

Churer Schriften  
zur Informationswissenschaft

Herausgegeben von  
Wolfgang Semar

---

Arbeitsbereich  
Informationswissenschaft

**Schrift 114**

**Einflüsse auf das Vertrauen und die  
Nutzerakzeptanz von Voice Commerce in  
der Schweiz**

**Livia Mosberger**

---

Chur 2020



# **Churer Schriften zur Informationswissenschaft**

Herausgegeben von Wolfgang Semar

Schrift 114

## **Einflüsse auf das Vertrauen und die Nutzerakzeptanz von Voice Commerce in der Schweiz**

**Livia Mosberger**

Diese Publikation entstand im Rahmen einer Thesis zum Bachelor of Science FHGR in Digital Business Management.

Referent: Philipp Liebreuz

Korreferent: Prof. Armando Schär

**Verlag:** Fachhochschule Graubünden

**ISSN:** 1660-945X

**Ort, Datum:** Chur, September 2020





## Abstract

Sprachassistenten sind in vielen Ländern bereits fest in den Alltag der Nutzer integriert und werden unter anderem dafür benutzt, Produkte einzukaufen. Dieser neue Absatzkanal bietet für Unternehmen spannende Möglichkeiten, aber auch Herausforderungen. In der Schweiz bietet sich diese Option nicht, da Sprachassistenten noch nicht weit verbreitet sind und die Möglichkeit, mit der Stimme einzukaufen, aktuell nicht besteht. Dies könnte sich in den nächsten Jahren jedoch schnell ändern, beispielsweise mit einem Eintritt von Amazon in den Schweizer Markt.

Bisher ist nicht bekannt, welche Faktoren einen Einfluss auf die Nutzerakzeptanz und das Vertrauen von Voice Commerce in der Schweiz haben. Ziel dieser wissenschaftlichen Arbeit war deshalb, diese Faktoren zu identifizieren. Das wurde durch eine ausführliche Literaturanalyse sowie durch eine empirische Befragung von Experten erreicht. Die wissenschaftliche Fragestellung, die als Basis der Arbeit diente, lautet:

*«Welche Faktoren beeinflussen das Vertrauen und die Nutzerakzeptanz von Voice Commerce und Voice-Commerce-Anwendungen in der Schweiz?»*

Die Ergebnisse aus der Literaturanalyse haben gezeigt, dass das Vertrauen in die Technologie Voice Assistants und in die Anwendungen von Voice Commerce massgeblich dadurch beeinflusst wird, wie gross ein wahrgenommener Nutzen ist. Wenn die Nutzung von Voice Commerce einen genügend grossen Vorteil bringt, werden auch mögliche Bedenken, beispielsweise im Bereich Datenschutz, zurückgestellt. Weiter ist entscheidend, wie die Nutzer informiert werden. Unternehmen müssen transparent kommunizieren und Wege finden, wichtige Informationen sinnvoll zu den Nutzern zu transportieren. Der Kaufprozess selbst kann durch einen ergänzenden Bildschirm informativer und damit vertrauenswürdiger gestaltet werden. Die technologische Weiterentwicklung und die Verbesserung der Spracherkennung sind weitere wichtige Einflussfaktoren auf die Nutzerakzeptanz. Je natürlicher die Interaktion ist, desto eher haben Nutzende Vertrauen in die Anwendung. Dabei spielt auch der Faktor Zeit eine Rolle, da technologische Entwicklung immer ein stetiger Prozess ist.

Die Befragung von Experten hat ergeben, dass die Schweiz sich nicht von anderen europäischen Ländern unterscheidet, was die Akzeptanz von Innovationen und neuen Technologien betrifft. Es herrscht Einigkeit darüber, dass sich Voice Assistants und Voice Commerce auch in der Schweiz durchsetzen werden.

## Inhaltsverzeichnis

1	Einleitung .....	1
1.1	Ausgangslage und Themenrelevanz .....	1
1.2	Themenabgrenzung und wissenschaftliche Fragestellung .....	3
1.3	Ziel der Arbeit .....	3
1.4	Aufbau der Arbeit .....	4
2	Theoretische Grundlagen .....	5
2.1	Begriffserklärung und -abgrenzung .....	5
2.2	Conversational Interfaces .....	5
2.2.1	Geschichte von Conversational Interfaces .....	6
2.2.2	Technische Grundlagen von Conversational Interfaces .....	8
2.3	Voice Assistants .....	9
2.4	Smart Speaker .....	9
2.4.1	Amazon .....	10
2.4.2	Google .....	11
2.5	Voice Commerce .....	12
2.5.1	Stand Voice Commerce .....	13
2.5.2	Kritische Betrachtung .....	16
2.5.3	Nichtnutzung von Voice Commerce .....	17
2.5.4	Stand Voice Commerce in der Schweiz .....	18
2.6	Nutzerakzeptanz und Vertrauen .....	19
2.6.1	Akzeptanz von Technologien .....	20
2.6.2	Vertrauen in digitalen Transaktionen .....	22
2.6.3	Entstehung von Vertrauen in Voice Commerce .....	25
3	Methode .....	31
3.1	Methodisches Vorgehen .....	31
3.1.1	Auswahl der Expertinnen und Experten .....	32
3.1.2	Interviewaufbau und -durchführung .....	33
3.1.3	Aufbau Interviewleitfaden .....	33
3.1.4	Interviewauswertung .....	34
4	Ergebnisse .....	35
4.1	Voice Assistants in der Schweiz .....	35
4.2	Potenzial von Voice Assistants in der Schweiz .....	35
4.3	Nutzerakzeptanz .....	36
4.3.1	Nutzen und Wert .....	36
4.3.2	Information und Kommunikation .....	36

4.3.3	Nutzerakzeptanz in der Schweiz .....	37
4.4	Voice Commerce .....	37
4.5	Voice Commerce in der Schweiz .....	38
5	Fazit .....	39
5.1	Einflüsse auf die Nutzerakzeptanz von Voice Commerce in der Schweiz .....	39
5.1.1	Nutzen .....	39
5.1.2	Information und Kommunikation .....	39
5.1.3	Technologische Entwicklung .....	40
5.2	Betrachtung der Schweiz .....	40
5.3	Beantwortung der Teilfragestellung .....	41
6	Diskussion und Reflexion .....	43
7	Quellenverzeichnis .....	45
8	Anhang .....	49
	Rechercheprotokoll .....	49
	Interviewleitfäden .....	50
	Interviewtranskripte .....	54
	Interviewauswertung .....	76

## Abbildungsverzeichnis

Abbildung 1: Hype Cycle for Emerging Technologies 2018 .....	2
Abbildung 2: Grundsätzliche Funktionsweise von Voice Assistants .....	9
Abbildung 3: Total Alexa Skills nach Ländern .....	11
Abbildung 4: Anzahl Smart Speaker bis 2018 in den USA nach Hersteller .....	12
Abbildung 5: Touchpoints in Voice Commerce .....	14
Abbildung 6: Wahrnehmung von Conversational Commerce .....	16
Abbildung 7: Verbesserungspotenzial von Voice Commerce .....	17
Abbildung 8: Technology Acceptance Model nach Davis .....	21

## Tabellenverzeichnis

Tabelle 1: Arten von vertrauensbildenden Diensten mit Beispielen.....	25
Tabelle 2: Angefragte Expertinnen und Experten.....	32

## Abkürzungsverzeichnis

AI	Artificial Intelligence
CI	Conversational Interface
GUI	Graphical User Interface
KI	Künstliche Intelligenz
NLP	Natural Language Processing
NLU	Natural Language Understanding
PC	Personal Computer
STT	Speech-to-Text
TAM	Technology Acceptance Model
TTS	Text-to-Speech
VA	Voice Assistant

# 1 Einleitung

Sie ist ein beliebtes Thema in Science-Fiction-Filmen: die Beziehung zwischen Mensch und Computer. Nicht selten sind die Computer dabei die Bösewichte und wollen nichts Anderes als das Ende der Menschheit. Bekanntes Beispiel dafür ist der Supercomputer HAL 9000 aus dem Film «2001: Odyssee im Weltraum», der sich gegen seine Crew richtet und versucht, diese komplett auszulöschen. Ein extremer Gegensatz dazu ist Samantha, die Künstliche Intelligenz (KI) im Film «Her» aus dem Jahr 2013. Auch in diesem Film ist die KI so weit fortgeschritten, dass die Menschen mit ihr so interagieren wie mit realen Personen. Beim Protagonisten Theodore geht diese Interaktion und das Vertrauen in die Technologie sogar so weit, dass er sich in Samantha verliebt und umgekehrt (Evans, 2015). In naher Zukunft ist keines der beiden Szenarien besonders realistisch. Doch die Technologie von Voice Assistant Systems hat sich in den vergangenen Jahren stark entwickelt und viele spannende Fragestellungen aufgeworfen. Eine dieser Fragestellungen ist die Grundlage dieser wissenschaftlichen Bachelorarbeit.

## 1.1 Ausgangslage und Themenrelevanz

Voice Assistant Systems, zu Deutsch Sprachassistenten genannt, sind das Oberthema dieser Bachelorarbeit. Mit der Einführung von Siri auf dem Apple iPhone im Jahr 2010 sind Sprachassistenten das erste Mal richtig im breiten Markt angekommen. Längst ist Apple aber nicht mehr der einzige Anbieter auf diesem Gebiet. Amazon mit Alexa, Google mit dem Google Assistant, Microsoft mit Cortana oder Samsung mit Bixby - sie alle stellen dem Nutzer einen Sprachassistenten zur Verfügung, der eine Nutzerinteraktion ohne Tastatur ermöglicht (Gentemann, 2018, S. 10). Potenzial sehen nicht nur die grossen Anbieter, auch auf dem Hype Cycle for Emerging Technologies von Gartner wird den Voice Assistants unter dem Oberbegriff «Virtual Assistants» eine erfolgreiche Zukunft prognostiziert. Die Technologie der Voice Assistants hat gemäss Abbildung 1 den grössten Hype bereits hinter sich und wird voraussichtlich in zwei bis fünf Jahren das Plateau der Produktivität erreichen. Daraus lässt sich schliessen, dass in den nächsten Jahren die Entwicklung der Voice Assistants weiterhin zügig voranschreiten wird.

Diese Entwicklung angestossen haben die Voice Assistants auf Smartphones. Sie waren der Wegbereiter für die sogenannten Smart Speaker, die an einem festen Standort installiert sind, meistens im eigenen Zuhause. Je nach Land sind diese smarten Lautsprecher unterschiedlich etabliert. So hatten in den USA im Jahr 2018 rund 26 Prozent der Erwachsenen Zugang zu einem Smart Speaker (Kinsella & Mutchler, 2019, S. 6). Und auch in Deutschland gab jeder vierte Erwachsene an, bis Ende 2018 einen

Smart Speaker besitzen zu wollen (Accenture, 2018). In der Schweiz hingegen sind Smart Speaker noch kaum verbreitet, da die beiden grössten Anbieter Amazon und Google die Lautsprecher in der Schweiz nicht offiziell verkaufen. Ob und wann ein Markteintritt bevorstehen könnte, ist unklar. Und auch wenn die bekanntesten Anbieter den Schweizer Markt erschliessen würden, bleibt offen, wie die smarten Lautsprecher bei den Schweizer Kunden ankommen und akzeptiert werden würden. Gerade diese Nutzerakzeptanz wird ausschlaggebend sein, ob eine Voice-Anwendung in der Schweiz erfolgreich etabliert werden kann.

## Hype Cycle for Emerging Technologies, 2018

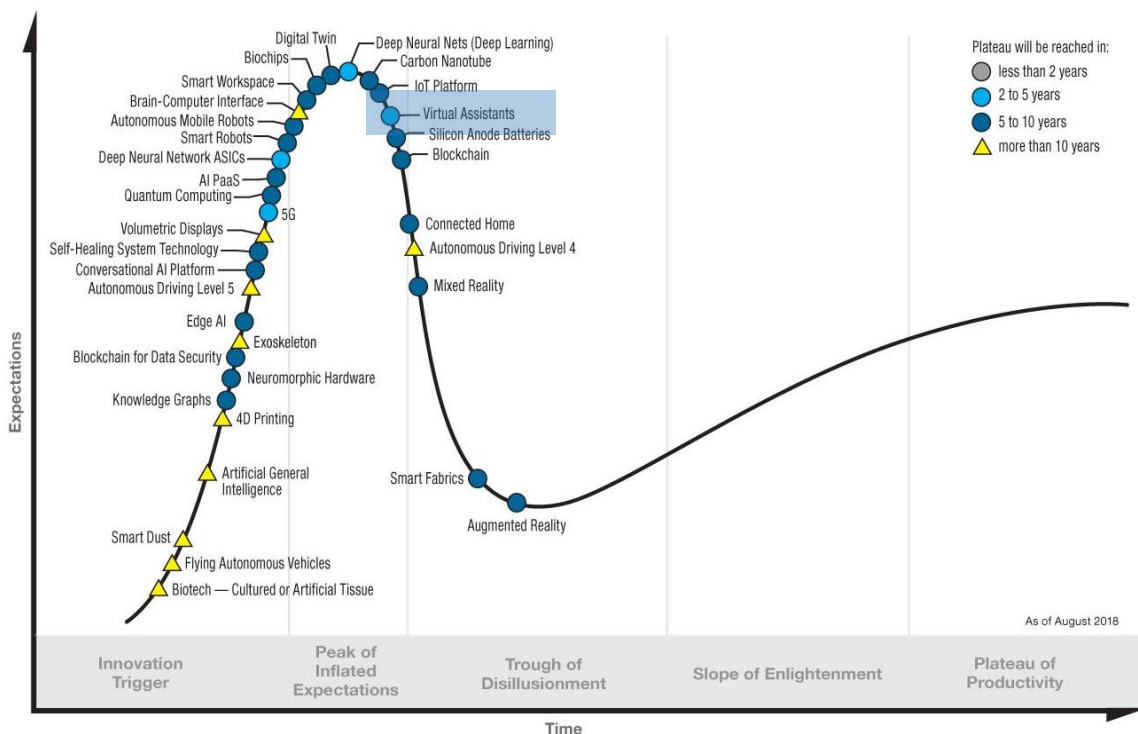


Abbildung 1: Hype Cycle for Emerging Technologies 2018

Quelle: Gartner (2018).

Ein für viele Unternehmen spannendes Feld im Bereich der Voice Assistants ist das sogenannte Voice Commerce. Voice Commerce ist ein Teil von E-Commerce, bei dem der Informationsaustausch sowie die Kommunikation zwischen dem Kunden und dem Anbieter im Stil einer natürlichen Konversation geführt wird (Kruse Brandão & Wolfram, 2018, S. 260). Voice Commerce bildet dabei alle Stadien des Einkaufsprozesses ab. Dieser neue Absatzkanal kann im Gegensatz zur Schweiz bereits in vielen Ländern, beispielsweise in Deutschland oder den USA, genutzt werden.

## 1.2 Themenabgrenzung und wissenschaftliche Fragestellung

Voice Assistants sind das Oberthema dieser Arbeit. Innerhalb des Voice Assistant Systems wird der Fokus auf Voice Commerce gelegt.

Um eine Anwendung im Voice Commerce-Bereich zu entwickeln und erfolgreich auf dem Markt zu etablieren, muss diese von vielen Personen genutzt werden. Wird der Anwendung oder der Technologie, auf der die Anwendung aufbaut, hingegen nicht vertraut und nur wenig Nutzerakzeptanz entgegengebracht, kann diese kaum Erfolg haben. Für Unternehmen ist das entscheidend, da sich die Entwicklung einer Voice Commerce-Anwendung sonst nicht lohnt. Der zweite Fokus dieser Arbeit liegt darum auf dem Vertrauen und der Nutzerakzeptanz dieser Anwendungen.

Der dritte Fokus, der in dieser Arbeit berücksichtigt wird, ist die geographische Eingrenzung auf die Schweiz. Grund dafür ist, dass Smart Speaker in der Schweiz zwar nicht offiziell verkauft werden, viele Unternehmen aber bereits über Anwendungen nachdenken. Eines dieser Unternehmen ist die Verlagsgruppe Tamedia AG, die diese Bachelorarbeit als Partnerin unterstützt. Für diese Unternehmen ist es wichtig zu wissen, wie die potenziellen Schweizer Kunden gegenüber Voice Commerce eingestellt sind.

Die übergeordnete wissenschaftliche Fragestellung, die sich aus diesen Einschränkungen ergibt, lautet:

*«Welche Faktoren beeinflussen das Vertrauen und die Nutzerakzeptanz von Voice Commerce und Voice-Commerce-Anwendungen in der Schweiz?»*

Zusätzlich zu dieser Fragestellung wurden im Verlauf der Arbeit drei weitere Teilfragestellungen aufgestellt, die am Ende dieser Arbeit beantwortet werden.

## 1.3 Ziel der Arbeit

Mit dieser Bachelorthesis soll wissenschaftlich erarbeitet werden, welche Faktoren das Vertrauen und die Nutzerakzeptanz von Voice Commerce in der Schweiz beeinflussen. Durch Literaturanalyse und Experteninterviews sollen ausserdem Eindrücke gewonnen werden, wie der aktuelle Stand von Voice Assistenten in der Schweiz ist. Im Bereich Voice in der Schweiz, speziell im Voice Commerce-Bereich, gibt es noch Wissenslücken. Es wird versucht, diese durch die folgende Bachelorthesis zu schliessen, indem die Bereiche Voice Commerce, Vertrauen in Technologie und Nutzerakzeptanz in der Schweiz miteinander verknüpft werden.

Diese Bachelorarbeit richtet sich an Personen, die sich für die Technologie der Voice Assistants interessieren und wissen möchten, wie das Vertrauen in diese zustande kommt. Ausserdem können die Ergebnisse für Personen interessant sein, die in einem Unternehmen arbeiten, welches eine Voice-Anwendung plant oder sich bereits in der Phase der Umsetzung befindet.

## **1.4 Aufbau der Arbeit**

In Kapitel 1 der vorliegenden Arbeit werden die Zielsetzung, die Abgrenzung des Themas und die Forschungsfragen behandelt. In Kapitel 2 werden die theoretischen Grundlagen zu Vertrauen in Voice Commerce erarbeitet sowie Begrifflichkeiten geklärt. Die durchgeführte Empirie zu Voice Commerce in der Schweiz wird in Kapitel 3 erläutert. Kapitel 4 befasst sich mit der Auswertung der Ergebnisse aus dieser empirischen Erhebung. Die gefundenen Einflussfaktoren werden in diesem Kapitel behandelt. Im 5. Kapitel werden die Ergebnisse aus Literatur und Experteninterviews kombiniert und ein Fazit daraus gezogen. Im letzten Kapitel werden die Ergebnisse diskutiert und die Arbeit reflektiert.



## 2 Theoretische Grundlagen

In diesem Kapitel werden die theoretischen Grundlagen dieser Arbeit vermittelt. Die folgenden Unterkapitel basieren auf einer Literaturrecherche und -auswertung.

### 2.1 Begriffserklärung und -abgrenzung

Die Begriffe, die im Zusammenhang mit Voice Assistenten oft verwendet werden, sind kaum einheitlich definiert. Smart Speaker, Sprachassistenten, Voice Assistant Systems, Digital Personal Assistant, Virtual (Personal) Assistant, Spoken Dialog Systems, Intelligent Virtual Assistant, Conversational Interfaces, und viele mehr, werden oft synonym verwendet, da auch die Forschung auf diesem Gebiet noch in einem frühen Stadium steht (Budzinski, Noskova & Zhang, 2018, S. 3). Eine klare Abgrenzung ist nicht möglich, denn sie alle definieren sich meist wie folgt: «With a conversational interface, people can speak to their smartphones and other smart devices in a natural way in order to obtain information, access Web services, issue commands, and engage in general chat» (McTear, Griol & Callejas, 2016, S. 11).

Je nach Publikation und Literatur wird ein Begriff verwendet, der im ganzen Text konsistent bleibt. In dieser Arbeit wird die Technologie Voice Assistants oder auf Deutsch Sprachassistenten genannt. Wenn von konkreten Geräten die Rede ist, werden diese Smart Speaker genannt.

### 2.2 Conversational Interfaces

«Conversational Interfaces sind innovative Benutzerschnittstellen, die Sprache, d. h. Dialog, in geschriebener oder gesprochener Form zur Kommunikation zwischen Menschen und Maschinen (Computer) verwenden» (Stanoevska-Slabeva, 2018, S. 26). Conversational Interfaces (CI) sind ein Paradigmenwechsel in der Interaktion zwischen Mensch und Maschine und der nächste Schritt nach dem Graphical User Interface (GUI) und dessen Vorgänger, dem Terminal Interface (Stanoevska-Slabeva & Lenz-Kesekamp, 2018). CI lassen sich nach der Dialogart in die zwei Kategorien Chatbots und Voice Assistants einteilen. Die geschriebene Form des Dialogs bezieht sich auf die Chatbots und wird in dieser Bachelorthesis nicht berücksichtigt. CI ist somit der Überbegriff aller oben genannten Begrifflichkeiten. Der Begriff CI ist für diese Bachelorthesis relevant, da viele Publikationen sich nicht nur explizit mit Voice Assistants befassen, sondern allgemein mit CI und somit auch die Interaktion mit Chatbots untersucht wurde.

### 2.2.1 Geschichte von Conversational Interfaces

Als Schlüsseltechnologie im Transformationsprozess vom GUI zum CI und somit zu Voice Assistants gilt die künstliche Intelligenz. Diese wurde bereits in den 1950er-Jahren vom Mathematiker Alan Turing erforscht. Der von ihm entwickelte Turing-Test gilt als bestanden, wenn ein Mensch nicht mehr unterscheiden kann, ob er gerade mit einem anderen Menschen oder mit einer Maschine interagiert (Sieber, 2019, S. 12). Doch nicht nur die künstliche Intelligenz wurde zu diesem Zeitpunkt erforscht, auch Sprachentwicklungssoftware, die auf künstlicher Intelligenz aufbaut, hat ihren Ursprung in den 50ern. Bereits 1952 konnte ein Prototyp von Bell Laboratories namens Audrey gesprochene Zahlen erkennen (Kruse Brandão & Wolfram, 2018, S. 262).

Zehn Jahre später präsentierte IBM den Rechner Shoebox, der 16 Worte verstehen konnte, davon die Ziffern 0 bis 9 sowie mathematische Funktionen wie «plus» und «minus». So konnten dem Rechner einfache Rechenaufgaben gestellt werden (Kruse Brandão & Wolfram, 2018, S. 262).

In den 70ern und 80ern wurde der Wortschatz der Sprachsoftware stetig vergrößert. Das System Harpy, entwickelt von der Carnegie-Mellon-Universität, hatte beispielsweise einen Wortschatz von rund tausend Wörtern. Die nächste Entwicklungsstufe wurde mit dem Hidden Markov Modell erreicht. Dieses konnte ganze Wortfolgen analysieren und statistisch berechnen, welches Wort auf ein anderes folgt (Kruse Brandão & Wolfram, 2018, S. 262).

Mit dem Aufkommen der Personal Computer (PC) haben viele Unternehmen parallel Sprachsoftware entwickelt, die über den PC bedient werden konnte. Ein wichtiger Meilenstein in der Entwicklung der künstlichen Intelligenz (KI) war 1996, als ein von IBM entwickelter Computer namens Deep Blue einen Schachweltmeister besiegen konnte (McTear et al., 2016, S. 16). Das wirkte sich auch auf die Weiterentwicklung von CI aus.

Ab den 2010er-Jahren stieg die Entwicklung von Conversational Interfaces exponentiell an (Stanoevska-Slabeva & Lenz-Kesekamp, 2018). Mit der Einführung von Siri auf dem iPhone 4s im Jahr 2011 wurde ein Sprachassistent das erste Mal der breiten Masse zugänglich. Obwohl Siri als Beta-Version eingeführt wurde, funktionierte sie schon sehr gut und konnte den Nutzer tatsächlich unterstützen, da sie auf die auf dem iPhone installierten Apps zugreifen und beispielsweise Anrufe starten konnte. Als erster Sprachassistent wirkte Siri tatsächlich menschlich, indem sie Humor zeigen und auch auf Scherzfragen gute Antworten finden kann (Martin, 2019). Beinahe zeitgleich lancierte Google den Sprachassistenten Google Now (seit 2016 Google Assistant genannt), der vor allem als Suchdienst angewendet wurde. Im Jahr 2014 kam Microsoft mit Cortana

auf den Markt, und auch Amazon lancierte mit Amazon Alexa eine Spracherkennungssoftware im gleichen Jahr (Kruse Brandão & Wolfram, 2018, S. 263). Im Gegensatz zu Konkurrent Google lieferte Amazon mit Alexa von Anfang an eine Hardware-Lösung, den Amazon Echo, den es heute in verschiedenen Ausführungen gibt. Ein weiterer Vorteil von Amazon waren die Synergien, die gekonnt genutzt wurden und so das eigene Produkt- und Dienstleistungsangebot noch näher an den Kunden brachten (Martin, 2019).

Nebst den amerikanischen Unternehmen haben auch viele asiatische Unternehmen eigene Sprachsoftware entwickelt. Bixby wurde 2017 von Samsung lanciert, die russische Suchmaschine Yandex etablierte Alice-KI im gleichen Jahr, genauso wie das chinesische Amazon-Pendant Alibaba mit AliGenie. Ein weiteres chinesisches Unternehmen, Xiaomi, ist seit 2018 mit seiner sogenannten Mi AI auf dem Sprachassistentenmarkt vertreten (Martin, 2019). Im Gegensatz zu Amazon oder Google haben diese Unternehmen den einfacheren und günstigeren Weg gewählt, um die Software zu lancieren, nämlich über die bestehende Hardware der Smartphones (Martin, 2019).

Im Jahr 2019 sind Voice Assistants soweit ausgereift, dass sie rund 95 Prozent der (englischsprachigen) Anfragen verstehen und korrekt ausführen können (Capgemini Digital Transformation Institute, 2018, S. 3). Und mit Google Duplex, der neusten Artificial Intelligence (AI) von Google, konnte während einer Live-Präsentation ein Coiffeur-Termin vereinbart werden, ohne dass die Gesprächspartnerin realisierte, dass sie nicht mit einem Menschen, sondern mit der AI sprach. Es gibt sogar Stimmen, die behaupten, dass Google damit den Turing-Test bestanden hat, auch wenn die Voraussetzungen für den Turing-Test nicht zu hundert Prozent erfüllt waren (Oppermann, 2018).

Dass sich die Technologie der Voice Assistants stark weiterentwickeln wird, steht ausser Frage. Interessanter ist hingegen, in welchem Tempo das geschieht, da die Systeme durch die immer grösser werdende Datenmenge stets dazulernen und sich verbessern. Oder wie es McTear et al. (2016) beschreiben:

«As a result of these technological advances, user acceptance of technologies such as the conversational interface has increased, leading to increased adoption and consequently producing more data from which systems can learn, in turn resulting in further improvements in the technology.» (S. 22)

Daraus wird die folgende Teilfragestellung abgeleitet, die zusätzlich zur Hauptfragestellung beantwortet wird:

*«Je besser die Technologie der Conversational Interfaces in Zukunft wird, desto mehr werden die Nutzer die Technologie akzeptieren und in ihre Anwendungen vertrauen.»*

## **2.2.2 Technische Grundlagen von Conversational Interfaces**

Die Technologie, auf der Conversational Interfaces aufbauen, nennt sich Natural Language Processing (NLP). NLP befasst sich mit der Verarbeitung der natürlichen Sprache, dem Verstehen sowie der Semantik von Wörtern und Sätzen, der Klassifizierung von Texten, der korrekten Aussprache und Betonung sowie der Syntaxanalyse und der Beantwortung von Fragen (ITWissen.info, 2019). NLP ermöglicht die Interaktion und Kommunikation zwischen Mensch und Maschine in einer natürlichen Sprache. Durch NLP-Algorithmen können Computer die menschliche Sprache und deren Inhalt erkennen und beispielsweise Anweisungen ausführen (ITWissen.info, 2019). NLP hängt eng mit Natural Language Understanding (NLU) zusammen. Das Ziel von NLU ist, den Kontext, Zusammenhänge und den Sinn einer Anfrage zu verstehen und aufgrund dessen die möglichen Antworten zu erkennen (Stanoevska-Slabeva, 2018, S. 29). Damit NLU und NLP funktionieren, muss eine Spracheingabe zuerst umgewandelt werden. Das Verfahren nennt sich Speech-to-Text (STT). Der Gegenpol dazu, sprich die Konversion von Text in die natürliche Sprache, nennt sich Text-to-Speech (TTS). TTS ist dafür verantwortlich, dass die Antwort auch wieder in einer Sprache an den Nutzer ausgegeben wird (Stanoevska-Slabeva & Lenz-Kesekamp, 2018, S. 2–3).

Eine Anfrage eines Nutzers wird also zuerst vom Mikrofon aufgezeichnet und über das Internet an die Server des entsprechenden Voice Assistants gesendet. Dort werden die Töne mittels STT konvertiert. Mit NLU und NLP wird der Zusammenhang und der Sinn der Frage erkannt und verarbeitet. Die Antwort, die in Textform erstellt wird, wird anschliessend mit TTS wieder in die natürliche Sprache konvertiert und an den Nutzer ausgegeben (Kruse Brandão & Wolfram, 2018, S. 264). Bei CI, die ohne Sprachinput und -output funktionieren (Chatbots), fallen die Konversionen STT und TTS weg. In Abbildung 2 ist eine vereinfachte Funktionsweise einer Voice-Anfrage dargestellt.

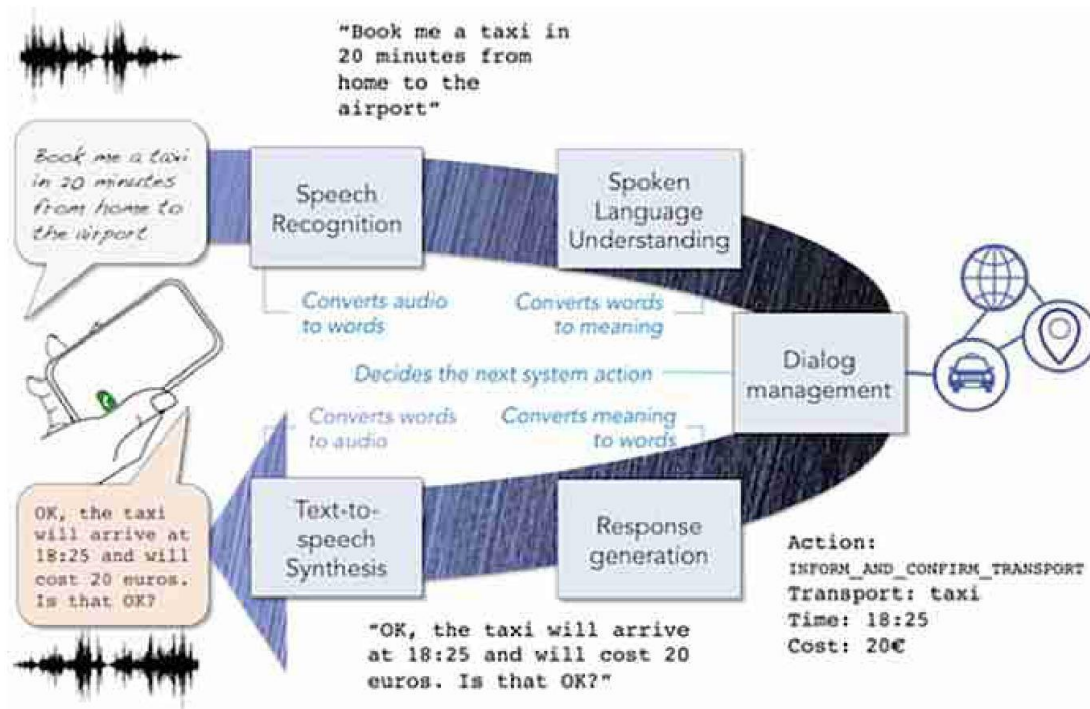


Abbildung 2: Grundsätzliche Funktionsweise von Voice Assistants

Quelle: McTear et al. (2016, S. 21).

## 2.3 Voice Assistants

Voice Assistants (VA) sind der Teilbereich von Conversational Interfaces, die in dieser Arbeit eingehender betrachtet werden. Als Definition von VA orientiert sich die Verfasserin an der Definition von TechTarget (2017), da diese eine sehr hohe Aktualität aufweist.

A voice assistant is a digital assistant that uses voice recognition, natural language processing and speech synthesis to provide aid to users through phones and voice recognition applications. [...] Voice assistants are built on artificial intelligence (AI), machine learning and voice recognition technology. As the end user interacts with the digital assistant, the AI programming uses sophisticated algorithms to learn from data input and better itself at predicting the user's needs.

Unter Voice Assistants wird also die Software verstanden, mit der User interagieren.

## 2.4 Smart Speaker

Wie im Kapitel «Geschichte von Conversational Interfaces» bereits beleuchtet, wurden VA vor allem durch das Aufkommen von Smartphones der breiten Masse zugänglich.

Das zweite Anwendungsfeld, in dem Voice Assistants heute vor allem genutzt werden, sind die sogenannten smarten Lautsprecher oder Smart Speaker. Viele Anwendungsgebiete sind identisch mit VA auf dem Smartphone. So lassen sich beispielsweise mit dem Smartphone-Sprachassistenten wie auch mit einem Smart Speaker Fragen beantworten, einen Timer stellen oder Termine abfragen. Viele Sprachassistenten auf dem Smartphone sind jedoch geschlossene Systeme, beispielsweise Siri von Apple. Die bekanntesten und am häufigsten verbreiteten Smart Speaker hingegen, nämlich der Amazon Echo und der Google Home, bieten jeweils eine dazugehörige Plattform an. Dadurch erst wurde das Thema für viele Unternehmen interessant, denn durch diese Plattformökonomie haben auch Firmen die Möglichkeit erhalten, Kunden über diesen neuen Kanal zu erreichen und eigene Funktionen oder Dienstleistungen über die Smart Speaker anzubieten (Elster, 2019). Aufgrund dieser Plattformökonomie liegt der Fokus dieser Arbeit primär auf den Sprachassistenten, die mit einem stationären Smart Speaker genutzt werden. Nachfolgend werden die wichtigsten Smart-Speaker-Anbieter vorgestellt.

#### **2.4.1 Amazon**

Die Smart Speaker, die vom Online-Versandhändler Amazon hergestellt werden, werden Echo genannt. Es gibt sie in verschiedenen Ausführungen und Grössen. Auch Modelle mit einem Bildschirm sind erhältlich und ergänzen die akustische Bedienung somit mit einer visuellen Komponente (Amazon, 2018). Die Sprachassistenten-Software von Amazon wird Alexa genannt. Alexa ist auch das Aktivierungswort für einen Amazon Echo. Erst nach der Nennung dieses Wortes kann der Smart Speaker bedient werden. (Lenz-Kesekamp & Weber, 2018, S. 20) Wie eingangs erwähnt, ist Alexa nach einem Plattform-Prinzip aufgebaut. Vergleichbar mit dem Apple App Store bietet Amazon einen kostenlosen Store an, auf dem die Kunden neue Alexa Sprachapplikationen, sogenannte Skills, aktivieren und hinzufügen können. Die Parallelen zum App Store gehen aber noch weiter, denn auch Unternehmen können eigene Skills entwickeln. Unterstützung bietet Amazon dabei mit dem Amazon Skill Kit und den Amazon Voice Services (Lenz-Kesekamp & Weber, 2018, S. 20). Der Kreativität sind dabei kaum Grenzen gesetzt. So können Personen in Deutschland beispielsweise einen Nachrichtenüberblick in 100 Sekunden erhalten, den nächsten Zug von A nach B abfragen oder das Fernsehprogramm des Abends erfahren (Freist, Dirscherl & Joos, 2019). Die Möglichkeit, eigene Alexa Skills zu entwickeln, ist aber nicht auf Unternehmen limitiert. Grundsätzlich kann jede Person eigene Alexa-Befehle entwickeln und diese nutzen. Diese Plattformökonomie von Amazon hat dazu geführt, dass die Anzahl an Skills ständig

wächst. Im Jahr 2018 hat sich die Anzahl Amazon Skills in den USA verdoppelt, und auch in Deutschland nahmen die Skills um 152 Prozent zu (Kinsella, 2019a).

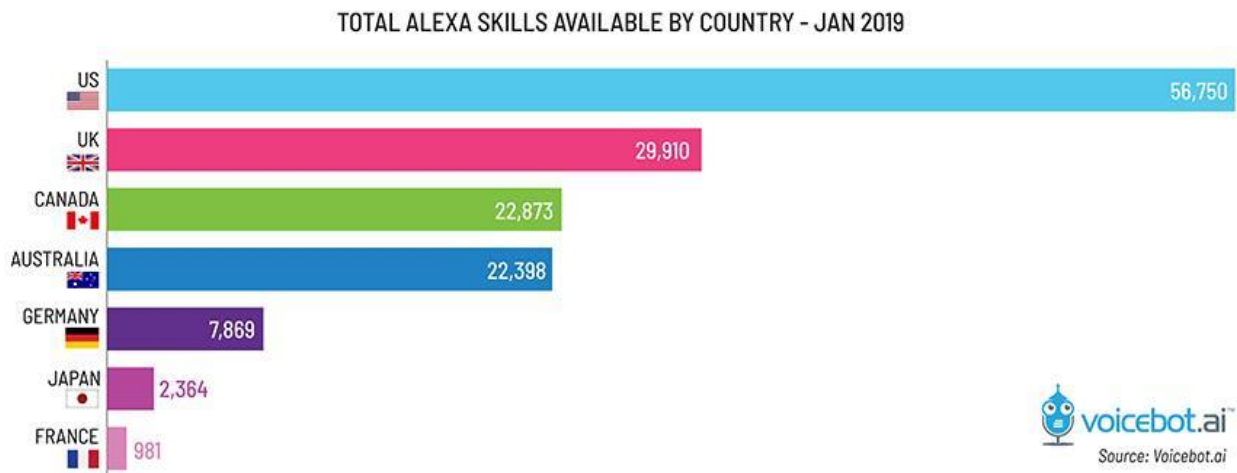


Abbildung 3: Total Alexa Skills nach Ländern

Quelle: Kinsella (2019a).

Amazon gilt in den USA wie auch in Deutschland als Marktführer von Smart Speakern. 61.1 Prozent der amerikanischen Smart Speaker sind von Hersteller Amazon (Kinsella & Mutchler, 2019, S. 7). Und auch in Deutschland besitzen 71 Prozent der Personen mit Smart Speaker ein Gerät von Amazon (Statista, 2019, S. 14).

Amazon bietet Alexa und die dazugehörige Infrastruktur auch anderen Hardwareherstellern an, die diese in ihre eigenen Geräte einbauen können. Auch das trägt zur weiten Verbreitung von Amazon Alexa bei (Kruse Brandão & Wolfram, 2018, S. 270). So findet sich Alexa beispielsweise auf Geräten von Sonos, Element oder Lenovo (Deloitte, 2018, S. 20).

## 2.4.2 Google

Google brachte seinen Smart Speaker Google Home im Jahr 2016 auf den Markt (Budzinski et al., 2018, S. 5) Der Sprachassistent Google Assistant, der sich im Google Home befindet, ist die Weiterentwicklung von Google Now aus dem Jahr 2011 (Kruse Brandão & Wolfram, 2018, S. 263). Der Google Assistant ist sehr verbreitet, da er auf jedem neueren Android-Smartphone vorinstalliert ist. Der Aktivierungsbefehl für den Google Assistant lautet «Ok Google» (McTear et al., 2016, S. 12). Wie auch Amazon hat Google eine Plattform, über die Unternehmen wie auch Privatpersonen eigene Applikationen hinzufügen können. Bei Google werden diese «Actions» genannt (Kruse Brandão & Wolfram, 2018, S. 268).

Der Google Assistant ist im Vergleich zu Konkurrent Amazon sehr stark in Wissensfragen und kann auch mit komplexen Fragestellungen gut umgehen. Das ist auf die enorme Wissensbasis der Suchmaschine Google zurückzuführen, auf die der Assistant direkt zugreift (Kruse Brandão & Wolfram, 2018, S. 276).

Die Smart Speaker von Google sind nach Amazon Echo am häufigsten verbreitet, wobei die Verbreitung des Google Assistants ein stärkeres Wachstum aufweist als Amazon Alexa. Im ersten Quartal 2019 waren es immerhin 19 Prozent der deutschen Smart Speaker-Nutzer, die eines oder mehrere Geräte von Google besaßen (Statista, 2019).

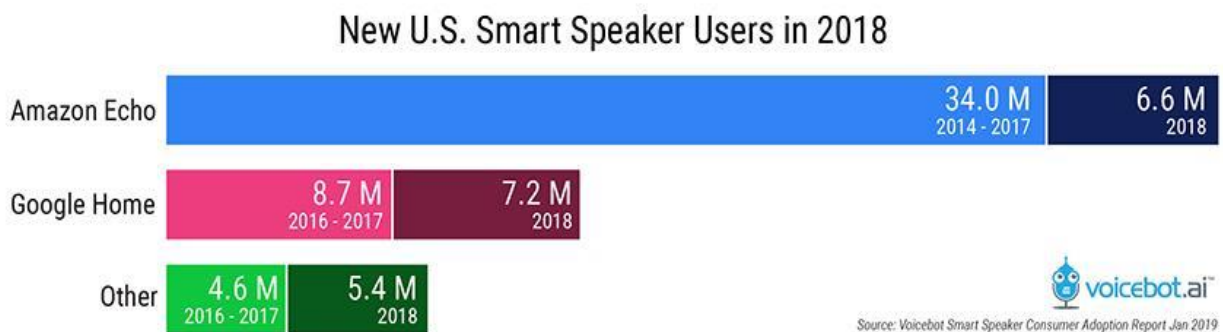


Abbildung 4: Anzahl Smart Speaker bis 2018 in den USA nach Hersteller

Quelle: Kinsella (2019b).

## 2.5 Voice Commerce

Die aktuell häufigsten Aktionen, die mit Smart Speakern gemacht werden, sind Musik abspielen, das Wetter abfragen und Nachrichten anhören (Hayllar & Coode, 2018, S. 6). Für Unternehmen bietet sich mit Voice aber auch die Möglichkeit, ihre Produkte durch diesen neuen Touchpoint zu vertreiben. Interessant ist das sogenannte «Voice Commerce» – also das Bestellen via Spracheingabe – vor allem für Betreiber von Onlineshops. (Baumgartner & Dalla Vecchia, 2017, S. 30)

Voice Commerce und Conversational Commerce werden oft synonym verwendet, da die Begriffe noch sehr jung sind und noch nicht lange auf diesem Gebiet geforscht wird. Einer der ersten, der von Conversational Commerce sprach und als Erfinder dieses Begriffs gilt, war Chris Messina, ein Entwickler von Uber. Im Jahr 2016 definiert er Conversational Commerce folgendermassen (Messina, 2016):

Conversational Commerce pertains to utilizing chat, messaging, or other natural language interfaces (i.e. voice) to interact with people, brands, or services and bots that have had no real place in the bidirectional, asynchronous messaging context.



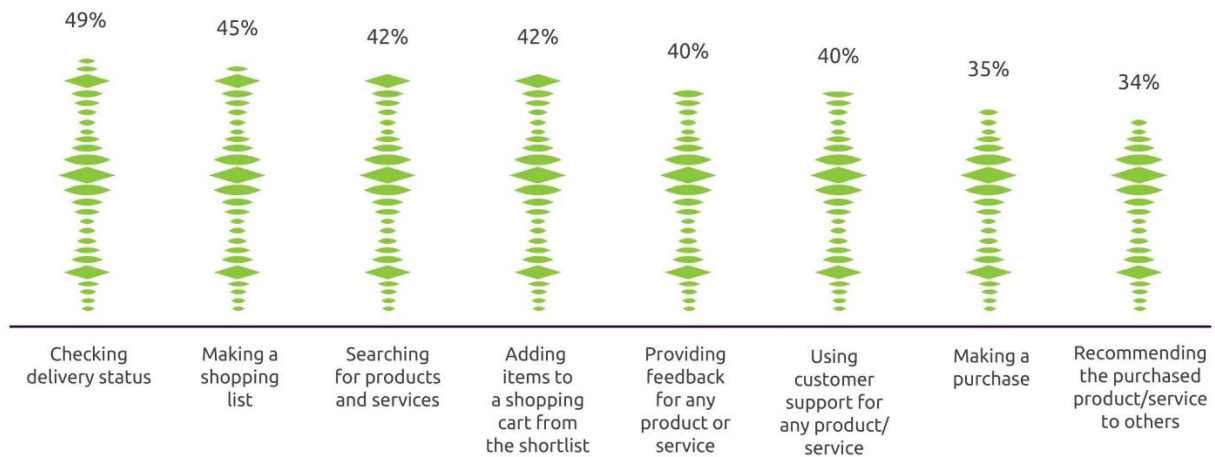
Eine weitere Definition von Conversational Commerce, die aber eigentlich eher Voice Commerce definiert, stammt aus einem Report von Capgemini: «...consumer purchase of products and services via voice assistants such as Google Assistant, Amazon's Alexa, and Apple's Siri...» (2018, S. 2). Und Eeuwen (2017) definiert Conversational Commerce schlicht als die Nutzung von Conversational Interfaces wie Chatbots oder Voice Assistants zur kommerziellen Nutzung im E-Commerce (S. 3). Im weiteren Verlauf dieser Arbeit wird jeweils der Begriff «Voice Commerce» verwendet, da dieser im Gegensatz zu Conversational Commerce die textbasierten Interfaces wie Chatbots nicht miteinschliesst.

### 2.5.1 Stand Voice Commerce

Voice Commerce gilt als die nächste grosse Disruption im E-Commerce. Das Beratungsunternehmen OC&C schätzt, dass in den USA bis 2022 40 Milliarden US-Dollar via Voice Commerce ausgegeben werden. Das entspräche rund sechs Prozent der gesamten E-Commerce-Ausgaben (Hayllar & Coode, 2018, S. 2). Voice Commerce umfasst nicht nur das Einkaufen von Produkten wie Lebensmittel oder Elektronikartikel, sondern auch das Bestellen von zubereitetem Essen, das Buchen eines Taxis oder Uber, und in einigen Ländern ist sogar das Senden von Geld möglich, was auch zu Voice Commerce gezählt wird (Capgemini Digital Transformation Institute, 2018, S. 6). Ausserdem gehört nicht nur der reine Kauf zu Voice Commerce dazu. Kundinnen und Kunden nutzen Voice bereits vor, während und nach dem Kauf (Narvar, 2018, S. 14). Es beginnt also mit der Produkt- und Informationssuche, geht weiter über das Hinzufügen zum Warenkorb bis zur definitiven Kaufentscheidung. Nach dem Kauf ist es je nach Anbieter möglich, den Sendungsstatus via Voice abzufragen, das Produkt zu bewerten, den Support zu kontaktieren oder eine Nachbestellung zu tätigen (Capgemini Digital Transformation Institute, 2018, S. 27; Narvar, 2018, S. 14). Durch diese vielen Touchpoints ist es nicht ganz einfach zu beurteilen, wie sehr Voice Commerce ein Kundenbedürfnis darstellt. In Abbildung 5 wurden Personen befragt, bei welchen Touchpoints sie sich vorstellen könnten, mit Voice Assistants zu interagieren. Die Umfrage zeigt deutlich, dass nicht alle Schritte im Verkaufsprozess via Voice gleich akzeptiert sind.

Trotzdem sind die meisten Kunden gegenüber Voice Commerce nicht komplett abgeneigt. 35 Prozent der Personen, die einen Voice Assistant benutzen, haben bereits mehrmals oder sogar regelmässig Konsumgüter gekauft. Der Anteil der Personen, die schon öfter Mahlzeiten via Voice bestellt haben, liegt mit 34 Prozent nur knapp darunter

(Capgemini Digital Transformation Institute, 2018, S. 6). Diese Zahlen zeigen, dass Voice Commerce zwar genutzt wird, es lässt sich aber noch deutliches Potenzial erkennen.



**Source:** Capgemini Digital Transformation Institute, Conversational Commerce Survey, October–November 2017, N = 5,041 consumers in the US, UK, France, and Germany.

Abbildung 5: Touchpoints in Voice Commerce

Quelle: Capgemini Digital Transformation Institute (2018, S. 27).

Grundsätzlich sind bei Voice Commerce kaum Grenzen gesetzt, welche Produkte über diesen Kanal vertrieben werden. In der Praxis sieht es aber bislang noch anders aus. Produkte, die momentan via Smart Speaker eingekauft werden, befinden sich in einem tiefen Preissegment und sind vor allem Alltagsprodukte (Baumgartner & Dalla Vecchia, 2017, S. 31). Dass vor allem günstige Produkte via Voice eingekauft werden, wird durch den durchschnittlichen Einkaufswert via Voice sichtbar. In rund zwei Dritteln der getätigten Einkäufe liegt dieser unter fünfzig US-Dollar (Kinsella & Mutchler, 2018, S. 23). Ein gutes Beispiel für ein beliebtes Voice-Commerce-Produkt sind Windeln. Diese sind im tiefen Preissegment, bei der entsprechenden Zielgruppe ein Alltagsgegenstand und werden in regelmässigen Abständen eingekauft. Ausserdem müssen bei einem Produkt wie Windeln keine grossen Informationen und Erklärungen eingeholt werden, da es sich um ein sehr einfaches und austauschbares Produkt handelt. Die Produkte, die bevorzugt werden, sind also aktuell solche, bei denen die Kundinnen und Kunden schon im Voraus ganz genau wissen, was sie kaufen wollen (Hayllar & Coode, 2018, S. 8). Das trifft natürlich auch auf Produkte zu, die erneut eingekauft werden. Es erstaunt darum nicht, dass auch das Nachbestellen eines Produkts ein beliebter Use Case ist (Narvar, 2018, S. 16). Nebst den Artikeln des täglichen Bedarfs besteht das grösste Interesse in den Produktgruppen Consumer Electronics und Lebensmittel (Sievers, 2019, S. 4).

Um den Stand von Voice Commerce beurteilen zu können, muss auch ein Blick auf die Anbieter von Voice-Commerce-Applikationen geworfen werden. Der naheliegendste Anbieter ist der Onlinehändler Amazon. Die eigene Sprachassistentin Alexa ist direkt mit dem Amazon-Shop verknüpft (Baumgartner & Dalla Vecchia, 2017, S. 30). Zu Beginn konnten Amazon-Prime-Kunden lediglich frühere Bestellungen nachbestellen, seit Längerem können aber auch neue Käufe getätigt werden. Amazon unternimmt einiges, um die Kunden zu Voice-Käufen zu animieren. So gibt es beispielsweise sogenannte Daily Deals für Voice-Shopper, bei denen eine kleine Auswahl an Produkten vergünstigt eingekauft werden kann (Kinsella & Mutchler, 2018, S. 28). Auch wenn Konkurrent Google keinen eigenen verknüpften Store hat, ist Voice Commerce durch den Google Assistant problemlos möglich. Die Funktion nennt sich Google Express und ist nicht nur auf dem Google Home, sondern auch via Smartphone abrufbar (Kinsella & Mutchler, 2018, S. 28). Google setzt stark auf Kooperationen, beispielsweise mit den US-amerikanischen Handelskonzernen Walmart, Target und Costco (Bruhn & Hadwich, 2018, S. 97). Mit dieser Strategie hat Google im Jahr 2018 sogar erreicht, dass mehr Personen die Voice Commerce-Funktion ausprobiert haben als bei Konkurrent Amazon (Kinsella & Mutchler, 2018, S. 28).

Die naheliegendste Form für Unternehmen, ihren Kundinnen und Kunden eine Voice Commerce-Applikation zur Verfügung zu stellen, ist aber nach wie vor via Alexa Skill oder Google Action. Vor allem in den USA machen das schon viele Unternehmen, vor allem bei Suchanfragen in der Nähe wird ein Potenzial erkannt. Grosse Gastronomie- und Convenience-Food-Ketten wie Dunkin' Donuts, Starbucks oder Domino's bieten ihren Kundinnen und Kunden die Möglichkeit, eine Bestellung via Alexa-Skill zu tätigen (Capgemini Digital Transformation Institute, 2018, S. 8). Es erstaunt darum nicht, dass der Befehl «Alexa, bestelle mir eine Cheese Pizza» einer der häufigsten Befehle ist (Langer, 2019). Alexa lernt auch schnell dazu und merkt sich, von welchem Anbieter man die Pizza denn nun am liebsten hat. Auch das verbessert das Nutzererlebnis wiederum (Langer, 2019).

In der Abbildung 6 bieten Tuzovic und Paluch (2018, S. 91) eine gute Zusammenfassung zum Stand von Conversational Commerce und somit Voice Commerce. Die Darstellung ist das Ergebnis ihrer Studie zur Kundenwahrnehmung von Conversational Commerce und zeigt die Herausforderungen, die Chancen und die Auswirkungen auf den Kaufprozess aus. Auch diese Studie kommt zum Ergebnis, dass die grosse Chance von Voice Commerce der Nutzen ist, der durch die bequeme, zeitsparende Art des Einkaufens generiert wird, aber auch durch die persönlichen Vorschläge, die dem Nutzer aufgrund seiner Daten und Präferenzen gemacht werden. Auf die Herausforderungen,

die sie in der Studie herausgefunden haben, wird im folgenden Kapitel genauer eingegangen.

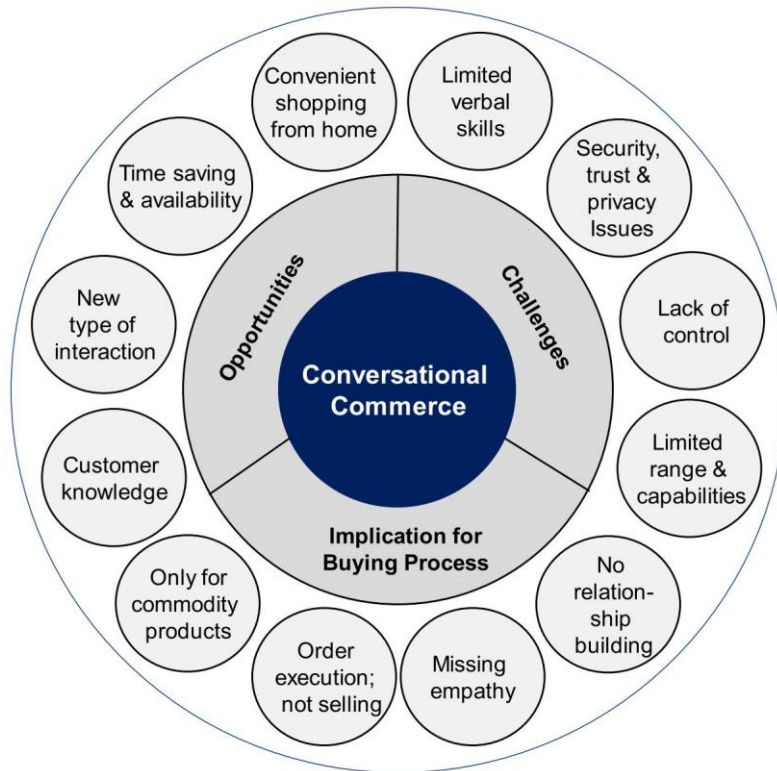


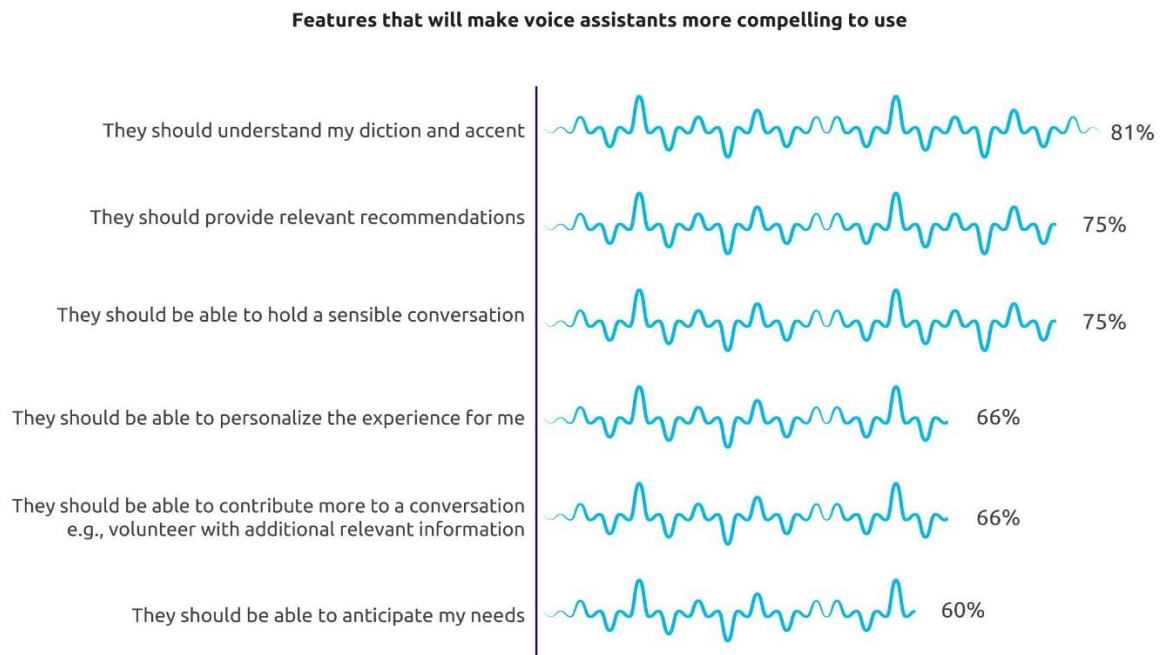
Abbildung 6: Wahrnehmung von Conversational Commerce

Quelle: Tuzovic & Paluch, (2018, S. 91).

## 2.5.2 Kritische Betrachtung

Trotz der vielen positiven Berichte: Voice Commerce ist noch nicht massentauglich in Europa, beispielsweise in Deutschland. Das liegt vor allem daran, dass es generell noch eine sehr junge Technologie ist und mehr als Trend denn als grosse Disruption wahrgenommen wird. In Deutschland sind die Unternehmen, die bereits Voice-Commerce-Applikationen anbieten, noch kaum über den Testbetrieb hinaus (Weidemann, 2019). Und auch in den USA liegt Voice Commerce gemäss den aktuellsten Reports deutlich hinter den Erwartungen zurück, da die Spracherkennung zu wenig zuverlässig funktioniert und sich darum nur wenige Produkte wirklich einfach via Spracheingabe einkaufen lassen (Pardes, 2019). Es ist also abzuwarten, in welche Richtung sich Voice Commerce entwickeln wird und ob auch die Kunden die Angebote über das reine Ausprobieren hinaus verwenden. Bis Voice Commerce einen fixen Platz im Alltag von deutschen oder Schweizer Kundinnen und Kunden haben wird, muss

primär die Spracherkennung verbessert werden, vor allem bei komplexen Anfragen, die über das reine Nachbestellen hinaus gehen (Weidemann, 2019).



**Source:** Capgemini Digital Transformation Institute, Conversational Commerce Survey, October–November 2017, N = 2,558 users in the US, UK, France, and Germany.

Abbildung 7: Verbesserungspotenzial von Voice Commerce

Quelle: Capgemini Digital Transformation Institute (2018, S. 16).

### 2.5.3 Nichtnutzung von Voice Commerce

In dieser kritischen Betrachtung von Voice Commerce dürfen die Nichtnutzer von Voice Assistants oder Voice Commerce nicht aussen vorgelassen werden. Denn die Personen, die sich bewusst oder auch unbewusst nicht mit Voice Assistants auseinandersetzen, wurden in dieser wissenschaftlichen Arbeit bisher kaum berücksichtigt. Dabei gibt es durchaus plausible Gründe, sich gegen eine Nutzung von Voice Commerce zu entscheiden. Der mit Abstand am häufigsten genannte Grund sind die Bedenken bezüglich Datenschutz und Datensicherheit (Capgemini Digital Transformation Institute, 2018, S. 17; EARSandEYES GmbH, 2018; Gentemann, 2018, S. 14). In die gleiche Richtung geht auch die Angst, dass man von Dritten abgehört wird. Weitere wichtige Gründe, die Nichtnutzer von der Nutzung abhalten und in beinahe allen Umfragen und Reports als häufigste Ablehnungsgründe auftauchen, sind die, dass schlicht ein zu geringer Nutzen erkennbar ist, sich keine Vorteile ergeben und das Sprechen mit einer künstlichen Stimme als unangenehm empfunden wird (Capgemini Digital Transformation Institute, 2018, S. 18). Diese genannten Gründe treffen generell auf Voice Assistants zu.

Wer also aus einem solchen Grund keinen Voice Assistant nutzt, wird auch nicht damit einkaufen wollen. Doch auch im Bereich des Voice Commerce gibt es Ablehnungsgründe. Das bedeutet also, dass es Personen gibt, die einen VA nutzen, jedoch keine Einkäufe damit tätigen wollen. Die Gründe dieser Personen unterscheiden sich von denen, die grundsätzlich nicht mit VA interagieren wollen. Der am häufigsten genannte Grund ist der, dass sich Konsumenten schlicht nicht wohlfühlen. Direkt dahinter ist das fehlende Vertrauen bezüglich den Bezahlinformationen sowie der fehlende Screen (Kinsella & Mutchler, 2018, S. 22). Gerade aus der Aussage mit dem fehlenden Screen lässt sich schliessen, dass die Informationen, die alleine via Voice vermittelt werden, oft nicht ausreichen, um den Kundinnen und Kunden ein vertrauensvolles Gefühl zu geben.

#### **2.5.4 Stand Voice Commerce in der Schweiz**

Die Nutzung von Voice Commerce ist in der Schweiz nicht möglich. Dafür ist zum einen verantwortlich, dass man die Smart Speaker von Amazon und Google in der Schweiz nicht offiziell erwerben kann. Über das Web kann man die Geräte zwar bestellen, viele ortsgebundene Funktionen stehen aber nicht zur Verfügung. Zum anderen liegt die verunmöglichte Nutzung daran, dass es in der Schweiz auch kaum Voice-Applikationen gibt, und die wenigen, die es gibt, sind nicht im Bereich Voice Commerce. Eine Anwendung, die sich noch am ehesten in den Bereich von Voice Commerce einteilen lässt, ist der Alexa Skill «Bring!» des gleichnamigen Schweizer Unternehmens. «Bring!» bietet schon seit längerem eine App an, auf der die Nutzenden ihre Einkaufsliste führen und teilen können. Mit dem Alexa-Skill kann das nun auch ohne Display erledigt werden (Kruse Brandão & Wolfram, 2018, S. 274). Das ist vor allem ein Vorteil in Situationen, in denen man die Hände gerade nicht frei hat, beispielsweise wenn man während des Kochens merkt, dass man neue Milch kaufen sollte, weil man gerade die letzte Packung aufgebraucht hat.

Bislang wurde kaum zu Voice Assistants in der Schweiz geforscht. Die erste Studie, die sich spezifisch mit Voice Assistants auf dem Schweizer Markt befasst hat, ist im Jahr 2018 erschienen. Ein Ergebnis der Studie war unter anderem, dass zwar bereits ein Drittel der befragten Personen Voice Assistants benutzt, doch gerade einmal 1 Prozent nutzt einen Smart Speaker für die Voice-Interaktion. Bis Ende 2019 planen jedoch 14 Prozent der Befragten die Nutzung eines Smart Speakers (Kunath, Hofstetter, Jörg & Demarchi, 2019, S. 4). Daraus lässt sich schliessen, dass viele Befragte nicht grundsätzlich Ablehnung gegenüber Smart Speakern empfinden, sondern durchaus interessiert sind. Die tiefen Nutzungszahlen im Jahr 2018 bestätigen, dass die

eingeschränkte Funktionalität von Smart Speakern in der Schweiz diese interessierten Personen von einer Nutzung abhält. Aus dem prognostizierten Anstieg um 13 Prozentpunkte lässt sich schliessen, dass viele Personen die Hoffnung auf einen offiziellen Launch der Smart Speaker von Amazon und Google im Jahr 2019 haben.

Auch im Bereich Voice Commerce haben bereits dreissig Prozent der Befragten erste Erfahrungen gemacht. Konkrete Kaufhandlungen wurden zwar erst von vier Prozent durchgeführt: Der Grossteil der Erfahrungen wurde im Bereich Produktinformation gemacht, was ja ebenfalls zu Voice Commerce gezählt wird (Kunath et al., 2019, S. 12). Aus der Studie geht hervor, dass sich die Einstellung gegenüber Voice Assistants und Voice Commerce kaum von anderen Ländern wie beispielsweise Deutschland unterscheidet. Auch dieser Report bestätigt, dass Mahlzeiten- und Lebensmittelbestellungen zwei der häufigsten Use Cases sind und dass vor allem wenig riskante Käufe im tiefen Preissegment für einen Kauf via Voice infrage kämen (Kunath et al., 2019, S. 13–14). Auch die Bereitschaft von Nichtnutzern, in Zukunft einen VA zu benutzen und sogar damit einzukaufen, ist in der Schweiz erstaunlich hoch (Kunath et al., 2019, S. 18).

Die Studie konnte keine signifikanten Unterschiede zwischen der Schweiz und anderen Ländern feststellen, was die Ablehnung gegenüber Voice im Allgemeinen betrifft. Wie zu erwarten sind potenzielle Nutzer skeptisch, was das Verständnis von Schweizerdeutsch betrifft und haben Angst, nicht verstanden zu werden (Kunath et al., 2019, S. 17). Der am häufigsten genannte Grund, der gegen die Nutzung von Voice Assistants spricht, ist jedoch der, dass der Nutzen schlicht noch zu klein ist und die (potenziellen) Nutzenden keinen Vorteil in VA sehen. Beinahe gleich oft genannt wird der Grund, dass man nicht möchte, dass ständig «jemand» mithört. Diese beiden Gründe wurden sogar noch häufiger als Datenschutzbedenken genannt (Kunath et al., 2019, S. 16). Auch diese Zahlen bestätigen, dass Voice Assistants in der Schweiz grundsätzlich auf Interesse stossen. Dazu müssen sich den Nutzenden genug Vorteile eröffnen und die Spracherkennung insofern verbessert werden, dass die Personen möglichst natürlich mit dem VA sprechen können, nämlich in ihrem jeweiligen schweizerdeutschen Dialekt.

## 2.6 Nutzerakzeptanz und Vertrauen

In diesem Kapitel wird der Frage nachgegangen, wie das Vertrauen und die Nutzerakzeptanz von Voice Commerce zustande kommt beziehungsweise was die Einflüsse darauf sind. Um das herauszufinden, muss zuerst grundsätzlich betrachtet werden, wie Vertrauen in Technologien entsteht. Ohne ein allgemeines Vertrauen in

Technologien wären Innovationen wie eben auch Voice Assistants und Smart Speaker nie möglich geworden.

### 2.6.1 Akzeptanz von Technologien

Kittl (2009) definiert Akzeptanz als «Widerspruch zur Ablehnung und ist gleichzusetzen mit der positiven Annahme-Entscheidung (Adoption) einer technologischen Innovation sowie deren weitere, andauernde Nutzung» (S.52). Ginner (2018) hingegen unterscheidet die Adoption klar von der Akzeptanz. Bei einer Adoption muss eine kontinuierliche Nutzung gegeben sein. Für die Akzeptanz hingegen reicht es schon aus, wenn die Innovation und deren Gebrauch grundsätzlich in Frage kommt. Akzeptanz sagt ihm zufolge nichts über die eigentliche Nutzung aus, sondern eher über die Nutzungsabsicht der Innovation (S. 151).

Eine wichtige Errungenschaft in der Akzeptanzforschung ist das Technology Acceptance Model (TAM) von Davis, Bagozzi und Warshaw (1989, S. 985). Das Modell dient zur Vorhersage der Akzeptanz von Technologien und wurde speziell für Innovationen im Bereich der Informationstechnologien und -systeme entwickelt (Davis et al., 1989, S. 985). Das TAM geht davon aus, dass die Haltung eines Nutzers gegenüber einer Technologie und damit deren Akzeptanz vor allem auf zwei zentralen Faktoren beruht, die jeweils von externen Variablen beeinflusst werden. Externe Variablen sind beispielsweise Persönlichkeitsmerkmale oder demographische Merkmale, aber auch Merkmale des Systems oder der Innovation. Die beiden zentralen Faktoren sind zum einen der «wahrgenommene Nutzen» (Perceived Usefulness), zum anderen die «wahrgenommene Benutzerfreundlichkeit» (Perceived Ease of Use). Die Faktoren werden wie folgt definiert:

«Perceived Usefulness is defined as the prospective user's subjective probability that using a specific application system will increase his or her job performance» (Davis et al., 1989, S. 985).

«Perceived Ease of Use defines the degree to which the prospective user expects the target system to be free of effort.» (Davis et al., 1989, S. 985).

Der wahrgenommene Nutzen wird also dadurch definiert, wie wahrscheinlich es ist, dass sich die individuelle Leistung aufgrund der Anwendung verbessert (Ginner, 2018, S. 159). Je mehr ein Nutzer also merkt, dass es durch die Innovation bei seiner Arbeit unterstützt wird, desto höher ist der wahrgenommene Nutzen und somit auch automatisch seine Akzeptanz für die Innovation (Kittl, 2009, S. 53).



Die wahrgenommene Benutzerfreundlichkeit ist ebenfalls individuell abhängig und betrachtet, in welchem Ausmass die Anwendung einfach zu handhaben ist. Nach diesem Kriterium wird eine Innovation am ehesten angenommen, wenn der Nutzende einen möglichst kleinen Aufwand bei der Nutzung betreiben muss. Im Idealfall ist der Aufwand durch die technologische Innovation sogar kleiner, als er ohne diese wäre (Kittl, 2009, S. 53).

Wie in Abbildung 8 zu sehen, beeinflusst die wahrgenommene Benutzerfreundlichkeit den wahrgenommenen Nutzen. Das ist dadurch zu erklären, dass eine einfachere Benutzung die Leistung des Nutzers steigert und somit automatisch ein höherer Nutzen empfunden wird (Ginner, 2018, S. 159).

Die «Attitude Towards Use» beschreibt die grundsätzliche Einstellung einer Person gegenüber der technologischen Innovation. Diese ist entweder eher positiv oder eher negativ, sprich zustimmend oder ablehnend. «Behavioral Intention to Use» bezeichnet die Absicht, die Innovation wirklich zu nutzen oder eben auch nicht. Diese Faktoren wirken sich schliesslich auf «System Use» aus. Damit ist die effektive Nutzung der Innovation gemeint (Kittl, 2009, S. 54).

Die Kernaussage des TAM lautet also: Je höher die die individuell wahrgenommene Nützlichkeit und je einfacher die Bedienung, desto wahrscheinlicher ist es, dass die technologische Innovation akzeptiert und genutzt wird.

Daraus lässt sich die folgende Teilfragestellung ableiten:

*«Je grösser die Anzahl an Nutzungsmöglichkeiten, die einem Voice-Assistent-Nutzer aufgezeigt werden, desto eher akzeptiert dieser die Technologie.»*

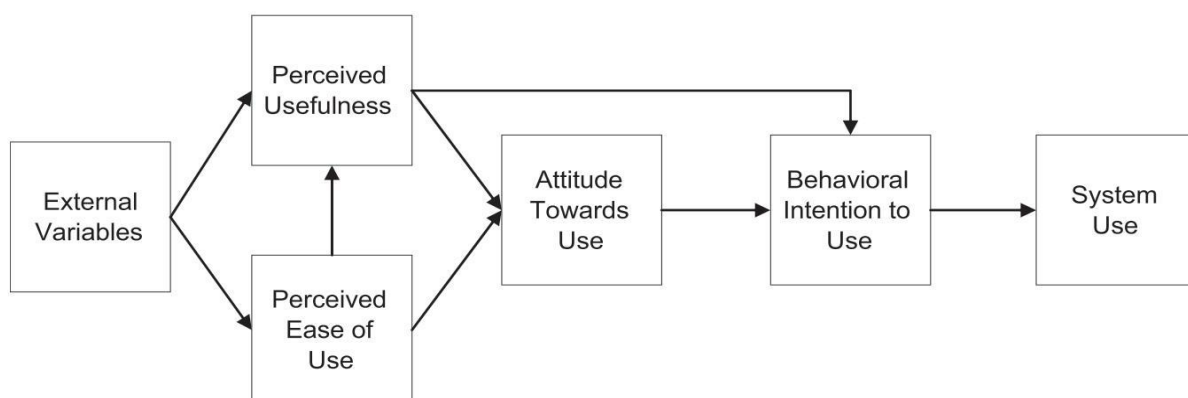


Abbildung 8: Technology Acceptance Model nach Davis

Quelle: Kittl (2009, S. 54).

## 2.6.2 Vertrauen in digitalen Transaktionen

Vertrauen ist sehr eng mit der Nutzerakzeptanz verwoben und eigentlich ein Teil derselbigen. Vertrauen ist vor allem bei Transaktionen mit einem oder mehreren Partnern relevant. Petrovic, Fallenböck, Kittl und Wolking (2003) definieren Vertrauen wie folgt: «Vertrauen ist die Bereitschaft zur Erbringung einer riskanten Vorleistung, die auf der Erwartung basiert, dass Vertrauensobjekte (Personen, Systeme) die entstandene Abhängigkeit nicht opportunistisch ausnutzen bzw. sich als funktionsfähig erweisen» (S.54). Eine weitere Definition von Vertrauen lautet: «Vertrauen ist in der Erwartung, dass der Vertrauensnehmer freiwillig auf opportunistisches Verhalten verzichtet» (Eggs, 2001, S. 73). Ohne Vertrauen wären Transaktionen im digitalen Raum ein unmögliches Unterfangen. Ist kein Mindestmass an Vertrauen vorhanden, wird sich keiner der Transaktionspartner bereiterklären, eine riskante Vorleistung zu erbringen. Ausserdem resultiert aus dieser Vorleistung eine Abhängigkeit vom Transaktionspartner (Kittl, 2009, S. 61). Wie beim TAM ist es auch hier ausschlaggebend, dass in der Transaktion ein Nutzen gesehen wird. Das Vertrauen kann noch so hoch sein: Wenn der erwartete Nutzen zu gering ist, wird die Transaktion nicht zustande kommen. Im Umkehrschluss bedeutet das, dass eine Transaktion durch grosses Vertrauen zustande kommen kann, auch wenn eine objektive Unsicherheit besteht. Diese wird aber in der subjektiven Wahrnehmung weniger als solche wahrgenommen (Petrovic et al., 2003, S. 55).

Der Vertrauensvorschuss, den eine Partei gegenüber der anderen Partei und dem System vorstrecken muss, erscheint im ersten Moment als hohe Hürde. Kuhlen (2008, S. 49) jedoch schreibt, dass in der deutschen Gesellschaft der Vertrauensvorschuss in informationstechnische Systeme grundsätzlich gewährt wird. Entscheidend ist, dass dieser Vertrauensvorschuss über einen längeren Zeitraum bestätigt wird, da er sich schnell aufbraucht. Die Bildung von Vertrauen ist also ein Prozess und darf nicht nur als Ergebnis von Handlungen und Massnahmen angesehen werden. Eggs (2001, S. 81–82) schreibt dazu, dass die Häufigkeit, mit der Transaktionen mit dem gleichen Partner durchgeführt werden, massgeblich zur Entstehung von Reputation und Vertrauen beiträgt. Dies wird schrittweise aufgebaut. Das bedeutet, dass mit jeder neuen Transaktion zwischen den gleichen Partnern das Vertrauen gesteigert wird.

### 2.6.2.1 Entstehung von Vertrauen

Ein Adoptionsprozess einer Technologie oder Innovation läuft immer auf einer individuellen Ebene ab. Das bedeutet, dass jede Person individuell den Adoptionsprozess durchläuft und am Ende dessen entscheidet, ob sie die Technologie

akzeptiert oder ablehnt (Rogers, 2003, S. 169). Rogers (2003) unterteilt diesen Prozess in fünf Phasen:

- In der *Wissensphase* erfährt eine Person von der Existenz einer Technologie, was sie macht und wie sie funktioniert.
- In der *Überzeugungsphase* entwickelt die Person eine Einstellung gegenüber der Innovation. Diese ist entweder positiv, neutral oder negativ.
- In der *Entscheidungsphase* wird der Entscheid gefällt, ob die Technologie adaptiert wird oder nicht.
- In der *Implementierungsphase* wird die Innovation angewendet.
- In der letzten Phase, der *Bewertung*, wird die gefällte Entscheidung noch einmal bewertet und allenfalls revidiert.

Zusätzlich definierte Rogers (2003, nach Bröcker, 2018) fünf zentrale Faktoren, die einen Einfluss auf den Adoptionsprozess haben:

- *Relativer Vorteil*: Ein Vorteil, den ein Individuum aus der potenziellen Adoption der Technologie zieht. Je grösser der Vorteil, desto schneller die Adaption.
- *Kompatibilität*: Die Innovation ist kompatibel mit bestehenden Systemen und den persönlichen Werten und Erfahrungen, die eine nutzende Person bereits gemacht hat. Je grösser die Kompatibilität, desto schneller die Adaption.
- *Komplexität*: Die Einfachheit der Nutzung, die subjektiv empfunden wird. Muss ein Nutzer viel Lernaufwand investieren, sinkt die Wahrscheinlichkeit einer Übernahme der Innovation.
- *Erprobbarkeit*: Die Möglichkeit des Ausprobierens der Innovation. Je mehr ein Nutzer die Innovation ausprobieren kann, desto kleiner sein Risiko, dass er enttäuscht wird. Die Wahrscheinlichkeit zur Adoption steigt mit erhöhter Erprobbarkeit an.
- *Beobachtbarkeit*: Das Mass, inwieweit ein Nutzen der Innovation sichtbar wird. Je sichtbarer der Nutzen, desto höher ist die Wahrscheinlichkeit einer Übernahme.

Wenn eine Innovation wie Voice Assistants diese Faktoren berücksichtigt und erfüllt, sollte eine erfolgreiche Adaption durchaus möglich sein. Anbieter von Voice-Applikationen, sollten diese Faktoren bereits bei der Erstellung der Anwendung berücksichtigen.

### 2.6.2.2 Klassifizierung von Vertrauen

Petrovic et al. (2003, S. 59) haben versucht, vertrauensbildende Signale in digitalen Transaktionen in Klassen einzuteilen. Die Einteilung erfolgt in die Klassen Information, Reputation und Garantie. Ausserdem wurden auch Vertrauensobjekte in drei Kategorien

eingeteilt. Vertrauensobjekte sind beispielsweise Fragen des Verbraucherschutzes, der Lieferzeiten und -gebiete, Schutz von privaten, personenbezogenen Daten (Eggs, 2001, S. 9). Die Kategorien sind Systemvertrauen, Vertrauen in den Transaktionspartner und Vertrauen in das Kontrollsystem. Jedes vertrauensbildende Signal kann also einem Vertrauensobjekt und in eine Vertrauensart eingeteilt werden (Petrovic et al., 2003, S. 59). In der Tabelle 1 ist die Klassifizierung mit Beispielen dargestellt.

Systemvertrauen beschreibt das Vertrauen, das eine nutzende Person in ein technologisches System haben muss. Das ist beispielsweise die Erwartung, dass das System fehlerfrei funktioniert, dass es tut, was es soll und dass das System sicher ist und somit die persönlichen Daten vor Missbrauch durch Dritte geschützt werden (Petrovic et al., 2003, S. 59). Vertrauen in den Transaktionspartner findet auf einer persönlicheren Ebene statt. Nutzende erwarten von einem Transaktionspartner, dem sie einen Vertrauensvorschuss gewähren, dass dieser nicht ausgenutzt wird. Beispiele für Vertrauen in den Transaktionspartner sind unter anderem, dass die vereinbarten Lieferbedingungen eingehalten werden oder auch, dass der Transaktionspartner die eigenen persönlichen Daten nicht missbräuchlich verwendet (Petrovic et al., 2003, S. 59). Die dritte Kategorie betrifft das Vertrauen in eines oder mehrere Kontrollsysteme. Ein Beispiel für ein Kontrollsystem ist die Verschlüsselungstechnik bei der Kreditkartenbezahlung (Petrovic et al., 2003, S. 59).

Die zentrale Aufgabe von vertrauensbildenden Signalen ist, dass eine Transaktion zustande kommt, auch wenn eine gewisse Unsicherheit besteht. Signale in der Kategorie Information haben das Ziel, Informationsdefizite zu verringern und so das Vertrauen zu stärken. Dabei muss jedoch beachtet werden, dass mehr Information nicht automatisch besser ist, da ein Kunde oder eine Kundin so die Transaktion als zu komplex und kompliziert wahrnehmen könnte (Petrovic et al., 2003, S. 59). Die zweite Kategorie wird von Petrovic et al. (2003, S. 59) folgendermassen definiert: «Reputation kann als Ruf eines Systems oder einer Person angesehen werden, inwieweit Dritte in der Vergangenheit ein vertrauenswürdiges Verhalten feststellten.» Ist beispielsweise die Reputation eines Transaktionspartners sehr gut, kann daraus geschlossen werden, dass sich der Partner auch in Zukunft vertrauenswürdig verhalten wird. Reputation soll verhindern, dass durch die digitale Transaktion ein Schaden ausgelöst wird. Im Gegensatz dazu sollen Signale in der Kategorie Garantie gewährleisten, dass im Falle eines bereits eingetretenen Schadens dieser möglichst begrenzt werden kann (Petrovic et al., 2003, S. 60). Das kann ein Schadenersatz oder auch Bestimmungen zur Rücknahme bei einem Schaden sein.

Aus diesen Erkenntnissen und der Klassifizierung von Petrovic et al. (2003, S. 60) wurde die folgende Teilfragestellung in Bezug auf Voice Commerce abgeleitet:

«Einflussfaktoren auf das Vertrauen in Voice Commerce lassen sich in die Kategorien Information, Reputation und Garantie einteilen.»

Vertrauensobjekte	Vertrauensbildendes Signal		
	Information	Reputation	Garantie
<b>System</b>	Darstellung der Funktionsweise	Ratingdienste, Gütesiegel	Garantien und Versicherungen, Zertifizierungen
<b>Transaktionspartner</b>	Produktinformationen	Ratingdienste, Gütesiegel	Gütesiegel, Versicherungen
<b>Kontrollsystem</b>	Darstellung der Funktionsweise	Gütesiegel	Garantien und Versicherungen

Tabelle 1: Arten von vertrauensbildenden Diensten mit Beispielen

Quelle: Eigene Darstellung, nach Petrovic et al. (2003, S. 60).

### 2.6.3 Entstehung von Vertrauen in Voice Commerce

Die vorhergehenden Kapitel haben sich allgemein mit Vertrauen und Nutzerakzeptanz in Technologien und Transaktionen befasst. In diesem Kapitel stehen nun explizit Smart Speaker und Voice Assistants im Vordergrund und innerhalb dessen die Anwendung des Voice Commerce. Da jedoch zu wenig einschlägige Literatur zum Thema Vertrauen in Voice Commerce gefunden werden konnte, werden die Theorien aus den vorherigen Kapiteln und der genutzten Literatur zum Thema Vertrauen in technologische Innovationen auf Voice Commerce abgeleitet. Dies, weil Voice Assistants ja auch eine technologische Innovation sind und darum die meisten Kriterien automatisch auch zu Voice Assistants passen. In diesem Kapitel wird nun auf die effektive Entstehung von Vertrauen in Voice Commerce eingegangen sowie auf die Faktoren, die das Vertrauen begünstigen oder erschweren.

#### 2.6.3.1 Vertrauen in den Plattformanbieter

Die Risiken, mit denen Nutzende von Voice Commerce konfrontiert sind, sind oft die gleichen wie in allen E-Commerce-Transaktionen. Der Austausch von Gütern und Geld erfolgt nicht zeitgleich; es herrscht eine räumliche Distanz und somit eine gewisse

Anonymität zwischen Anbieter und Käufer. Ausserdem kann der Käufer die effektive Qualität des Produktes nicht ausreichend beurteilen und er weiss auch nicht, wie die persönlichen Daten, die beim Kauf entstanden sind, beim Verkäufer gespeichert und verwendet werden (Grabner-Kräuter & Kaluscha, 2003, S. 133). Der Erfolg von E-Commerce der letzten Jahre zeigt jedoch, dass diese Risiken akzeptiert und eingegangen werden. Demnach müssten die Nutzer von E-Commerce also auch automatisch bereit sein, Voice Commerce zu nutzen. Das ist jedoch nicht der Fall, denn im Gegensatz zu den üblichen Online-Einkäufen ist bei Voice Commerce noch ein dritter Akteur beteiligt, nämlich der Hersteller der Sprachassistentensoftware, beispielsweise Google oder Amazon. Ein Nutzer von Voice-Commerce-Anwendungen muss also nicht nur Vertrauen in den Anbieter der Anwendung, sondern auch in die darunterliegende Infrastruktur der Software haben (Stanoevska-Slabeva, 2018, S. 34). Unternehmen, die Voice-Anwendungen planen, müssen sich also vorgängig überlegen, welchem Sprachassistenten sie vertrauen und ob auch die potenziellen Kunden Vertrauen in diesen Anbieter haben. In der Klassifikation nach Petrovic et al. (2003) lässt sich dies in die Kategorie Reputation einordnen. (Petrovic et al., 2003)

### **2.6.3.2 Vertrauen in den Kooperationspartner**

«Je besser man den Kooperationspartner kennt, desto geringer wird das Risiko eingeschätzt, dass dieser Informationen zurückhält» (Eggs, 2001, S. 27). Wenn das Unternehmen, das eine Voice-Applikation plant, bekannt und etabliert ist, wird also automatisch mehr Vertrauen gewährt. In Voice-Commerce ist die Nutzungschance also am wahrscheinlichsten, wenn der Anbieter bereits einen Onlineshop etabliert hat und dieser dem Kunden somit bekannt und vertraut ist (Sievers, 2019, S. 5).

Ein Kooperationspartner kann das Vertrauen in seine Anwendung erhöhen, indem er sich beispielsweise am TAM oder an den Kriterien von Rogers (2003) orientiert. Auch das Vertrauen in den Kooperationspartner wird der Kategorie Reputation zugeordnet.

### **2.6.3.3 Vertrauen durch Nutzen**

Gemäss dem TAM kann die Nutzerakzeptanz gesteigert werden, wenn die Bedienung möglichst einfach und der Nutzen für die anwendende Person möglichst gross ist. Für Voice Assistants bedeutet das, dass mit jeder neu dazukommenden Funktion und Anwendung der Nutzen für die User potenziell gesteigert wird. Anbieter von Smart Speakern könnten beispielsweise in der Werbung die Qualität sowie die grosse Bandbreite von Anwendungen hervorheben, die durch einen Smart Speaker erschlossen werden (Kowalczyk, 2018, S. 427). Durch das Aufzeigen von vielen Use Cases können

ausserdem mehr Nutzergruppen angesprochen werden, was die allgemeine Akzeptanz wiederum fördert.

Auch die Anbieter von Voice-Applikationen müssen einen Weg finden, den Kundinnen und Kunden zu zeigen, welchen Nutzen die Anwendung stiftet und wie einfach die Nutzung ist. Werden diese zwei Punkte aufgezeigt und auch erfüllt, kann gemäss dem TAM die Nutzerakzeptanz als gegeben erachtet werden. Erste Schritte zur Generierung von Nutzen sind Aktionen und Vergünstigungen, durch die die Anwendenden Geld einsparen können (Capgemini Digital Transformation Institute, 2018, S. 10). So wird für die Nutzer sofort ein Wert erkennbar, nämlich durch das eingesparte Geld gegenüber herkömmlichen E-Commerce-Transaktionen. Ein weiterer Nutzen, der langfristig Vertrauen und Akzeptanz generiert, ist die Zeiteinsparung, die bei Voice Commerce im Vergleich zu herkömmlichem Online-Shopping gemacht wird (Capgemini Digital Transformation Institute, 2018, S. 10). Auch hier wird der Nutzen für den Kunden sofort sichtbar und steigt mit einer wiederholenden Nutzung immer stärker an. Anbieter von Voice-Commerce-Applikationen müssen sich deshalb der Herausforderung stellen, wie sie den Prozess des Einkaufens via Stimme möglichst schnell und einfach gestalten können. Die Anzahl Schritte und die dialogische Interaktion soll dabei auf ein Minimum reduziert werden, um Fehler und Missverständnisse, die zu einem Abbruch von Seiten des Nutzers führen würde, zu reduzieren (Kraus, Reibenspiess & Eckhardt, 2019, S. 10). Zeitintensive Aktionen, wie beispielsweise die Eingabe einer Lieferadresse oder die Übermittlung der Zahlungsinformationen, sollten dabei möglichst nicht via Voice eingegeben werden müssen.

#### **2.6.3.4 Vertrauen durch Gewohnheit und Gewöhnung**

Anbieter von Voice-Applikationen und Hersteller von Voice Assistants können theoretisch extreme Anstrengungen unternehmen, um die Akzeptanz ihrer Systeme bei den Nutzenden zu fördern. Ein Faktor, der hingegen nicht beeinflusst werden kann, ist die Zeit. Durch den schrittweise verlaufenden Aufbau von Vertrauen ist es normal, dass es eine gewisse Zeit dauert, bis die ersten Nutzenden wirklich Vertrauen gefasst haben. Denn Vertrauen entsteht durch Gewohnheit und Gewöhnung (Kuhlen, 2008, S. 42). Wenn der Einkaufsprozess via Voice das erste Mal getätigt wird, ist es in den meisten Fällen nur ein Ausprobieren, grosses Vertrauen ist in den meisten Fällen eher nicht vorhanden. Verläuft jedoch alles gemäss den Erwartungen des Nutzenden, wird sich das Ausprobieren mit jedem weiteren Mal der Nutzung in Vertrauen wandeln.

### 2.6.3.5 Vertrauen durch Visualisierung

In diversen Literaturquellen wird Voice Commerce ein grosses Potenzial vorausgesagt, jedoch unter der Voraussetzung, dass beim Einkaufsprozess eine visuelle Komponente miteingebunden wird (Elster, 2019; Pardes, 2019; Sievers, 2019, S. 12). Das kann ein Gerät wie ein Smartphone oder ein Smart-TV sein, an das Informationen übermittelt werden, oder ein Screen, der direkt in den Smart Speaker integriert ist. Die grossen Hersteller Amazon und Google haben beide bereits solche Geräte mit Screen im Angebot. Gemäss Pardes (2019) kann Voice Commerce so sein volles Potenzial entfalten, da sich so mehr Nutzungsmöglichkeiten ergeben. Mit Smart Speakern, die einen Screen integriert haben, können beispielsweise auch Produkte eingekauft werden, die über die Alltagsprodukte hinausgehen und etwas komplexer sind. Denn durch den Screen wird es erlaubt, verschiedene Produkte zu vergleichen, was ohne Bildschirm und nur über die Stimmausgabe sehr mühsam und umständlich ist. Ist kein Bildschirm vorhanden, ist ein Kauf zwar möglich, Voice Commerce wird sich dann aber wohl nie so weit entwickeln, dass Nutzende wirklich auch Produkte vergleichen und Informationen zu einem Produkt suchen (Tuzovic & Paluch, 2018, S. 94).

Der Grund, warum viele Personen die Ergänzung mit einem Screen bevorzugen ist der, dass viele Menschen Informationen besser visuell als auditiv aufnehmen können (Sievers, 2019, S. 12). Wenn den Nutzenden also die Informationen auf einem Screen angezeigt werden, können sie sich sicher sein, alle für den Kauf notwendigen Informationen erhalten zu haben. Das wiederum schafft Vertrauen und steigert die Chance auf eine nachhaltige Akzeptanz von Voice Commerce. Der Screen kann dabei über alle Phasen des Einkaufsprozesses hinweg verwendet werden. Der Nutzer könnte beispielsweise den Voice Assistant per Spracheingabe damit beauftragen, nach den besten Kaffeemaschinen zum Preis unter hundert Franken zu suchen und drei Vorschläge zu machen. Diese Vorschläge werden auf dem Screen präsentiert und der Nutzer kann sich für ein Modell entschieden, da er die angezeigten Produktinformationen gut miteinander vergleichen kann. Nun kann er die Bestellung entweder selbst auf dem Screen tätigen oder via Voice ausführen lassen. Ein weiterer Vorteil eines Screens ist, dass damit weitere Möglichkeiten zur Authentifizierung geschaffen werden, beispielsweise durch die Gesichtserkennung (Deloitte, 2018, S. 53). So kann verhindert werden, dass fremde Personen oder Kinder Produkte im Namen des Nutzers bestellen können. Diese zusätzliche Sicherheit ist essentiell in der Generierung von Vertrauen und Akzeptanz.

Der Faktor Vertrauen durch Visualisierung wird der Klasse Information zugeteilt.



### **2.6.3.6 Vertrauen durch Verbesserung der Technologie**

Wenn man einen Blick in die Entwicklungsgeschichte von Voice Assistants wirft, sieht man die extreme Entwicklung, die diese Technologie in den letzten Jahren durchgemacht hat. Ein Ende dieser Entwicklung ist bislang nicht in Sicht, ständig gibt es neue Verbesserungen. Dennoch gibt es immer noch viele zu komplexe Dialoge, die die Voice Assistants überfordern. Auch die Eigenschaft, Verhaltensmuster und Charaktereigenschaften der Nutzer zu erkennen, zu analysieren und sich danach auszurichten, ist noch kaum ausgereift (Stanoevska-Slabeva, 2018, S. 34). Der «Human Touch» ist noch kaum gegeben. Das heisst, es fühlt sich für die meisten Nutzer nicht wirklich nach der Interaktion mit einem Menschen an, sondern dem Nutzenden ist ständig bewusst, dass er mit einer Maschine spricht. Wenn die Voice Assistants also persönlicher und «menschlicher» erscheinen, kann auch eher Vertrauen aufgebaut werden (Stanoevska-Slabeva, 2018, S. 34). Auch eine stimmabhängige Erkennung des Sprechers ist aktuell noch nicht zuverlässig und muss verbessert werden (Deloitte, 2018, S. 54). So könnte der VA in Zukunft den Nutzer sofort authentifizieren und gewisse Aktionen verweigern, wenn er merkt, dass es sich nicht um den korrekten Nutzer handelt. Dieser Faktor der Technologieverbesserung spielt eng mit dem Faktor Vertrauen durch Nutzung und Vertrauen durch Gewöhnung zusammen. Verbessert sich die Technologie, steigt der individuelle Nutzen. Ausserdem sind technologische Verbesserungen immer mit einem unbestimmten Zeitaufwand verbunden, weshalb der Faktor der Gewöhnung langfristig auch das Vertrauen stärkt.

### **2.6.3.7 Vertrauen durch Transparenz und Aufklärung**

Gerade in der Anfangszeit nach dem Launch einer Voice-Applikation ist es wichtig, dem Kunden klar zu machen, dass die Sprachtechnologie wie auch die Anwendung noch nicht perfekt funktioniert (Elster, 2019). Die Nutzer müssen sich darauf einstellen, dass nicht alles reibungslos funktioniert, um nicht enttäuscht zu werden, wenn sie die Anwendung oder den Voice Assistant das erste Mal nutzen. Wenn der Kunde mit diesen realistischen Erwartungen die Funktion ausprobiert, wird seine Einstellung eher positiv ausfallen, auch wenn die Anwendung noch kleine Fehler aufweist. Der Nutzende war jedoch darauf vorbereitet und wird das darum nicht zwingend als negativ empfinden. Umgekehrt wird ein sehr positives Bild vermittelt, wenn die Anwendung die Erwartungen des Kunden übertrifft.

Vertrauen entsteht durch Transparenz (Kuhlen, 2008, S. 42). Diese Transparenz darf nicht einseitig sein, sondern muss auf Gegenseitigkeit beruhen. Ein Nutzer von Voice-Commerce gibt viel von sich preis und wird somit für den Transaktionspartner

transparent. Diese Transparenz wird auch vom Anbieter eingefordert. Transparenz ist dabei sehr vielfältig und auch sehr individuell geprägt. Beispielsweise sollte dem Kunden klar sein, welche Funktionen und Möglichkeiten ihm bei der Nutzung offenstehen, was das Produkt kann und was nicht. Datenschutzbedenken sind der häufigste Hinderungsgrund, der von Nichtnutzern genannt wird (Capgemini Digital Transformation Institute, 2018, S. 17), deshalb müssen Anbieter auch hier Transparenz zeigen und offenlegen, was mit den gesammelten Daten passiert und wo diese gespeichert und ausgewertet werden. Die Nutzer müssen aktiv aufgeklärt werden (Deloitte, 2018, S. 56).

Dabei müssen Anbieter zusätzlich beachten, dass die Komplexität möglichst nicht ansteigt, denn das würde gemäss dem TAM die Einfachheit und somit die Akzeptanz wieder verringern (Ginner, 2018, S. 159). Trotzdem werden die Security-Bedenken wohl nie ganz aus der Welt geschaffen werden können. Es wird immer Personen geben, die deswegen von einer Nutzung absehen. Denn dass das Mikrofon ständig aktiviert sein muss, um das Erkennungswort zu hören, oder dass die Daten auf meist ausländischen Servern gespeichert und ausgewertet werden, ist nun einmal die Funktionsweise von Voice Assistants (Martin, 2019). Entscheidend ist, dass die Kundinnen und Kunden sich dieser Tatsachen bewusst sind und wissen, wo sie sich informieren können, falls Fragen auftauchen. Es muss den Kunden ausserdem auch aufgezeigt werden, dass die Daten, die von ihnen produziert werden, auch dazu dienen, die Anwendungen und die Spracherkennung zu verbessern, was das Vertrauen und die Akzeptanz wiederum erhöhen würde (McTear et al., 2016, S. 22).

Vertrauen durch Transparenz wird in der Klassifizierung nach Petrovic et al. (2003) in die Kategorie Information eingeteilt.

### 3 Methode

Diese wissenschaftliche Arbeit beruht auf zwei verschiedenen Forschungsstrategien. Der erste Teil basierte auf einer systematischen Literaturrecherche nach Latzer (2016) und Brocke et al. (2009). Durch dieses Vorgehen konnte wissenschaftliche Literatur gefunden und der aktuelle Forschungsstand dargelegt werden.

Der nun folgende empirische Teil bildet den zweiten Teil dieser wissenschaftlichen Arbeit. Ziel dieser Empirie ist es, die aus der Literatur gewonnen Erkenntnisse zur Akzeptanz und Vertrauen in Voice Commerce zu bestätigen oder zu verwerfen. Ausserdem soll durch diese Erhebung der Stand von Voice Assistants in der Schweiz erhoben werden, da dazu kaum wissenschaftliche Literatur vorhanden ist.

#### 3.1 Methodisches Vorgehen

Für den empirischen Teil wurde eine qualitative Erhebung gewählt. Im Gegensatz zu quantitativen Erhebungsformen wird bei der qualitativen Erhebung anstelle von Daten Text erhoben (Aghamanoukjan, Meyer & Buber, 2009, S. 417) Dies wurde in Form von Experteninterviews erreicht. Als Zielgrösse wurden drei bis vier Interviews definiert, die zwischen dem 10. Juni 2019 und 5. Juli 2019 durchgeführt werden sollten.

Die Methode der Experteninterviews wurde gewählt, da mit Expertinnen und Experten die Zielsetzung, den Stand von Voice Assistants in der Schweiz zu erfassen, am besten realisiert werden konnte. Interviews mit Expertinnen und Experten ermöglichen einen sehr aktuellen Blick auf ein Thema. In einem sich rasant verändernden Gebiet wie Voice Assistants dauert die Publikation von wissenschaftlicher Literatur oft zu lange. Expertinnen und Experten, die in Theorie und Praxis auf dem Gebiet tätig sind, können ihr generiertes Wissen so direkt weitergeben. Bogner, Littig und Menz (2014) definieren Expertinnen und Experten wie folgt:

«Experten lassen sich als Personen verstehen, die sich – ausgehend von einem spezifischen Praxis- oder Erfahrungswissen, das sich auf einen klar begrenzbaren Problemkreis bezieht – die Möglichkeit geschaffen haben, mit ihren Deutungen das konkrete Handlungsfeld sinnhaft und handlungsleitend für Andere zu strukturieren.»  
(S. 13)

Sinn und Zweck der Interviews war es nicht, konkrete Hypothesen zu testen, die aus der Theorie generiert wurden. Es stand vielmehr im Zentrum, weiteres Wissen und Material zu generieren, das bisher aus der Literatur nicht oder nicht sehr detailliert ersichtlich war (Bogner et al., 2014, S. 32). Ausserdem stand oft die persönliche Meinung oder

Einschätzung des Experten oder der Expertin im Vordergrund, da diese meistens einen konkreten Praxisbezug hatten.

### 3.1.1 Auswahl der Expertinnen und Experten

Das Ziel bei der Auswahl der Expertinnen und Experten war, für jedes Themenfeld dieser wissenschaftlichen Arbeit ein Interview durchführen zu können. Dies, da es schon im Vorherein unmöglich erschien, eine Person zu finden, die sich mit Vertrauen in Voice Commerce in der Schweiz beschäftigt. Die Themenbereiche wurden darum aufgeteilt und im Idealfall hat eine Expertin oder ein Experte mehrere Teilbereiche abgedeckt. Durch Kontakte und Webrecherche war es möglich, vier Personen ausfindig zu machen, die diese Anforderung erfüllten. Die vier Personen wurden per E-Mail angefragt. Erfreulicherweise haben drei Personen nach kurzer Zeit bereits die Zusage für ein Interview erteilt. Die vierte Person hat auch nach mehrmaligem Nachfragen nicht auf die Anfrage reagiert.

Person	Schwerpunkt der Befragung	Durchführung
Urs Bucher	Voice Commerce in der Schweiz	11.06.2019
Karen Kaushansky	Vertrauen, Voice Commerce	03.07.2019
Riccardo Lopetrone	Vertrauen in Voice in der Schweiz	05.07.2019
Katarina Stanoevska-Slabeva	Forschung Vertrauen in CI	Keine Antwort

Tabelle 2: Angefragte Expertinnen und Experten

Quelle: eigene Darstellung

Urs Bucher aus Zürich arbeitete bei Farner Consulting AG und war dort für den Aufbau des Farner Lab zuständig, das sich unter anderem mit Voice-Technologie befasste. Ausserdem hat er in Zusammenarbeit mit Tamedia AG Workshops durchgeführt, die sich mit dem Thema Voice-Anwendungen im Medienbereich befassten (U. Bucher, persönl. Mitteilung, 11.06.2019).

Karen Kaushansky stammt aus den USA und arbeitet seit 23 Jahren an und mit Voice-Technologien. Seit zwei Jahren wohnt sie in der Schweiz, wo sie für Google tätig ist und sich im Google Assistant-Team im Bereich der Produktivitätssteigerung beschäftigt (K. Kaushansky, persönl. Mitteilung, 03.07.2019).

Riccardo Lopetrone arbeitet bei der Swisscom als Product Manager in der Abteilung Swisscom TV & Entertainment. Er beschäftigt sich dort mit sprachgesteuerten Produktfeatures (R. Lopetrone, persönl. Mitteilung, 05.07.2019).

Katarina Stanoevska-Slabeva arbeitet an der Universität St. Gallen und forscht dort im Bereich der Conversational Interfaces. Sie hat zu dem Thema Studien und wissenschaftliche Publikationen verfasst, die auch in dieser Bachelorarbeit verwendet wurden. Sie konnte nicht als Expertin für diese Bachelorthesis gewonnen werden.

### **3.1.2 Interviewaufbau und -durchführung**

Experteninterviews, die im Rahmen der qualitativen Sozialforschung durchgeführt werden, sind immer teilstrukturiert und werden somit mit Leitfäden durchgeführt (Bogner et al., 2014, S. 27). Ein Leitfaden dient während des Interviews als Orientierung, die Fragen müssen aber nicht wortwörtlich gemäss dem Leitfaden gestellt werden. Das erlaubt es der interviewenden Person, bei Bedarf genauer auf die Antworten des Experten oder der Expertin einzugehen und nachzufragen.

Die teilstrukturierten Fragebögen wurden für jede interviewte Person neu erstellt, um dem Teilgebiet, auf dem sich die Person auskennt, Rechnung zu tragen. Die Struktur des Leitfadens, nämlich Informationen zur Arbeit und zum Ablauf, einleitende Fragen, Hauptteil und Schluss, wurde hingegen bei allen Interviews beibehalten. Die einzelnen Frageblöcke hingegen wurden jeweils ausgetauscht oder angepasst.

Alle drei Interviews fanden in Form eines Einzelinterviews statt und wurden mündlich durchgeführt. Zwei Interviews fanden face-to-face statt, eines via Skype.

Um den Anforderungen der Dokumentation zu genügen, wurden die Interviews aufgezeichnet und transkribiert. Das sollte auch die Auswertung der Interviews erleichtern. Während der Gespräche wurden die wichtigsten Aussagen pro Frage trotzdem noch handschriftlich festgehalten, falls es ein technisches Problem mit der Aufnahme geben sollte. Die Interviewtranskripte finden sich im Anhang.

### **3.1.3 Aufbau Interviewleitfaden**

Der Interviewleitfaden startete jeweils mit einigen einleitenden Fragen zur Person, zum Beruf und zur privaten Nutzung von Voice Assistants. Das Ziel war, herauszufinden, wie lange sich der Experte oder die Expertin schon mit dem Thema Voice Assistants befasst, wie es dazu gekommen ist und wie die Person nebst dem beruflichen Kontext über diese

Technologie denkt. Diese Einleitungsfragen waren für alle Interviewten grösstenteils gleich.

Im Hauptteil wurden jeweils die entsprechenden, zur Person passenden, Frageblöcke gestellt. Die Themen gliederten sich folgendermassen:

- *Praxisbezug und aktuelle Projekte*: Erfahrungen zu aktuellen und vergangenen Projekten im Bereich Voice Assistants, Kundenfeedback und Reaktionen auf das Projekt / Produkt / Funktion
- *Voice Commerce*: Potenzial des Einkaufens mit der Stimme, Herausforderungen und Voraussetzungen für den Erfolg von Voice Commerce
- *Vertrauen und Nutzerakzeptanz in der Schweiz*: Unterschiede in der Akzeptanz von Technologien im Vergleich zum Ausland, Begünstigung und Hindernisse für Nutzerakzeptanz in der Schweiz, Aufgeschlossenheit in der Schweiz gegenüber Innovationen
- *Vertrauen und Nutzerakzeptanz in Voice / Human Computer Interaction*: Kriterien für die Entstehung von Vertrauen und Nutzerakzeptanz, Minimierung von Security-Bedenken, Verhinderung von Vertrauensverlust

Die Interviewleitfäden bestanden mehrheitlich aus offenen Fragen, um den Experten oder die Expertin möglichst nicht in eine Richtung zu lenken. Sollte das Gespräch zu sehr in eine Richtung schwenken, kann durch erneutes Nachfragen oder eine Präzisierung wieder zurück zum eigentlichen Thema gefunden werden

Die kompletten drei Interviewfragebögen finden sich im Anhang.

#### **3.1.4 Interviewauswertung**

Die Interviewauswertung erfolgte durch ein Kodierungsverfahren. Dabei werden die Aussagen, die im Interview gemacht wurden, zu Kernaussagen zusammengefasst und in Kategorien eingeordnet (Mayring & Brunner, 2009, S. 674). Das Kodieren findet auf der Ebene des Einzelinterviews statt. Zur Auswertung hingegen werden die Kodierungen zwischen den Gesprächspartnern verglichen. So können gleiche oder ähnliche Aussagen zwischen den Gesprächspartnern einfacher erkannt und analysiert werden (Bogner et al., 2014, S. 78). Die Kodierung wurde in einer Excel-Tabelle erstellt und ist im Anhang einsehbar.

## **4 Ergebnisse**

In diesem Kapitel werden die Ergebnisse behandelt, die aus den durchgeführten Experteninterviews hervorgingen. Die Erkenntnisse sind thematisch gegliedert und orientieren sich an den Frageblöcken aus den Interviews. Teilweise gibt es Überschneidungen zwischen den Themen.

### **4.1 Voice Assistants in der Schweiz**

Gemäss der Umfrage von Kunath et al. (2019, S. 4) nutzen bereits ein Drittel der befragten Personen in der Schweiz Voice Assistants. Diese hohe Zahl war auch zwei Interviewpartnern bekannt und sorgte für Erstaunen, da dieser Wert fast so hoch ist wie in Deutschland. Hingegen nutzen gerade einmal ein Prozent der Befragten einen Smart Speaker. Dass es jedoch möglich ist, in der Schweiz einen oder mehrere Smart Speaker zu besitzen, haben alle drei Interviewpartner bestätigt, da sie selbst private Nutzer dieser Geräte sind. Es wurde bestätigt, dass die Anwendung von lokalen Applikationen nicht möglich ist, diese aber auch gar nicht vorhanden sind. Die Befragten nutzen ihre Smart Speaker und Voice Assistants vor allem um Musik zu hören, die Einkaufsliste zu erstellen, das Smart Home zu steuern, für News und allgemeine Wissensfragen.

### **4.2 Potenzial von Voice Assistants in der Schweiz**

Alle Befragten sind sich einig, dass Voice Assistants in der Schweiz Potenzial haben. Es wird geschätzt, dass die Schweiz die gleiche Entwicklung durchmachen wird wie beispielsweise Deutschland. Jedoch eher längerfristig gesehen, es könnte also gut noch fünf Jahre gehen, bis die Schweiz auf dem Stand ist wie Deutschland heute. Möglichkeiten werden vor allem in Nischen gesehen, die sehr lokal und schweizerisch sind und die für die grossen internationalen Anbieter nicht von primärem Interesse sind. Ein gutes Beispiel dafür ist die Spracherkennung auf Schweizerdeutsch. Diese ist technisch möglich und wird von einigen Schweizer Unternehmen auch bereits eingesetzt. Dagegen ist es unrealistisch, dass Google oder Amazon ein grosses Interesse daran haben, dass ihre Voice Assistants Schweizerdeutsch verstehen und auf Schweizerdeutsch interagieren können. Der Schweizer Markt ist dafür einfach zu klein und zu uninteressant.

### 4.3 Nutzerakzeptanz

Nutzerakzeptanz wird schrittweise aufgebaut. Neukunden von Voice Assistants werden nicht sofort alle Funktionen nutzen, sondern sich langsam vorantasten. Mit diesem Ansatz des Ausprobierens wird mit jeder erfolgreichen Anwendung mehr Vertrauen gewonnen und bei der nächsten Anwendung mehr «riskiert».

Die Nutzerakzeptanz kann gefördert werden, wenn grosse Player, denen man bereits online oder auch offline vertraut, vorausgehen. Das wären beispielsweise Medienhäuser, bekannte Onlineshops oder auch grosse Detailhändler. Je mehr Anwendungsgebiete es gibt, desto schneller steigt auch das Vertrauen von Anwendern in die Technologie. Ein Beschleuniger sind hier Unternehmen, die Voice Assistants in alle möglichen Produkte integrieren, beispielsweise in einen Spiegel oder ein Bett. So wird der Kunde an vielen Stellen parallel an Voice Assistants gewöhnt, egal ob er diese effektiv nutzt oder nicht.

#### 4.3.1 Nutzen und Wert

Ein sehr grosser Impact auf die Akzeptanz einer Technologie hat der gestiftete Nutzen und der Wert, der mit der Technologie generiert wird. Es gibt Anwendungen, bei denen Voice ein enormes Potenzial hat, beispielsweise weil man viel Zeit einsparen kann. Umgekehrt gibt es Use Cases, bei denen die Voice-Technologie keinen Sinn macht, weil der Nutzen und der generierte Wert für den Anwendenden schlicht zu klein ist. Es muss den Anbietern gelingen, einen so grossen Nutzen zu generieren, dass die Datenschutzbedenken dabei zurückgestellt werden, wie das heute bereits mit Plattformen wie beispielsweise Facebook geschieht. Die Datenschutzbedenken können also nicht unbedingt aufgehoben werden, sondern sie werden bewusst in Kauf genommen, um vom bereitgestellten Nutzen profitieren zu können.

Ausschlaggebend für den Nutzen ist ausserdem, dass das System oder die Anwendung funktioniert und die Aufgabe ausgeführt werden kann. Das klingt simpel, doch macht eine Anwendung nicht, was sie soll, löst das sofort Frustration aus und mindert somit den Nutzen und das Vertrauen. Nebst dieser Verlässlichkeit muss die Anwendung auch einfach verständlich und logisch sein. Ausserdem muss die Spracherkennung weiter verbessert werden und auch komplexe Anfragen bearbeiten können.

#### 4.3.2 Information und Kommunikation

Eine transparente Kommunikation kann gemäss den Experten ein Schlüssel zum Erfolg sein. Ein Unternehmen, sei das der Anbieter des Voice Assistants oder der Anbieter der



Voice-Applikation, muss dem Kunden klar kommunizieren, was die Anwendung kann und was nicht, welche Daten gesammelt werden und welche nicht und was mit den Daten passiert. Das muss auf eine Art geschehen, die der Nutzende auch versteht. Es nützt also nichts, die Informationen in den Allgemeinen Geschäftsbedingungen zu «verstecken», da diese in den meisten Fällen nicht gelesen werden. Und durch ein Vorlesen lassen der Nutzungsbedingungen wird der Nutzen für den Anwendenden massiv geschmälert. Hier sind die Unternehmen gefragt, aber auch die Medien. Diese tragen mit ihrer Berichterstattung nicht unwesentlich zur Einstellung gegenüber der Technologie bei. So gibt es beispielsweise viele negative Beispiele über Versuche mit Voice, die unklar kommuniziert wurden und darum auf Ablehnung oder Unverständnis stiessen. Mangelnde Information ist einer der wichtigsten Gründe für eine negative Einstellung gegenüber Voice Assistants. Wenn die Bevölkerung die Technologie noch kaum kennt, ist man automatisch dazu geneigt, ablehnend zu reagieren und eine Nutzung kategorisch auszuschliessen.

### **4.3.3 Nutzerakzeptanz in der Schweiz**

In der Schweiz ist der Datenschutz ein sehr präsent Thema. Die Entwicklung und Akzeptanz von neuen Technologien dauert darum oft länger als beispielsweise in den USA. In den USA herrscht eine Opt-out-Mentalität, sprich die Technologie wird einmal ausprobiert und wenn man nicht zufrieden ist oder Bedenken hat, beendet man die Nutzung. In Europa hingegen ist man vorsichtiger und gewichtet Datenschutz höher. Man möchte möglichst alle Fakten schon zu Beginn kennen und sich dann, wenn man mit allem einverstanden ist, bewusst für die Technologie entscheiden, also Opt-in. Die Schweiz unterscheidet sich hingegen nicht von den restlichen europäischen Ländern, was die Offenheit und Akzeptanz von neuen Technologien betrifft.

## **4.4 Voice Commerce**

Die Interviewten waren sich einig, dass Voice Commerce längst nicht für alle Produkte geeignet ist. Am meisten Potenzial sehen sie in alltäglichen Produkten im tiefen Preissegment und in Artikeln, die man nachbestellt. Ausserdem eignet sich diese Möglichkeit des Einkaufens für sehr spezifische Produkte, bei denen sich der Nutzer schon sicher ist, welches Produkt er bestellen möchte und dass er es bestellen möchte. Ein Beispiel dafür sind Bücher. Zwei Experten sind sich einig, dass es bei Voice Commerce einen Medienbruch benötigt, um das Vertrauen zu stärken. Sobald es um Geld geht, herrscht allgemein mehr Vorsicht und Misstrauen. Mit einem Medienbruch,

beispielsweise die Anzeige der Bestellinformationen vor dem Kauf, kann die Unsicherheit genommen werden, dass man beispielsweise das falsche Produkt bestellt. Eine ähnliche Möglichkeit wäre, die Bestellbestätigung nach dem Kauf visuell anzuzeigen. Wenn sich der Kunde oder die Kundin mit der Zeit an das Vorgehen gewöhnt hat und Vertrauen gefasst hat, könnte dieser Medienbruch auch weggelassen werden.

Die Experten sind sich uneins, ob es einen kritischen Punkt im Kaufprozess gibt, der überschritten werden muss. Bucher (2019) sieht denn effektiven Bestellabschluss am Schluss des ganzen Prozesses als kritischen Punkt an. Gemäss ihm braucht es an dieser Stelle eine möglichst einfache und schnelle Bestätigung der Bestellung, möglichst mit einem Medienbruch. Auch Lopetrone (2019) findet, dass es ein visuelles Feedback braucht. Kaushansky (2019) hingegen glaubt nicht, dass es einen kritischen Punkt gibt. Gemäss ihr muss der dargebotene Nutzen von Voice Commerce nur gross genug sein. Gemäss ihren Erfahrungen werden Personen, die bereits heute die Einkaufsliste via Voice erstellen, in Zukunft via Voice einkaufen, sobald sich diese Möglichkeit eröffnet.

#### **4.5 Voice Commerce in der Schweiz**

Die Schweiz ist, was Voice Commerce angeht, nicht wirklich ein attraktiver Markt für Google oder Amazon. Die Interviewten sehen aber durchaus Potenzial in der Technologie. Gerade Personen, die technologie-affin sind und bereits heute Online-Shopping nutzen, dürften schnell Vertrauen in Voice Commerce fassen, solange der Nutzen gross genug ist. Das grösste Problem, das Voice Commerce in der Schweiz nach wie vor hat, ist die mangelnde Verfügbarkeit von Voice Assistants und Anwendungen. Daraus ergibt sich ein Henne-Ei-Problem: Potenzielle Anbieter von Voice-Commerce-Applikationen entwickeln diese nicht, da noch keine Infrastruktur vorhanden ist, und die Anbieter der Infrastruktur lancieren ihre Produkte nicht in der Schweiz, da es kaum Anwendungen und darum auch kaum Nutzer gibt. Wenn dieser Kreis aber durchbrochen werden kann, steht einem erfolgreichen Start von Voice Commerce eigentlich nur noch wenig im Weg.

## 5 Fazit

In diesem Kapitel werden die Erkenntnisse aus Theorie und Empirie verwendet, um die Fragestellung zu beantworten, die zu Beginn dieser Arbeit aufgestellt wurde. Ausserdem werden die drei Teilfragestellungen, die im Verlauf dieser Arbeit aufgestellt wurden, beantwortet.

### 5.1 Einflüsse auf die Nutzerakzeptanz von Voice Commerce in der Schweiz

*«Welche Faktoren beeinflussen das Vertrauen und die Nutzerakzeptanz von Voice Commerce und Voice-Commerce-Anwendungen in der Schweiz?»*

In dieser wissenschaftlichen Arbeit konnten diverse Einflüsse auf die Nutzerakzeptanz von Voice Commerce in der Schweiz eruiert werden. Diese Faktoren lassen sich in die Hauptfaktoren Nutzen, Information und technologische Entwicklung zusammenfassen, die jedoch sehr eng miteinander verknüpft sind.

#### 5.1.1 Nutzen

Um Voice Commerce erfolgreich etablieren zu können, muss der Nutzen für den Kunden möglichst maximiert werden. Je mehr Nutzen Anwendende in Voice Commerce sehen, desto eher wird Vertrauen und Akzeptanz aufgebaut. Dabei können sogar allfällige Bedenken, beispielsweise gegenüber Datenschutz, in den Hintergrund rücken. Nutzen entsteht unter anderem durch einen einfachen, klaren, schnellen und sicheren Kaufprozess, durch vielfältige Anwendungsmöglichkeiten des Smart Speakers im Allgemeinen und durch Produktangebote, die sich sinnvollerweise via Voice einkaufen lassen. Diese Erkenntnisse bestätigen somit die dritte Teilfragestellung:

*«Je grösser die Anzahl an Nutzungsmöglichkeiten, die einem Voice-Assistent-Nutzer aufgezeigt werden, desto eher akzeptiert dieser die Technologie.»*

#### 5.1.2 Information und Kommunikation

Dass ein grosser Nutzen vorhanden ist, nützt aber niemandem etwas, wenn dieser von den Nutzenden oder Interessierten nicht erkannt wird. Der zweite wichtige Faktor, der die Akzeptanz von Voice Commerce nachhaltig beeinflusst, ist Information und Kommunikation. Unternehmen müssen offen und transparent informieren, gerade bei Themen wie der Datenverarbeitung. Die Personen, die Voice Commerce nutzen, sollten

im Idealfall genau wissen, welche Daten gesammelt und verarbeitet werden und sich bewusst dafür entscheiden. Auch durch die Medien werden Kunden über Voice Assistants und Voice Commerce informiert. Gibt es auf diesem Gebiet viel negative Berichterstattung, schmälert das das Vertrauen in die Technologie. Information ist auch wichtig beim eigentlichen Kaufprozess. Die Kunden sollen sich hundertprozentig sicher sein, dass sie den Artikel kaufen, den sie möchten, zu dem Preis, der vorgeschlagen wurde. Ist der Smart Speaker mit einem Bildschirm ergänzt, können diese Informationen visuell vermittelt werden, was das Vertrauen und die Sicherheit in den Prozess erhöht.

### 5.1.3 Technologische Entwicklung

Der dritte Faktor, der einen massiven Einfluss auf die Akzeptanz von Voice Commerce in der Schweiz hat, ist die technologische Entwicklung. Das bezieht sich zum einen auf die Spracherkennung, die sich im Vergleich zu heute wohl noch einmal deutlich verbessern wird und in Zukunft auch komplexe Anfragen und Kontext verstehen sowie zwischen verschiedenen Nutzern unterscheiden kann. Für Schweizer Kunden, die möglichst in ihrer natürlichen Sprache mit Voice Assistants interagieren möchten, wäre das Verständnis ihres Dialekts ein starkes vertrauensbildendes Argument. Allgemein, sobald ein Voice Assistant immer «menschlicher» wird, kann davon ausgegangen werden, dass das Vertrauen gestärkt wird. Mit diesen gesammelten Erkenntnissen kann die erste Teilfragestellung *«Je besser die Technologie der Conversational Interfaces in Zukunft wird, desto mehr werden die Nutzer die Technologie akzeptieren und in ihre Anwendungen vertrauen.»* bestätigt werden.

## 5.2 Betrachtung der Schweiz

Die Hauptfragestellung legte ihren Fokus auch auf die Schweiz. In der empirischen Untersuchung hat sich ergeben, dass sich die Schweiz, sollten Voice Assistants im Markt ankommen, nicht von anderen europäischen Ländern unterscheidet, was die Nutzerakzeptanz betrifft. So kann beispielsweise Deutschland als Referenz genommen werden, um zu sehen, in welche Richtung sich Voice Commerce bewegt. Voice Commerce hat somit durchaus Potenzial in der Schweiz, auch wenn es voraussichtlich noch rund fünf Jahre gehen wird, bis der Stand von heute von Voice Commerce in Deutschland erreicht wird.

### 5.3 Beantwortung der Teilfragestellung

Die zweite Teilfragestellung zielte darauf ab, die identifizierten Einflussfaktoren in Kategorien einzuteilen.

*«Einflussfaktoren auf das Vertrauen in Voice Commerce lassen sich in die Kategorien Information, Reputation und Garantie einteilen.»*

Diese Teilfragestellung kann nicht abschliessend beantwortet werden. So wurden wie eingangs erwähnt Faktoren identifiziert, die sich in die Kategorie Information einteilen lassen. Auch Faktoren im Bereich Reputation haben sich im Verlauf der Arbeit herauskristallisiert, beispielsweise dass grosse, bekannte Unternehmen mehr Chancen mit einer Voice-Applikation haben als unbekannte Anbieter. Im Bereich der Garantie hingegen konnten in dieser Arbeit keine eindeutigen Faktoren identifiziert werden, was jedoch nicht heisst, dass es diese nicht gibt. Diese Teilfragestellung bleibt darum auch nach dieser wissenschaftlichen Arbeit noch offen. Die Verfasserin geht jedoch davon aus, dass es auch Einflusskriterien auf das Vertrauen in Voice Commerce in dieser Kategorie gibt, da das Kategorienmodell für alle digitalen Transaktionen aufgestellt wurde und darum auch kategorisch für Voice Commerce gelten sollte.



## 6 Diskussion und Reflexion

Die vorliegende Bachelorarbeit hat sich mit dem Thema befasst, welche Faktoren die Nutzerakzeptanz von Voice Commerce in der Schweiz beeinflussen. Dazu wurden eine breite Literaturanalyse sowie empirische Experteninterviews durchgeführt. Die Ergebnisse aus beiden Forschungsmethoden zeigen, dass Voice Commerce durchaus Potenzial hat, auch in der Schweiz. Hierbei muss betont werden, dass viele Erkenntnisse aus dieser Arbeit auf Einschätzungen und Voraussagen beruhen, die sich erst nach einigen Jahren effektiv belegen lassen. Durch die Analyse von Voice Commerce in anderen Ländern konnte dieser Umstand jedoch etwas abgefedert werden. Ausserdem konnte durch qualifizierte Experten ein umfassenderer Blick auf die Schweiz geworfen werden, was nur durch die reine Literaturanalyse nicht möglich gewesen wäre. Somit ergänzt diese Arbeit den bereits vorliegenden Wissensstand und kann die Wissenslücke schliessen.

Für diese wissenschaftliche Arbeit konnten zwei Experten und eine Expertin gewonnen werden, die jeweils einen anderen Hintergrund hatten und auf unterschiedliche Gebiete spezialisiert waren. Dadurch konnte das ganze Wissensfeld abgedeckt werden. Auf der anderen Seite gab es so teilweise keine Zweitmeinung zu einem Thema oder einer Aussage. Bei einer weiterführenden Forschung sollte darum das Feld der Experten ausgeweitet und mehr Personen interviewt werden.

Die Grenzen dieser Arbeit liegen darin, dass keine effektiven Handlungsempfehlungen an Unternehmen abgegeben werden. Die Arbeit befasst sich vor allem mit dem theoretischen Aspekt der Vertrauensbildung. Um konkrete Handlungsempfehlungen abgeben zu können, hätte sich die Verfasserin auf bestimmte Unternehmen oder Branchen fixieren müssen. Da jedoch der Anspruch bestand, eine möglichst generalisierbare Forschung zu betreiben, wurde auf Handlungsempfehlungen verzichtet.

Eine weitere Limitation ist, dass auf die Befragung von Nutzern verzichtet wurde. Hauptgründe dafür waren, dass es in der Schweiz nur eine kleine Anzahl von Nutzern gibt und es sehr schwierig wäre, diese anzusprechen. Die Befragung von potenziellen Nutzern hingegen erschien zu unpräzise, da viele Personen in der Schweiz noch kaum in Kontakt mit Voice Assistants, geschweige denn Voice Commerce, gekommen sind. Durch das Fehlen einer solchen Befragung wurde ein verminderter Praxisbezug bewusst in Kauf genommen. Weiterführende Forschung sollte diesen Umstand einbeziehen und verstärkt einen Fokus auf die Nutzenden legen. Dabei muss jedoch berücksichtigt werden, ob Voice Commerce in der Schweiz bereits etabliert ist oder nicht.





## 7 Quellenverzeichnis

- Accenture. (2018). *The Continued Rise of Digital Voice Assistants*. Zugriff am 10.06.2019. Verfügbar unter <https://www.accenture.com/us-en/insights/software-platforms/continued-rise-voice-assistants>
- Aghamanoukjan, A., Meyer, M. & Buber, R. (2009). Qualitative Interviews. In R. Buber & H. H. Holzmüller (Hrsg.), *Qualitative Marktforschung. Konzepte - Methoden - Analysen* (Lehrbuch, 2., überarb. Aufl., S. 415–436). Wiesbaden: Gabler.
- Amazon. (2018). *Echo & Alexa-Geräte - Amazon.de*. Zugriff am 05.07.2019. Verfügbar unter <https://www.amazon.de/b?ie=UTF8&node=14100226031>
- Baumgartner, A. & Dalla Vecchia, M. (2017). Alexa bestellt Milch. *UnternehmerZeitung*, 11, 30–31. Zugriff am 24.04.2019. Verfügbar unter <https://web.fhnw.ch/plattformen/blogs/iwi/2017/10/30/voice-commerce/>
- Bogner, A., Littig, B. & Menz, W. (2014). *Experteninterviews. Eine Einführung in Theorie und Praxis einer Forschungsmethode* (Qualitative Sozialforschung, Aufl. 2014). Wiesbaden: Springer Fachmedien Wiesbaden GmbH. <https://doi.org/10.1007/978-3-531-19416-5>
- Brocke, J. vom, Simons, A., Niehaves, B., Reimer, K., Plattfaut, R. & Cleven, A. (2009). RECONSTRUCTING THE GIANT: ON THE IMPORTANCE OF RIGOUR IN DOCUMENTING THE LITERATURE SEARCH PROCESS. *ECIS 2009 Proceedings*. Verfügbar unter <https://aisel.aisnet.org/ecis2009/161>
- Bröcker, F. (2018). *Die Anwendung der Diffusionstheorie nach Rogers auf die Adoption von Brain-Computer-Interfaces (BCI) bei Exoskeletten*. Seminararbeit. Universität Karlsruhe (TU). Zugriff am 16.07.2019. Verfügbar unter <https://www.grin.com/document/427610>
- Bruhn, M. & Hadwich, K. (Hrsg.). (2018). *Service Business Development. Strategien - Innovationen - Geschäftsmodelle : Band 1*. Wiesbaden, Germany: Springer Gabler. <https://doi.org/10.1007/978-3-658-22426-4>
- Bucher, U. (11.06.2019). *Voice Commerce in der Schweiz* (Interview).
- Budzinski, O., Noskova, V. & Zhang, X. (2018). *The Brave New World of Digital Personal Assistants: Benefits and Challenges from an Economic Perspective*. Technische Universität Ilmenau, Ilmenau. Zugriff am 06.07.2019. Verfügbar unter <http://hdl.handle.net/10419/191021>
- Capgemini Digital Transformation Institute. (2018). *Conversational Commerce. Why Consumers Are Embracing Voice Assistants in Their Lives*. Zugriff am 09.05.2019. Verfügbar unter <https://www.capgemini.com/resources/conversational-commerce-dti-report/>
- Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science*, 35(8), 982–1003.
- Deloitte. (2018). *Beyond Touch – Voice Commerce 2030. Wie Voice-assisted Interfaces den Handel in Europa revolutionieren werden*. Zugriff am 19.07.2019. Verfügbar unter <https://www2.deloitte.com/de/de/pages/consumer-business/articles/sprachassistent.html>
- EARSandEYES GmbH. (2018, April). *Gründe für die Nichtnutzung von Sprachassistenten in Deutschland 2018* (Statista, Hrsg.). Zugriff am 17.04.2019. Verfügbar unter <https://de-statista-com.ezproxy.fh-htwchur.ch/statistik/daten/studie/872316/umfrage/gruende-fuer-die-nichtnutzung-von-sprachassistenten-in-deutschland/>
- Eeuwen, M. van. (2017). *Mobile conversational commerce: messenger chatbots as the next interface between businesses and consumers*. Essay (Master). University of Twente, Enschede. Zugriff am 08.07.2019. Verfügbar unter <http://purl.utwente.nl/essays/71706>

- Eggs, H. (2001). *Vertrauen in Electronic Commerce: Herausforderungen und Lösungsansätze* (1. Auflage). Wiesbaden: Deutscher Universitäts-Verlag GmbH.
- Elster, M. (2019). Warum ist Voice das nächste große Interface? *marconomy*. Zugriff am 12.04.2019. Verfügbar unter [https://www.wiso-net.de/document/MARC\\_\\_45756808](https://www.wiso-net.de/document/MARC__45756808)
- Evans, C. (2015). *Spike Jonzes neuer Film „Her“*. Zugriff am 10.06.2019. Verfügbar unter <https://www.spikeartmagazine.com/de/artikel/du-bist-so-vollkommen-cecile-b-evans-uber-spike-jonzes-film-her>
- Freist, R., Dirscherl, H.-C. & Joos, T. (2019). *Amazon Echo: Die besten Alexa-Skills*, IDG Tech Media GmbH. Zugriff am 06.07.2019. Verfügbar unter <https://www.pcwelt.de/a/die-10-besten-skills-fuer-amazon-alexa,3422718>
- Gartner. (2018). *5 Trends Emerge in the Gartner Hype Cycle for Emerging Technologies, 2018*. Zugriff am 16.04.2019. Verfügbar unter <https://www.gartner.com/smarterwithgartner/5-trends-emerge-in-gartner-hype-cycle-for-emerging-technologies-2018/>
- Gentemann, L. (2018). *Zukunft der Consumer Technology – 2018. Marktentwicklung, Trends, Mediennutzung, Technologien, Geschäftsmodelle* (Bitkom e.V., Hrsg.). Berlin. Zugriff am 10.06.2019. Verfügbar unter <https://www.digitalconnection.de/wp-content/uploads/2018/09/180822-CT-Studie-2018-online.pdf>
- Ginner, M. (2018). *Akzeptanz von digitalen Zahlungsdienstleistungen*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-19706-3>
- Grabner-Kräuter, S. & Kaluscha, E. A. (2003). Ansatzpunkte für die Entwicklung einer „Pattern-Language“ zum Aufbau von Konsumentenvertrauen im Electronic Commerce. *der markt*, 42(3-4), 132–146. <https://doi.org/10.1007/BF03032139>
- Hayllar, W. & Coode, M. (2018). *The talking shop. The rise of voice commerce* (OC&C Strategy Consultants, Hrsg.). Zugriff am 11.04.2019. Verfügbar unter [https://www.occstrategy.com/media/1285/the-talking-shop\\_uk.pdf](https://www.occstrategy.com/media/1285/the-talking-shop_uk.pdf)
- ITWissen.info. (2019). *NLP (natural language processing)*. Zugriff am 02.07.2019. Verfügbar unter <https://www.itwissen.info/NLP-natural-language-processing.html>
- Kaushansky, K. (03.07.2019). *User acceptance in voice commerce* (Skype-Interview).
- Kinsella, B. (2019). *Amazon Alexa Skill Counts Rise Rapidly in the U.S., U.K., Germany, France, Japan, Canada, and Australia - Voicebot*. Zugriff am 06.07.2019. Verfügbar unter <https://voicebot.ai/2019/01/02/amazon-alexa-skill-counts-rise-rapidly-in-the-u-s-u-k-germany-france-japan-canada-and-australia/>
- Kinsella, B. (2019). *Google Home Added 600,000 More U.S. Users in 2018 Than Amazon Echo, But Amazon Echo Dot is Still the Most Owned Smart Speaker - Voicebot*. Zugriff am 07.07.2019. Verfügbar unter <https://voicebot.ai/2019/03/07/google-home-added-600000-more-u-s-users-in-2018-than-amazon-echo-but-amazon-echo-dot-is-still-the-most-owned-smart-speaker/>
- Kinsella, B. & Mutchler, A., Voysis (Mitarbeiter) (voicebot.ai, Hrsg.). (2018). *Voice Shopping Consumer Report. June 2018*. Zugriff am 17.04.2019. Verfügbar unter <https://voicebot.ai/research/>
- Kinsella, B. & Mutchler, A. (voicebot.ai, Hrsg.). (2019). *Smart speaker consumer adoption report. March 2019*. Zugriff am 17.04.2019. Verfügbar unter <https://voicebot.ai/research/>
- Kittl, C. (2009). *Kundenakzeptanz und Geschäftsrelevanz*. Wiesbaden: Gabler. <https://doi.org/10.1007/978-3-8349-9425-7>
- Kowalczyk, P. (2018). Consumer acceptance of smart speakers: a mixed methods approach. *Journal of Research in Interactive Marketing, Vol.12 Issue: 4*, 418–431.

- Kraus, D., Reibenspiess, V. & Eckhardt, A. (2019, Februar). *How Voice Can Change Customer Satisfaction: A Comparative Analysis between E-Commerce and Voice Commerce*. Siegen. Zugriff am 12.04.2019. Verfügbar unter [https://www.researchgate.net/publication/331357421\\_How\\_Voice\\_Can\\_Change\\_Customer\\_Satisfaction\\_A\\_Comparative\\_Analysis\\_between\\_E-Commerce\\_and\\_Voice\\_Commerce](https://www.researchgate.net/publication/331357421_How_Voice_Can_Change_Customer_Satisfaction_A_Comparative_Analysis_between_E-Commerce_and_Voice_Commerce)
- Kruse Brandão, T. & Wolfram, G. (2018). *Digital Connection. Die bessere Customer Journey mit smarten Technologien - Strategie und Praxisbeispiele*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-18759-0>
- Kuhlen, R. (2008). Vertrauen in elektronischen Räumen. In D. Klumpp, H. Kubicek, A. Roßnagel & W. Schulz (Hrsg.), *Informationelles Vertrauen für die Informationsgesellschaft* (S. 37–51). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Kunath, G., Hofstetter, R., Jörg, D. & Demarchi, D. (2019). *Voice Barometer Schweiz 2018*. Universität Luzern.
- Langer, U. (2019). Alexa, wo bleibt meine Pizza? *Horizont*, (13), 38. Zugriff am 12.04.2019. Verfügbar unter [https://www.wiso-net.de/document/HOR\\_\\_20190328491100%7CAHOR\\_\\_20190328491100](https://www.wiso-net.de/document/HOR__20190328491100%7CAHOR__20190328491100)
- Latzer, M. (2016). *Protokoll Literaturrecherche und Erarbeitung des Forschungsstands. Leitfaden für Studierende an der Abteilung Medienwandel & Innovation* (IPMZ – Institut für Publizistikwissenschaft und Medienforschung, Hrsg.). Zürich: Universität Zürich. Zugriff am 02.06.2019. Verfügbar unter [http://www.mediachange.ch/media/medialibrary/2016/05/Leitfaden\\_Recherche\\_und\\_Literatur%C3%BCbersicht\\_Mai2016.pdf](http://www.mediachange.ch/media/medialibrary/2016/05/Leitfaden_Recherche_und_Literatur%C3%BCbersicht_Mai2016.pdf)
- Lenz-Kesekamp, V. & Weber, T. (2018). Alexa Skills: Welche Chancen und Risiken sind damit verbunden? *Wirtschaftsinformatik & Management*, 10(6), 18–25. <https://doi.org/10.1007/s35764-018-0115-9>
- Lopetrone, R. (05.07.2019). *Nutzerakzeptanz von Voice-Anwendungen in der Schweiz* (Interview).
- Martin, D. (2019). “OK Google – Schreibe Teil 2 des Blogartikels über Sprachassistenten!”, nextlevel Blog. Zugriff am 02.07.2019. Verfügbar unter <https://www.nextlevel.de/blog/ok-google-schreibe-teil-2-des-blogartikels-uber-sprachassistenten>
- Mayring, P. & Brunner, E. (2009). Qualitative Inhaltsanalyse. In R. Buber & H. H. Holzmüller (Hrsg.), *Qualitative Marktforschung. Konzepte - Methoden - Analysen* (Lehrbuch, 2., überarb. Aufl., Bd. 50, S. 669–679). Wiesbaden: Gabler. [https://doi.org/10.1007/978-3-8349-9441-7\\_28](https://doi.org/10.1007/978-3-8349-9441-7_28)
- McTear, M., Griol, D. & Callejas, Z. (2016). *The conversational interface. Talking to smart devices*. Cham: Springer. Retrieved from <http://lib.myilibrary.com/detail.asp?ID=924473>
- Messina, C. (2016). *2016 will be the year of conversational commerce*. Zugriff am 07.07.2019. Verfügbar unter <https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991>
- Narvar. (2018). *Connecting with Shoppers in the Age of Choice. Consumer Report 2018*. Zugriff am 08.07.2019. Verfügbar unter [https://see.narvar.com/rs/249-TEC-877/images/Connecting%20With%20Shoppers\\_Narvar%20Consumer%20Report%20Q1%202018.pdf](https://see.narvar.com/rs/249-TEC-877/images/Connecting%20With%20Shoppers_Narvar%20Consumer%20Report%20Q1%202018.pdf)
- Oppermann, A. (2018). *Did Google Duplex beat the Turing Test? Yes and No.*, <https://www.facebook.com/towardsdatascience>. Zugriff am 02.07.2019. Verfügbar unter <https://towardsdatascience.com/did-google-duplex-beat-the-turing-test-yes-and-no-a2b87d1c9f58>
- Pardes, A. (2019). *Hey Alexa, Why Is Voice Shopping So Lousy?*, WIRED. Zugriff am 10.07.2019. Verfügbar unter <https://www.wired.com/story/why-is-voice-shopping-bad/>

- Petrovic, O., Fallenböck, M., Kittl, C. & Wolking, T. (2003). Vertrauen in digitale Transaktionen. *Wirtschaftsinformatik*, 45(1), 53–66. <https://doi.org/10.1007/BF03250883>
- Rogers, E. M. (2003). *Diffusion of innovations* (5. Aufl.). New York: Free press.
- Sieber, A. (2019). *Dialogroboter. Wie Bots und künstliche Intelligenz Medien und Massenkommunikation verändern*. Wiesbaden: Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-24393-7>
- Sievers, M. (2019). *Consumer Barometer 2/2018 – Thema: Voice Commerce*. KPMG Deutschland. Zugriff am 09.07.2019. Verfügbar unter <https://home.kpmg/de/de/home/themen/2018/08/consumer-barometer-2-2018-voice-commerce.html>
- Stanoevska-Slabeva, K. (2018). Conversational Interfaces — die Benutzerschnittstelle der Zukunft? *Wirtschaftsinformatik & Management*, 10(6), 26–37. <https://doi.org/10.1007/s35764-018-0117-7>
- Stanoevska-Slabeva, K. & Lenz-Kesekamp, V. (2018). *Studie zu Einsatzpotenzialen und Anwendungsbeispielen von Conversational Interfaces* (Institut für Medien- & Kommunikationsmanagement der Universität St.Gallen, Hrsg.). Zugriff am 03.06.2019. Verfügbar unter [http://www.smama.ch/wp-content/uploads/2018/04/Conversational\\_Studie\\_A4\\_gruen\\_v1.pdf](http://www.smama.ch/wp-content/uploads/2018/04/Conversational_Studie_A4_gruen_v1.pdf)
- Statista. (2019). *Welchen Smart-Speaker mit integriertem Assistenten besitzen Sie? Umfrage in Deutschland zu beliebten Smart-Speaker-Marken 2019*. Zugriff am 08.06.2019. Verfügbar unter <https://de.statista.com/prognosen/999790/umfrage-in-deutschland-zu-beliebten-smart-speaker-marken>
- TechTarget. (2017). *Definition Voice Assistant*. Zugriff am 04.07.2019. Verfügbar unter <https://whatis.techtarget.com/definition/voice-assistant>
- Tuzovic, S. & Paluch, S. (2018). Conversational Commerce – A New Era for Service Business Development? In M. Bruhn & K. Hadwich (Hrsg.), *Service Business Development. Strategien - Innovationen - Geschäftsmodelle: Band 1* (S. 81–100). Wiesbaden, Germany: Springer Gabler.
- Weidemann, T. (2019). *Voice Commerce - nur ein Hype oder das nächste große Ding?*, DMEXCO. Zugriff am 10.07.2019. Verfügbar unter <https://dmexco.com/de/stories/voice-commerce-nur-ein-hype-oder-das-naechste-grosse-ding/>

## 8 Anhang

### Rechercheprotokoll

Datenbank	Suchalgorithmus	Einschränkungen	Anzahl Ergebnisse	Anzahl relevante Ergebnisse	Weitere Schritte
ABI Inform /ProQuest	"voice commerce" OR "voice shopping" AND trust	ab 2013, nur Volltext, nur Fachmagazine	37	1	Download der zitierten Quelle im Artikel
	"voice assistant" OR "voice assistant systems" AND trust	ab 2013, nur Volltext, nur Fachmagazine	33	1	Download der zitierten Quelle im Artikel
ACM Digital Library	"smart speaker" AND "acceptance"	ab 2013, nur Volltext, nur wissenschaftliche Zeitschriften	1	1	Download des Volltextes
	"voice commerce"	keine Einschränkungen	1	0	ungeeignete Datenbank
Business Source Premier	"voice commerce" OR "voice shopping" AND acceptance	ab 2013, nur Volltext	11	1	Download des Volltextes
	"user experience" AND "voice" AND "acceptance"	ab 2013, nur Volltext	3	0	
De Gruyter	"conversational commerce" AND trust	nur Volltext	64	0	
	"voice commerce"	keine Einschränkungen	1	0	
econbiz	sprachassistent	2010-2019	28	3	kein Volltextzugang
	"smart speaker"	keine Einschränkungen	4	1	Download des Volltextes
Emerald Journals	"conversational commerce" OR "voice commerce"	ab 2013	3	0	
	"smart speaker"	ab 2013	2	0	klings sehr relevant, leider kein Vollzugriff--> andere Datenbank suchen für Zugriff? Consumer acceptance of smart speakers: a mixed methods approach - Pascal Kowalczuk
OECDiLibrary	"smart speaker"	keine Einschränkungen	7	0	
Taylor and Francis Journals	"voice assistant"	keine Einschränkungen	12	1	relevanter Artikel sehr kurz und oberflächlich, wird nicht genauer betrachtet
Web of Science	"voice commerce"	letzte 5 Jahre	1	1	relevanter Artikel sehr kurz und oberflächlich, wird nicht genauer betrachtet
	"voice assistant*"	nur Volltext	3	1	Volltext
WISO	"smart speaker" AND "commerce"	ab 2013, nur Fachzeitschriften	46	4	2 relevante Artikel, die auf relevante Studien verwiesen haben
	"voice commerce" AND "vertrauen"	ab 2013, nur Fachzeitschriften	14	1	Verweis auf relevante Studie
Springer E-Books	"conversational commerce" AND "vertrauen"	ab 2013, nur Fachzeitschriften	16	2	Verweisen auf bereits bekannte Studien
	"voice commerce"	nur Volltext	18	3	Download des Volltextes
	"voice assistant*"	nur Volltext	48	5	Teilweise Überschneidungen
	sprachassistent AND vertrauen	nur Volltext	16	0	
	"smart speaker" AND vertrauen	nur Volltext	8	0	
	user experience AND vertrauen AND voice	nur Volltext	41	1	Download des Volltextes
	"voice shopping" AND vertrauen	nur Volltext	3	0	
Google Scholar	"natural interaction"	keine Einschränkungen	9	2	Download des Volltextes
	"voice commerce" AND "vertrauen"	ab 2013	3	0	
	"voice commerce" AND trust	ab 2013	20	5	alle relevanten Texte bereits in anderen Datenbanken gefunden und gespeichert
	"voice shopping" AND akzeptanz	ab 2013	6	1	Volltext via Springer heruntergeladen
	"smart speaker" AND akzeptanz	ab 2013	15	0	
Hinweis: Suchanfragen, die zu keinem Resultat führten, werden in dieser Liste nicht aufgeführt.					
Literatur, Artikel und Berichte, die bereits im Vorfeld gefunden wurden oder die von Drittpersonen gesendet wurden, sind in dieser Tabelle nicht ersichtlich.					

## Interviewleitfäden

### Interviewleitfäden

Interviewpartner: Urs Bucher

Datum: 11. Juni 2019

Uhrzeit Beginn: 14 Uhr

Uhrzeit Ende:

#### Einleitung

- Vorstellung Fragestellerin
- Vorstellung des Themas der Bachelorarbeit
- Aufbau des Interviews
- Informationen zum Zeitaufwand, Aufnahme des Gesprächs, Transkription und Auswertung
- Offene Fragen soweit?

#### Warm-up

- Bitte beschreiben Sie sich in einigen Sätzen.
- Wo sind Sie aktuell beschäftigt / was machen Sie beruflich?
- Wie sind Sie das erste Mal in Kontakt mit Voice Assistants gekommen?
- Nutzen Sie selbst ein Voice Assistant System oder haben Sie schon einmal einen Smart Speaker benutzt?
- Wofür haben sie den Assistenten / Speaker benutzt und wie zufrieden waren sie?

#### Hauptteil

##### Projekte

- Sie haben mit der Tamedia Workshops zum Thema Smart Audio durchgeführt. Wie ist es dazu gekommen und was wurde gemacht?
- Was war das Ziel der Workshops?
- Ist es korrekt, dass sie zusammen mit der Uni Luzern am Voice Barometer Schweiz mitgewirkt haben? Wie ist es dazu gekommen?
- Was war ihre Rolle in diesem Projekt?
- Was hat sie an den Resultaten am meisten überrascht?

##### Voice Commerce

- Wie gross sehen sie das Potenzial, dass zukünftig über die Stimme eingekauft wird?
- Wo sehen sie die grössten Herausforderungen in diesem Bereich?
- Welche Voraussetzungen müssten gegeben sein, dass Voice Commerce akzeptiert wird?
- Was am Kaufprozess schreckt die Kunden ab, den Kauf via Voice zu tätigen?

##### Vertrauen und Nutzerakzeptanz in der Schweiz

- Unterscheidet sich die Nutzerakzeptanz von Personen aus der Schweiz gegenüber Personen beispielsweise Deutschland oder England?
- Spielt es eine Rolle bezüglich der Nutzerakzeptanz, ob man einen Einkauf tätigt oder nur Informationen abfragt?
- Wäre es möglich, dass Misstrauen nur vorhanden ist, weil man die Technologie hier noch gar nicht kennt?
- Was begünstigt die Nutzerakzeptanz in der Schweiz?
- Was erschwert sie?
- Was ist ihrer Meinung nach ausschlaggebend, dass Voice-Anwendungen in der Schweiz Erfolg haben können?

##### Schluss

- Denken Sie, dass sich Voice Assistenten und Smart Speaker in der Schweiz im grossen Stil durchsetzen werden?
- Abschliessende Worte?
- Fragen ihrerseits?

## Interviewleitfaden

Interviewpartner: Karen Kaushansky

Datum: 3. Juli 2019

Uhrzeit Beginn: 14 Uhr

Uhrzeit Ende:

### Introduction

- Vorstellung Fragestellerin
- Vorstellung des Themas der Bachelorarbeit
- Aufbau des Interviews
- Informationen zum Zeitaufwand, Aufnahme des Gesprächs, Transkription und Auswertung
- Offene Fragen soweit?

### Warm-up

- Please describe yourself in a few sentences.
- When did you first get in touch with voice assistants / voice technology?
- Do you use a voice assistant or a smart speaker?
- In which situations do you use the voice assistant?
- You are working with voice technology for many years now. How did you get into this?
- You are working at Google as a Conversation Designer. Can you tell me something about this position or your current projects?

### Main Part

#### Trust and user acceptance in voice / Human computer interaction

- Which is the most important criteria for user acceptance in technology in general? Is there one?
- Do you think that people in Switzerland have less acceptance in voice technology than the USA per example?
- Amazon Alexa or Google Assistant are not officially released in Switzerland yet. Is it possible, that the only reason Swiss people do not trust this technology is because they do not know it yet?
- Many people have safety and security concerns in this technology. How can these concerns be minimized?
- Do you think that the more transparent a voice application is, the more trust it gets from users? As an example transparent in their data practice, what data is stored, where is the data stored etc.
- What is the best thing a company can do to increase user acceptance and trust in its voice application?
- What is the worst thing what a company with a voice application can do? Are there any no-go's that lead to trust being lost?
- Do you think that it is more relevant that people have trust in the hardware device (per example Amazon) than in voice application itself? Per example: If a swiss company would create their own smart speaker, would Swiss people automatically have more trust in the technology?

#### Voice Commerce

- Do you think that Voice Commerce has a great potential and is the next big evolution in e-commerce?
- What are the main challenges voice commerce has to face?
- Do you think that people have more trust issues in voice commerce than in other voice applications?
- What conditions must be given for voice commerce to be accepted? Are this the same conditions than for other voice applications?

#### Final questions

- Do you think that voice assistants and smart speakers will succeed on a large scale in Switzerland?
- Do you want to say any concluding words?
- Do you have open questions for me?

## Interviewleitfaden

Interviewpartner: Riccardo Lopetrone, Swisscom

Datum: 5. Juli 2019

Uhrzeit Beginn: 9 Uhr

Uhrzeit Ende:

### Einleitung

- Vorstellung Fragestellerin
- Vorstellung des Themas der Bachelorarbeit
- Aufbau des Interviews
- Informationen zum Zeitaufwand, Aufnahme des Gesprächs, Transkription und Auswertung
- Offene Fragen soweit?

### Warm-up

- Bitte beschreibe dich und deine Funktion bei der Swisscom in einigen Sätzen.
- Wie bist du zu dieser Funktion gekommen?
- Wie bist du das erste Mal in Kontakt mit Voice Assistenten gekommen?
- Beim letzten Mal hast du uns ja gezeigt, welche Assistenten du benutzt. Nutzt du diese auch privat und wenn ja, wofür?
- Wie zufrieden bist du generell mit den Voice Assistenten, die du benutzt?

### Hauptteil

#### Projekte

- Die Swisscom befasst sich im TV-Bereich mit Voice und bietet eine Voice-Funktion zur TV-Steuerung an. Wie ist es dazu gekommen?
- Was waren die Herausforderungen bei der Einführung der Funktion?
- Was lief gut bei der Einführung der Funktion?
- Wurden die Funktionen komplett selbst entwickelt oder wurden Kooperationen eingegangen?
- Wie ist das Kundenfeedback auf diese Funktion?
- Wie habt ihr das Kundenfeedback eingeholt? Konntet ihr die Nutzerakzeptanz messen / erfassen?
- Wisst ihr, wie häufig die Funktion genutzt wird und welche Anfragen beispielsweise gemacht werden?
- Gibt es weitere eingeführte Projekte im Voice-Bereich bei der Swisscom?
- Wohin möchte sich die Swisscom auf dem Gebiet Voice entwickeln?
- Wie steht die Swisscom zu den Plattformen von Amazon und Google? War es mal ein Thema, beispielsweise einen Alexa Skill zu entwickeln?
- Was würde es für die Swisscom bedeuten, wenn Amazon oder Google die Smart Speaker offiziell launchen würden?
- Wie betreibt die Swisscom Research auf dem Gebiet Voice?



### Vertrauen und Nutzerakzeptanz in der Schweiz

- Wie versucht die Swisscom, die Nutzerakzeptanz zu erhöhen? (nicht nur auf Voice, sondern allgemein, beispielsweise auch 5G etc.)
- Was bietet die Swisscom den Kunden an, die unsicher oder skeptisch sind? Gibt es Massnahmen?
- Denkst du, dass die Personen in der Schweiz weniger aufgeschlossen gegenüber neuen Technologien sind?
- Denkst du, die Personen in der Schweiz sind gegenüber Voice nur misstrauisch, weil sie es noch kaum kennen? Und sobald sie zb. Alexa offiziell kaufen können, verfliegt das Misstrauen?
- Wie minimiert die Swisscom Sicherheitsbedenken in ihre Technologien? Oft ist das der häufigste Grund für die Nichtnutzung.
- Denkst du, dass Transparenz der Schlüssel zum Erfolg ist? Je transparenter eine Anwendung, desto besser?
- Denkst du, dass die Schweizer Kunden einem Schweizer Unternehmen wie der Swisscom automatisch mehr vertrauen, als den amerikanischen Hardware-Herstellern und deren Anwendungen?
- Was ist deiner Meinung nach ausschlaggebend, dass Voice-Anwendungen in der Schweiz Erfolg haben können?
- Denkst du, dass das Misstrauen bei Voice Commerce höher ist als bei anderen Voice-Anwendungen?

### Schluss

- Denkst du, dass sich Voice Assistenten und Smart Speaker in der Schweiz im grossen Stil durchsetzen werden?
- Abschliessende Worte?
- Fragen deinerseits?

## Interviewtranskripte

### Interviewtranskript Urs Bucher

Dienstag, 11. Juni 2019, Café Bovelli, Zürich  
Start: 14 Uhr  
Ende: 14:45 Uhr

#### Warm-up

**Bitte beschreibe dich in einigen Sätzen.**

Ich bin seit den 80-ern in der Informatik unterwegs, zweite Hälfte der 90er bei Microsoft gewesen und dort den Partnerkanal mitaufgebaut, dann 2000-2007 bei Namics die Zürcher Filiale aufgemacht, Marketing und Verkauf dort aufgebaut, dann 2007 bis 2014 bei KPMG für alles zuständig gewesen, was selber gemacht wurde im Online-Bereich, dann vier Jahre Webagentur geleitet, die heisst Amazelabs, die macht Drupal-Open-source-Projekte und jetzt seit letztem März bis Ende April 2019 bei Farner Consulting gewesen und dort das Lab aufgebaut, bei dem wir uns um Themen gekümmert haben wie beispielsweise Voice.

**Warst du bei Farner nur für das Lab zuständig und deine Arbeit war nach dessen Aufbau erledigt?**

Kann man so sehen, wir waren uns aber auch nicht so einig, wie schnell dass das ein Business generieren soll und dann habe ich dann aufgehört. Das Lab geht aber weiter.

Uns sonst zu mir: Ich wohne in Zürich, bin verheiratet, Tochter ist 12, esse gerne gut, koche gerne, bin in Zürich in einer Zunft und bin bei TedX Zürich seit Anfang dabei, also als Gründungsmitglied. Selber darf ich als Organisator keine Vorträge halten, aber ich habe letztes Jahr TedxZürich moderiert den ganzen Tag lang und mache das dieses Jahr auch wieder.

**Wo bist du aktuell beschäftigt?**

Ende April habe ich bei Farner aufgehört und schaue mich jetzt um, was ich neues machen könnte.

**Wie bist du das erste Mal in Kontakt mit Voice Assistants gekommen?**

Ich hatte schon immer eine gewisse Affinität zu Gadgets. Ich meine es war 2016 als Google den Assistant lanciert hat, da liess ich mir von einem Mitarbeiter von Amazelab einen solchen mitbringen. Das war die erste Generation. Ich habe dann die übliche Annäherung gemacht, zuerst einige Scherzfragen, Sachen ausprobiert, Spotify verknüpfen, einfach ein bisschen herumgespielt und ausprobiert, was es damals schon konnte. Und nacher auch mal mit Amazon Alexa herumgespielt, mit Siri auch ein bisschen. Als es bei Farner ums Lab ging war Voice eine der Technologien, bei der wir das Gefühl hatten, und bei der ich auch jetzt immer noch das Gefühl habe, dass es Marketing und Verkauf massiv beeinflussen wird und darum habe ich mich dort auch weiter reingestürzt und mein Wissen vertieft. Beispielsweise was es kann, wie es aussieht, wer die Hersteller sind, alles was man sich vorstellen kann.

**04:05 Du hast erwähnt, dass du selbst den Google Assistant benutzt. Benutzt du noch andere oder hast du auch andere ausprobiert?**

Nein, ich bin dem Google Assistant relativ treu geblieben, weil ich das Gefühl habe, dass es offener ist als Alexa. Mit Cortana habe ich mal rumgespielt aber das gibt es ja mittlerweile auch nicht mehr. Solche Exoten wie Bixby usw. habe ich nicht angeschaut. Und spannend finde ich, dass jetzt die Zeit kommt, in der Sprachassistenten nicht nur von Google und Amazon selbst in Geräte gebracht werden, sondern sie verkaufen sie an andere, die sie integrieren. Und das geht von Auto bis zum smarten Spiegel im Badezimmer. Es gibt also wirklich eine ganze Bandbreiten von Sachen, die den Assistant nun enthalten. Das finde ich viel spannender.

**05:25 Und wie sieht es jetzt aus mit konkreten Anwendungen, die du nutzt? Auf die Schweiz bezogen ist es ja immer noch schwierig, oder?**

Erstens sind weder Alexa noch der Google Assistant in der Schweiz offiziell gelauncht, und darum ist es relativ schwierig mit lokalen Sachen weil es diese einfach nicht gibt. Aber wofür ich den Google Assistant brauche ist relativ simpel: News anhören, im Auto Google Maps, Siri zum Telefonieren und zuhause die Verknüpfung des Assistants mit Spotify. Und mit Sonos, das gibt es ja bald auch mit Amazon Alexa. Sie haben zwar vor zwei Monaten oder so angekündigt, dass sie offiziell einen Assistant nebst Alexa auf die kleinen Sonos-Boxen lancieren, aber man kann immer nur einen betreiben. Sehr wahrscheinlich wird Alexa dann rausgeworfen.

## Hauptteil

**06:29 Ich bin ja auf dich gekommen, weil du mit der Tamedia einen Workshop gemacht hat. Dort etwas mehr dazu, wie ist es dazu gekommen, was wurde gemacht, was hast du gemacht?**

Ich habe, als wir das Lab aufgemacht haben bei Farmer, da hat mich der Chef von Sebastian (Danielsson), der Marc Gisler, den kenne ich relativ gut und lange, und der hat mitbekommen dass das ein Thema ist, mit dem wir spielen, also Voice. Wir haben dann. 2-3 mal miteinander geredet, und Sebastian kam dann irgendwann auch dazu. Sie haben dann gesagt, dass sie gerne mit einer Gruppe von Leuten aus der Tamedia zusammengesetzt von Verlagsseite und Redaktionsseite, vor allem der Bezahlmedien, da wollten sie gerne einen Workshop machen, um den Leuten zu zeigen, was das Ding ist und was es überhaupt schon kann, wo wir in der Schweiz stehen. Dann war aber auch die konkrete Aufgabe, Projektideen zu generieren aus dem Workshop heraus. Mit diesen Projektideen sind dann Sebastian und Marc zur Geschäftsleitung und haben diese Ideen präsentiert. Ich habe jetzt eine Weile lang nichts mehr gehört von ihnen, aber ich nehme an, dass sie daran arbeiten, die eine oder andere Idee als Pilot umzusetzen.

**07:47 Ihr habt im Workshop die Ideen also gemeinsam erarbeitet.**

Genau es sind, Irrtum vorbehalten, 7-8 Ideen gewesen, die wir hatten. Es war recht spannend, durch das, dass es ein gut gemischtes Feld war von Verlagsleuten mit kommerziellen Interessen und Redaktionsleuten, die eher am Inhalt interessiert waren. Die Ideen, zu denen wir am Schluss gekommen sind, wurden dann von Marc und Sebastian irgendwann im März in einem kurzen Slot an der GL-Sitzung vorgestellt und sie konnten so das okay abholen für 1 oder 2 oder 3 Ideen. Da weisst du mittlerweile ja mehr als ich.

**08:40 Wie muss ich mir das vorstellen, wart ihr da einen Tag lang dran oder eine Woche?**

Ja wir haben einen Tag gemacht, ich habe den ganzen Workshop vorbereitet mit Market Research, was es in der Schweiz schon gibt, wer die Player sind, was es in Amerika gibt, was es in England schon gibt und in Deutschland. Der Google Assistant interessiert sie primär, Alexa eher weniger, und in diesen Ländern wurde geschaut, was es im Medienbereich bereits für Angebote gibt, die man als Inspiration brauchen könnte. Dann haben wir geschaut, was bei denen läuft in den einzelnen Medien. Themen, die man per Voice vermitteln könnte, die nicht allzu tief oder allzu lang sind, nicht allzu schwierig sind. Da haben sich einige Sachen herausgeschält während des Tages, und entsprechend haben wir die Vorschläge gemacht. Etwas war: Immer in der Sonntagszeitung einmal pro Woche schreibt Peter Schneider die Kolumne: Stilfrage, darf man das? Und diese wären sozusagen gemacht, dass man die eingeschickten Leserfragen vorliest und seine Antwort dann als Antwort vorlesen lässt. Da könnte man eine schöne Serie draus machen, man könnte Leseabende daraus machen, Podcasts. Solche Ideen sind ein paar entstanden, die mit relativ vernünftigem Aufwand umgesetzt werden könnten. Welche das sie jetzt machen, wie sie die machen, wann sie die machen, da musst du Sebastian fragen.

**Das Ziel war also wirklich die Ideengeneration und ein Blick in die Zukunft zu werfen. Wurde es auch konkret?**

Schlussendlich ist es schon sehr konkret geworden, wir haben relativ fest ausdetailliert, was die Ideen wären, sind das Menschen, eine synthetische Sprache, was ist die Frequenz, könnte man es monetarisieren, wenn ja wie, geht es zb in den Tagepass hinein, den man für 2 Franken lösen kann beim Tagesanzeiger. Es ist also schon relativ fest ins "Lebige" gegangen, aber natürlich nur ansatzweise, so weit wie man in einem Tag halt kommt. Die Umsetzung wurde aber bewusst aussen vor gelassen, in dieser Phase des Projektes würde man sich nur zu sehr im Weg stehen wenn man schon zu stark über die Umsetzung nachzudenken beginnt.

**Ist es korrekt, dass du bei Farner mit der Uni Luzern am Voice Barometer Schweiz mitgewirkt hast? Wie ist es dazu gekommen? Warst du beteiligt?**

Ja, wir haben diese Studie zusammen mit Reto Hofstetter von der Uni Luzern gemacht, das stimmt so. Zu den Anfängen: Wenn mich nicht alles täuscht hat uns eine Mitarbeiterin von Reto angefragt, ob wir nicht interesse hätten, das miteinander zu machen. Und dann haben wir, weil wir das Thema spannend fanden und weil es ein Thema war, das wir im Lab hatten, das miteinander gemacht. Wir haben geholfen beim Setup der Studie, bei den Fragen, wir haben das ganze zu einem Teil mitfinanziert. Die Umfrage selbst hat das Link-Institut gemacht, ich glaube das steht auch auf der Studie drauf. Das war eine repräsentative Studie mit irgendwie 1100 Teilnehmer. Sie haben die inhaltliche Aufbereitung gemacht und wir haben die Kommunikation dazu gemacht.

**13:00 Was war deine persönliche Rolle in diesem Projekt? Wofür warst du verantwortlich?**

Alle zwei Wochen mal draufgeschaut, ich war relativ am Rande beteiligt.

**Was hat dich an den Resultaten am meisten überrascht?**

Die Studie ist ja letzten Herbst herausgekommen, ziemlich zeitgleich mit dem Launch vom Farner Lab im November. Was uns alle und mich auch erstaunt hat, war, dass 37% vom repräsentativen Sample, von diesen 1100 Leuten gesagt haben, dass sie bereits täglich einen Sprachassistenten brauchen. Das wär natürlich der grösste Teil Siri und dann auch Alexa und

der Google Assistant, die ja beide noch nicht gelauncht sind in der Schweiz. Das ist eine massiv hohe Zahl, vor allem in Anbetracht dessen, dass eben diese zwei Produkte von Amazon und Google noch nicht gelauncht sind. Das hat mich massiv erstaunt. Dann war auch spannend: Der Altersunterschied zwischen denen, die schon einen Sprachassistenten nutzen und denen, die keinen nutzen, war ungefähr ein Dreivierteljahr. Im Durchschnitt. Die Nutzer, ich glaube 46 oder 47 Jahre alt, und die Nichtnutzer sind im Schnitt ein Dreivierteljahr älter. Sprich es ist eigentlich null Unterschied. Das habe ich super spannend gefunden. Und was ich als Bestätigung der Resultate vom Schweizer Voice Barometer spannend finde: In Deutschland hat es auch eine Studie gegeben, die letzte Woche erschienen ist. Die hat Business Insider letzte Woche veröffentlicht. Und diese Studie sagt auch, dass in Deutschland ein Drittel von allen Personen die in der Studie waren, täglich entweder Siri, Alexa oder Google Assistant. Das war für mich eine schöne Bestätigung, dass das, was wir in der Schweiz herausgefunden haben, nicht ganz neben den Schuhen sein kann, sondern dass das die Realität ist heutzutage.

**16:20 Spielt es eine Rolle bezüglich der Nutzerakzeptanz, ob man einen Einkauf tätigt oder nur Informationen abfragt?**

Ich denke, es ist ein massiver Unterschied ob man nur Informationen abfragt oder ob man etwas kauft, denn bei einem Kauf gibt man Geld aus. Und der Schweizer Konsument bei Geld ausgeben, da ist man gerne ganz sicher, dass es wirklich nur 37.55 Franken sind und nicht 3700.55 Franken. Das macht einen massiven Unterschied.

**Und von der Akzeptanz her? Vielfach ist der Kaufprozess ja über Voice eingeleitet, liegt es wirklich nur am Abschluss oder beginnen die Leute schon früher, dem Ganzen etwas zu misstrauen?**

Wenn ich mir überlege, wie ich Vertrauen gefasst habe in einen Voice Assistenten und wie ich das auch gesehen habe bei allen Personen, denen ich ein Google Home Mini gegeben habe oder die im Lab hatte: Man startet mit irgendwelchen Scherzfragen (wer ist dein Vater, was ist für Wetter, wo bist du), dann bewegt man sich zu Abfragen, zb des SBB-Fahrplans, News vom Tag, allgemein nützlichen Sachen. Dann beginnt man, das Ganze mit Home Automation zu verbinden, also zb mit dem Licht oder eben mit Spotify. Und in diesem Prozess beginnt man langsam, dem Ding zu trauen. Und so wird nach meiner eigenen "empirischen Forschung" sozusagen das Vertrauen aufgebaut. Und wenn man einmal an diesem Punkt angekommen ist, ist man bereit oder würde in Erwägung ziehen, etwas per Voice zu bestellen.

**18:18 Also nicht gleich von 0 auf 100?**

Nein, man muss sich wirklich vorarbeiten. Es ist eigentlich wie eine Treppe, die man abläuft. Und dann kommt es meiner Ansicht nach auch sehr darauf an, wie affin man zur Technologien ist. Also wenn man sonst schon bei LeShop, Coop oder Farny, bei Digitec online Sachen bestellt, dann ist die Wahrscheinlichkeit, dass man das auch mit Sprache macht, relativ gross. Wenn man sonst online eher nicht unterwegs ist und das Vertrauen nicht hat, dass zb die Kreditkartendaten sicher sind, dann kann man das ganze sowieso gleich vergessen.

Was ich auch glaube: es gibt eine ähnliche Treppe zum Vertrauen schaffen bei Voice Commerce wie generell mit einem Voice Device um Vertrauen zu fassen. Man hat zuerst beispielsweise die Bring-App, auf der man Items hinzufügt per Voice, und dann beginnt man, wenn überhaupt, diese via Voice zu bestellen. Aber eben, das kann man ja in der Schweiz noch gar nicht wirklich. Da sind wir auch wieder bei dem Huhn-Ei-Problem. Im ersten Schritt würde man dann beginnen, Produkte für kleinere Beträge in den Basket zu bestellen. Das sind diese Schritte: Zuerst öffnet man die App, legt dann das Produkt in den Basket, und erst wenn man sich ganz sicher ist, dann wird man die Bestellung auch per Voice abschliessen. Aber ich

wüsse im Moment keinen Shop in der Schweiz, bei dem man per Voice bestellen kann. Was auch spannend ist: es gibt eine Studie von PWC in Amerika zum Thema Voice Commerce von letzter Woche. Auf dieser sieht man wunderschön, dass bei kleinen Beträgen die Akzeptanz um Voice Commerce anzuwenden hoch ist, und je grösser der Betrag wird, umso kleiner wird die Bereitschaft, etwas zu bestellen. Und ich denke mir, dass das in der Schweiz genau so sein wird.

#### 21:09 Was begünstigt die Nutzerakzeptanz in der Schweiz?

Grosse Player und Brands, die das Vertrauen mitbringen. Wenn Migros, Coop, Digitec, Farny, Leshop, Marken mit Onlineshop, bei denen man dem Brand vertraut, dann macht man es eher als beim Nischenhändler XYZ, von dem man noch nie etwas gehört hat und auch nicht weiss, wer das jetzt ist und woher die Sachen gehen. Es braucht die, die vorauslaufen, um das Vertrauen zu schaffen. Als Phase 0 wäre es auch Sache der Medienhäuser wie Tamedia, Informationen bereitzustellen, um dort den Boden zu ebnet und dann können die, die etwas verkaufen wollen, darauf aufsetzen.

#### 22:03 Was ist die Schwierigkeit? Im Bereich der Nutzerakzeptanz?

Schlussendlich die Unsicherheit der Leute mit der Technologie. Dann das, dass es noch keine grossen und bekannten Anbieter gibt in der Schweiz, die etwas machen. Und darunter geht natürlich auch, dass es weder den Assistant noch Alexa in der Schweiz offiziell schon gibt. Sprich, weder Amazon oder Google haben irgendwelche Anstrengungen gemacht, ihre Produkte hier in den Markt zu drücken. Ich bin ziemlich genau vor einem Jahr in Hamburg gewesen als Google den Assistant gelauncht hat. Die haben Hamburg zugepflastert mit Plakaten. Das ist schon noch ein anderer Marktdruck als "hmm ja es gibt vielleicht etwas aber wir wissen eigentlich nicht genau was", was wir im Moment in der Schweiz haben. Und was da auch nicht hilft, da du ja noch zu Swisscom gehst, sind so Geschichten wie die Stimmerkennung, wo sogar Unterschriften gesammelt werden dafür dass man das nicht mehr machen darf. Die Post macht genau das gleiche, und solche Geschichten helfen der Technologieakzeptanz per se auch nicht. Und für den 0815-Benutzer ist das das gleiche, es geht um die Sprache. Ob es jetzt Stimmerkennung ist oder Voice Commerce, da muss man den Leuten schon relativ viel erklären.

#### 23:38 Zusammenfassend also: Damit es Erfolg haben kann braucht es die grossen Player, ein offizieller Verkauf der Geräte und das am besten natürlich zeitgleich?

Stell dir einmal vor, das würde diesen Herbst passieren.

**Ich denke, als ersten gäbe es natürlich die Early Adaptors, die nur darauf gewartet haben.**

Genau, die Early Adopters bis Laggards, in der Gartner-Terminologie. Was sicher auch spannend ist, wie du dein Feld noch etwas ausbreiten kannst, damit es nicht nur um Commerce geht, das ist der ganze Bereich vom Marketing. Wie mache ich Marketing im Bereich Voice, was muss ich verändern beim Ranking. Analog wenn man auf dem Telefon oder am PC etwas sucht auf Google, dort gibt es ja die Search Engine Result Page mit irgendwie 10 Resultaten, wenn man aber den Voice Assistant fragt, welches die beste Pizzeria ist in der Umgebung, dann wird nur ein Resultat ausgegeben. Was muss also der Werbetreibende oder der Pizzeriaeinhaber machen, damit er dort raufkommt? Wie kann /muss / soll man Werbung machen auf Voice? Macht man beispielsweise Werbung auf Spotify?

Macht man Werbung beim Blick auf der Voice-Seite? Bei der Zeitung? Wie muss man die Technik im Hintergrund anpassen, damit die SEO-Massnahmen für Voice greifen? Im Moment haben wir wieder die Phase, in der es Firmen gibt, die Leuten versprechen, dass sie auf den ersten Platz kommen wenn man ihnen Geld gibt, dass sie so optimiert werden. Das ist wie früher mit SEO für Webseiten, das kann man nicht machen. Es gibt best practices, es gibt Sachen, die man machen kann wie die Voice Snippets, die man richtig abfüllen muss, Meta-Tags, die man haben muss, die ganzen Informationen, die man haben muss zum Text dazu, damit es maschinenlesbar ist. Und dann hat man eine realistische Chance, dass dein Brand oder dein Produkt zu oberst genannt wird. Aber garantieren kann dir das niemand. Und das ist auch hier sehr früh, es bewegt sich sehr viel auf der Anbieterseite, es verändert sich alle zwei bis drei Monate wieder, und es hat noch relativ viel Platz für irgendwelche Schlaumeier die meinen, dass es jeden Tag wieder Leute gibt, denen man das Geld aus der Tasche ziehen kann, wenn man ihm verspricht, dass man ihn auf den ersten Platz bringt, was einfach nicht realistisch ist.

Und generell zum Marketing verändert der Punkt Voice auch wieder, dass man sich als Brand, vor allem als grössere Firma, überlegen muss, welche Stimme mein Brand hat. Mastercard hat beispielsweise vor zwei oder drei Monaten die Brand Personality ihres Brands kommuniziert. Und zwar von allem, von der Stimme über den Jingle, über 30-Sekunden-Spots, über 3-Minuten-Musik. Das ist super spannend, denn so hat mein einen stimmigen Brand vom Aussehen her, online und eben auch von der Sprache her.

**27:50 Wäre es möglich, dass Misstrauen nur vorhanden ist, weil man die Technologie hier noch gar nicht kennt?**

Ja, denn man baut das Vertrauen ja auf, genau gleich wie man das Vertrauen aufbaut in, aktuell Beispiel, selbstfahrende Fahrzeuge. Die Leute kennen es noch nicht, das ist so. Obwohl ja auch dagegen spricht, dass bereits gemäss der Studie ein Drittel der Leute sagt, dass sie es schon brauchen. Aber ich glaube es geht wirklich darum, dass die Leute in der Masse Vertrauen gewinnen, also der Durchschnitt, der Average, der zwischen dem Early Adopter und dem Laggard. Der muss der Technologie vertrauen und das Vertrauen aufbauen, über Scherzfragen, Sachabfragen, Verknüpfen mit anderen Diensten, und dann glaubt er langsam daran, dass das Ding vertrauenswürdige Resultate liefert. Und gleichzeitig nicht die ganze Zeit zuhört, das ist ja der andere Punkt. Und da ist meine Standardantwort immer: die Dinger haben entweder einen Knopf zum Abschalten oder man kann den Strom ausziehen und dann ist man ganz sicher.

**29:17 Unterscheidet sich die Nutzerakzeptanz von Personen aus der Schweiz gegenüber Personen beispielsweise Deutschland oder England?**

Zu England kenne ich mich zu wenig aus, da kann ich nichts dazu sagen. Aber in Deutschland scheint es mir ähnlich zu sein wie in der Schweiz. Der Proof dazu ist eben diese Studie, die etwa das gleiche aussagt. Ob es jetzt ein Drittel ist oder 37% macht ja nicht viel aus, das ist nur ein kleiner Unterschied.

**29:44 Man kann also nicht sagen, dass die Schweizer beispielsweise vorsichtiger sind oder Datenschutz höher gewichten?**

Sehr wahrscheinlich nicht. Da ist der Unterschied zwischen Deutschland und der Schweiz auch nicht gross. Ich denke mir aber, ohne es zu wissen, dass der Unterschied zwischen Amerika und der Schweiz im Bezug auf Sicherheitsbedenken gross bis sehr gross ist. Und

dann England und Schweiz wohl auch gross. Aber das kann ich nicht mit irgendwelchen Studien unterlegen.

Und wo es wirklich ans Eingemachte geht, ist dann, wenn man sich überlegt, dass es auch Voice Banking gäbe in Amerika, also dass sich Leute Geld per Voice Commerce überweisen. Da denke ich in der Schweiz in absehbarer Zeit eher weniger daran. Anders rum gesagt: eher in der Generation, die Twint braucht als die, die es nicht nutzen. Ich glaube von Informationen abfragen und Voice Commerce ist es noch ein sehr grosser Schritt zu Voice Banking, das ist noch sehr weit weg. Aber das ist sehr wahrscheinlich eine Alters- und Generationenfrage.

**32:23 Wie gross würdest du das Potenzial also einschätzen, dass in der Schweiz wirklich mit Voice eingekauft wird? Mal angenommen, die Player wären vorhanden?**

Ich denke, dass es für 0815-Sachen, Convenience-Sachen mit tiefem Preis, dass das funktionieren kann. Aber sobald es teurer wird und das Risiko grösser wird, dass das Geld dann irgendwo in der Kette aus der Sicht des Nutzers stecken bleibt, er also gezahlt hat und es kommt aber irgendwie nicht an, ich glaube da nimmt es dann rapide ab. Es geht wirklich um die Convenience. Schlussendlich ist reden einfacher als schreiben, und eben solche Leute, die z.B. Bring brauchen, um ihre Shoppingliste zu machen, die wöchentliche, und die, die heute schon viel Online shoppen, für die ist es ein kleiner Schritt, das dann auch per Sprache zu machen. Aber das ist wohl noch nicht der Durchschnitt in der Schweiz.

Es gibt auch noch so Indikatoren in Amerika, dort hat man auch relativ grosse Hoffnungen darauf gesetzt, dass Voice Commerce das nächste grosse Ding wird. Aber etwa vor 3 Monaten sind die ersten Resultate erschienen, da waren alle massiv enttäuscht, also die Anbieter, dass die Akzeptanz nicht so hoch war, wie man es eigentlich gerne hätte. Und wenn es dort schon nicht so funktioniert wie es soll, dann ist es hier wohl auch schwierig.

**34:07 Also nicht das nächste Mobile, wie man auch schon gehört hat?**

Nein.

**34:28 Wenn wir jetzt also über die Herausforderungen reden, ist es vor allem eine Preisfrage, so wie ich das jetzt herausgehört habe?**

Ja, das 1. ist Vertrauen, generell in die Technologie, dass man auch etwas ausprobieren möchte. Dann auch in Voice. Dann die Verlässlichkeit und die Einfachheit.

**34:58 Im Umkehrschluss also: Das muss gegeben sein, dass es funktionieren kann? Ja genau.**

**35:07 Wir haben es vorher schon kurz angetönt. Es ist also wirklich der Abschluss im Verkaufsprozess, der der kritische Punkt des ganzen ist?**

Als Anbieter müsste man einen Weg finden, dem Kunden am Schluss des Kaufprozesses möglichst schnell und einfach zu bestätigen, dass seine Bestellung angekommen ist und ausgeführt wird. Das kann etwas simples einfaches sein wie ein Bestätigungsmail, drei Sekunden nachdem man die Bestellung per Voice gemacht hat. Man macht einen Medienbruch und sieht dann, dass es angekommen ist. Und da müssen wir im Bereich UX noch relativ viel dürfen /müssen/ können forschen, was bei den Leuten am besten ankommt. Ob es reicht, dass man eine Bestellung noch einmal vorliest und eine Transaktions-ID, die



man sich eh nicht merken kann, oder ob es wirklich Sinn macht, den Leuten ein Bestätigungsmail oder Whatsapp oder SMS schickt, irgendetwas, damit man den Proof hat, dass das Ding angekommen ist und die Bestellung ist ausgelöst. Irgend so etwas muss es sein.

**36:17 Aber du denkst, dass es einen Medienbruch braucht? Ist es zu wenig glaubwürdig, wenn Alexa einfach sagt, dass alles funktioniert hat?**

Ich würde sagen: die ersten 10 Mal wäre es wohl gut, wenn es einen Medienbruch gibt. Und nachher traut man dem Ding sowieso. Technisch ist das relativ simpel, das ist kein Beinbruch

## Schlussfragen

**Denkst du, dass sich Voice Assistenten und Smart Speaker in der Schweiz im grossen Stil durchsetzen werden?**

Ja, das steht ausser Diskussion, und zwar nicht, weil Amazon, Google und Apple dahinter sind, sondern weil die Technologie wirklich in alles hineinkommt, was sich einigermaßen dafür eignet.

**37:39 Also wirklich aus der reinen Masse an Anwendungsgebieten?**

Wenn man zb die Consumer Electronic Show im Januar anschaut, das wäre sicher noch spannend für dich. Es gibt eine wunderschöne Liste, welche Anbieter von XYZ angekündigt haben, dass sie den Google Assistant oder Amazon Alexa in ihre Produkte einbauen. Da gibt es dann so Exoten wie Badezimmerspiegel, bei dem man am Morgen sagen kann "mach mir Licht, damit ich mich gut schminken kann". Oder der Betthersteller, die Elektrik drin haben, um die Lehne zu verstellen, damit mal also nicht mehr aufs Knöpfchen drücken muss. Oder Küchenutensilien bis zu Autos. Und dadurch, dass es überall drin ist, ist es dann schon einmal da. Und wenn es schon einmal da ist, ist die Wahrscheinlichkeit gross, dass es die Leute auch nutzen. Darum glaube ich schon daran.

Dann kommen dann auch die Basissachen dazu. Reden ist einfacher als schreiben, und auch schneller. Und reden ist halt unsere natürlichste und einfachste Kommunikationsform.

**39:01 Auch wenn die Schweizer Hochdeutsch sprechen müssten?**

Neinnein, es gibt eine Schweizer Firma, die heisst Spitch.ch, die machen Spracherkennung in Schweizerdeutsch, inklusive Walliserdeutsch. Und die haben das Produkt fertig, und die sagen, dass sie heute wirklich Speech-to-Text-Erkennung machen können von allen Schweizer Dialekten. Das ist schon implementiert bei der Swisscard in Horgen, ein Callcenter. Die geben Kreditkarten heraus. Und sie machen die Spracherkennung dort, und da kann man am Anfang mit einem System reden, das erkennt dann, was du für ein Anliegen hast und routet dich dann, mit ein bisschen künstlicher Intelligenz, zur richtigen Gruppe von Call Agents. Und sie sagen, dass es funktioniert und in Betrieb ist, es läuft richtig gut. Darum: Schweizerdeutsch ist eigentlich kein Problem mehr. Es ist halt einfach für die grossen Hersteller nicht wirklich interessant, der Markt ist einfach zu klein.

**40:16 Aber für spezifisch auf die Schweiz ausgerichtete Angebote wäre das kein Problem mehr? Also es gibt ja Amazon Skills für Alexa, wenn man einen Skill hat, der sowieso nur in der Schweiz relevant ist, wäre es ja je nach dem kein Problem mehr, diesen auf Schweizerdeutsch zu bedienen?**

Nein, eigentlich überhaupt nicht.

**40:50 Gibt es sonst noch etwas, das du erwähnen möchtest?**

Was ich noch vergessen habe, was ausschlaggebend ist, dass Voice-Anwendungen in der Schweiz Erfolg haben können: die Technologien müssen in der Schweiz von Google und Amazon releast werden, das ist mal das erste. Dann gute und sinnvolle Actions und Skills von Schweizer Brands. Das können Verlage sein, die SBB, Swiss, usw. Und dann auch Retail. Das sollte so der Mix sein, dass man einen Markt lancieren kann, den es jetzt noch nicht gibt. Das sind so die Zutaten, von denen ich denke, dass es sie braucht. Und dann ganz am Schluss die Banken.

**43:49 Hast du noch Fragen an mich?**

Nein eigentlich nicht. Was mich freuen würde, wäre, wenn du mir am Schluss der Arbeit eine Kopie schicken könntest, das fände ich noch spannend.

**Vielen Dank für das Gespräch.**

## Interview Karen Kaushansky

**Livia Mosberger** [00:00:00] So let's start with the first question. Could you just describe yourself in a few words?

**Karen Kaushansky** [00:00:08] My name is Karen Kaushansky. I started designing voice experiences in 1996 so it's been twenty three years. It's a very very long time. I've seen the technology and experiences really grow over the years so I have a deep experience. I did most of that in the US working for some of the big technology companies or working for smaller clients trying to use the technology. And then two years ago I moved to Switzerland with my husband and our daughter. Over the years I've been both in-house and a consultant and so two years ago I went back to being a consultant. Many of my clients are still in the US. Like Google is one of my clients here in Switzerland. So helping companies design natural interactions like experiences with natural interfaces and conversational interactions is part of that.

**Livia Mosberger** [00:01:20] So when you first got in touch with voice technology was in 1996?

**Karen Kaushansky** [00:01:34] Yes, good question. So after university I went to work at a company in Montreal Canada called Northern Telecom. This is a big telecommunications company that had a speech recognition division and I just started working for that group. And we were building their building voice activated dialing. So it was you picked 20 contacts you had to pick them ahead of time you say these are the people that I'm going to want to call with my phone so those models can be trained offline. You have to kind of say: OK mom dad David Johnny. And it had to be trained offline so that when you pick up your phone and you had to dial a special code like star 6 then you could say "call David".

**Livia Mosberger** [00:02:21] Do you use a voice assistant or some Smart Speaker for yourself in private?

**Karen Kaushansky** [00:02:29] I really do. I should probably turn off their microphones. They're going to hear me in the manner.

**Livia Mosberger** [00:02:35] In which situations do you use them?

**Karen Kaushansky** [00:02:39] I have both Amazon Echo and I have Google Home speaker and I have a Google Home hub, the one with the screen. I work in the industry so you know, I'm not a typical user I would say but we use them in my house we use them for timers, alarms, general questions, since we are relatively new to the metric system, a lot of conversions what's, you know, 360 degrees Fahrenheit in Celsius. We use it for weather, all of our music. We just started using the Google Translate feature so it could translate one language to another. So we're all learning German so we're able to kind of how Google interpret certain things in English and say it in German and say here we say it in German it says in English. My daughter who is 7, we created some Alexa skills specific for her like blueprints, we have a little skill which is who gets to put Duffy into bed so she says the keyword "Open whose turn" and the Amazon Echo tell us whose turn it is to put her to bed. The other big thing is that in our shopping list.

**Livia Mosberger** [00:04:26] So is it just the list or do you also shop with Voice Commerce?

**Karen Kaushansky** [00:04:38] No, just the list. But I will say if I was still in the US. Like in the US I would shop on Amazon via voice. I think it's just the fact that it is possible. I would definitely do it much more.

**Livia Mosberger** [00:05:00] I've seen on your LinkedIn profile that you're at Google as a conversation designer. Can you tell me something about current projects you are on at the moment?

**Karen Kaushansky** [00:05:14] Yeah. So for about six months I have contracted with Google as one of my clients as part of the assistant team and the assistant team is very big. Have you ever used the Google Assistant?

**Livia Mosberger** [00:05:29] No not yet. Just Amazon Alexa.

**Karen Kaushansky** [00:05:32] All right. I can't really talk about what I'm working on but I would say it falls into the realm of productivity. So like things that are helping be more productive. So whether that's things like your calendar, timer, lists, reminders, things like that. I work in the productivity part.

**Livia Mosberger** [00:05:56] Yeah okay. But you are not at Google, you're like an external?

**Karen Kaushansky** [00:06:02] I'm an external contractor. I work from their office in Zurich about two days a week.

**Livia Mosberger** [00:06:08] Okay that's great.

**Livia Mosberger** [00:06:12] So the next few questions are about trust and user experience in voice and in human computer interaction like in general. And to start with, maybe it's a hard question, but is there like a most important criteria for user acceptance in general? Is the one or two more than ones? Is there none?

**Karen Kaushansky** [00:06:50] To successfully complete a task. I mean the more that you can do and deliver, the more you're going to build trust. There's like being successful at helping users get things up their assistant. They're meant to help so complete a task. And if you can't complete a task, that's okay, if you kind of recognize and get the confidence of like why that might be right. So if I say to Google, you know, "buy me an airline ticket from Basel to Paris". Hearing something come back and say "I'm sorry I didn't get that" doesn't build trust. Even if it can't fulfill, even if it says "you know what? I'm sorry. I can't book airline tickets for you. But let me give you the information about those trips". So if it applies in a smart way to help the user move forward or explains what it can and cannot do, that builds trust.

**Livia Mosberger** [00:08:09] Here in Switzerland we do not have the possibility yet to use smart speakers as people in the US can do. But do you think people here have less acceptance in voice technology than in the USA?

**Karen Kaushansky** [00:08:24] Absolutely.

**Livia Mosberger** [00:08:27] Really? And why is that?

**Karen Kaushansky** [00:08:29] I think it's not only just Switzerland but I think it's Europe in general. I mean what we're seeing in the US: U.S. has been doing this for a long time. And Europe is just kind of coming online with this. It is very new to Europe. I would say that people don't necessarily understand what it can do for them. I mean I hear it all the time. Why would I want somebody listening to me all the time in my kitchen. If it doesn't feel like they're very open to new experiences and new possibilities of what it looks like. There is a little bit of a different mentality in the US that from Europe versus privacy which is very often. I feel like it's Variant This is a very

general statement. It's not true of everything but in in the US, it feels like people are willing to give up a little bit for convenience and privacy or things like this are more opt out. Like I'm going to try it and you know if it doesn't work for me I'm going to say no I don't want it. Where in Europe privacy is very much top of mind and it's very much more opt in. I have to allow you to do this. When you have those kind have that mentality how do when you have that thinking. It's often just hard to like try new things.

**Livia Mosberger** [00:10:31] OK so do you think that is possible that the only reasons people do not trust the technology because they do not know it yet. Would you think that if Amazon or Google would release their assistants here in Switzerland people would get used to it and maybe in two years we have the same conditions as in the US?

**Karen Kaushansky** [00:10:54] Just releasing it is not the reason people are going to buy it. I think that in Switzerland it's more than that like where can people try it and say oh wow now I get it. Like I understand life. Sure I keep my paper a grocery list but I like to experience like I open the fridge and I take the last egg and I just say you know "Amazon: Add eggs to my shopping list" while I am still standing in the kitchen. And then I go to the store and it's on my list. I feel like people need to experience this. I'm just saying that when it comes to the stores that's going to change everything, this is not happening.

**Livia Mosberger** [00:11:42] We have spoken about safety and security concerns. How can these concerns be minimized?

**Karen Kaushansky** [00:11:52] I think everyone needs to understand kind of their rights and what each of these companies does and be like "I understand that". People are concerned that what they say is being used in some way. That these companies do take these utterances to make the system better but it's all anonymised. I mean I think people that are concerned need to just do their research and understand what is true about the concerns and what maybe isn't true.

**Livia Mosberger** [00:12:40] So you would say that the customer has to inform himself about it and that it is not something a company can do to higher the trust in an application?

**Karen Kaushansky** [00:12:51] I mean I think the companies do what they can. It's all written like what they do with the data, what they collect, how long it's stored, the GDPR. Like definitely, companies need to and do what they can for each user to make sure that they understand that.

**Karen Kaushansky** [00:13:28] It's definitely not a thing that I want every voice skill to tell me like "here's what we do with your data" when I first use an app. Like the whole like disclaimers because nobody reads those. Each company needs a policy that they live by and then it's like part of their culture and it's part of everything that they do and then that's how they kind of gain trust over time.

**Livia Mosberger** [00:14:04] But there is no right way to inform the customers about transparency or what is happening with their data and what data is stored? It's about the customer who has to inform himself about these applications?

**Karen Kaushansky** [00:14:22] There's also going to be newer offerings on the market where the speech recognition is all embedded. So if you're really worried about your data there will be speakers that do not send your data anywhere that it is all local. You really care but you want voice in your home, you want whatever functionality that you want. So like snips.ai is a company that's building like all of this embedded into the speaker itself so that you can have similar experiences that are always local to your home.

**Livia Mosberger** [00:15:06] And what do you think is the best thing a company can do to increase the user acceptance or trust in its voice application? Is there anything they can do or can they just wait and hope for the best?

**Karen Kaushansky** [00:15:25] I think it's how they use cases that will change people's lives. To read like a new privacy statement it's not going to change like if I use this or not. It's what can it do for me. Does it really change my daily life. Is it something that I'm going to use on a daily basis. How can it change people's lives. And then it starts to be like "now I understand why it's so convenient, now I understand why people want these in their homes." Perhaps that's what it changes. It's all about the value, it needs to bring value.

**Livia Mosberger** [00:16:05] So if an application is so good and so convenient for the customers, they do not care the same about their security anymore?

**Karen Kaushansky** [00:16:18] I think they are willing to trust a little bit more. Of course security is still important but they now see the value, they see that tradeoff like "yes it's secure but I'm willing to use it as a value that it brings."

**Livia Mosberger** [00:16:37] And what is the worst the worst thing a company can do with a voice application? Are there any like no-go's that lead to trust being lost?

**Karen Kaushansky** [00:16:50] Well I think we've heard about it and we've heard about examples in the media like making a call to the wrong person or stuff like that, you know the mistakes that we hear about in the news.

**Livia Mosberger** [00:17:16] But this this is the hardware. These examples you've brought up are the hardware components. If this was a mistake from Amazon but if a company wants to create like an Alexa skill, people have to trust in this application. Besides that they have to trust in Amazon. So what can this company with this specific skill do right or do wrong in this acceptance topic?

**Karen Kaushansky** [00:17:51] Value, you have to deliver value.

**Livia Mosberger** [00:17:59] If a big Swiss company would create a Swiss smart speaker, do you think Swiss people would automatically have more trust in this technology because they know the company behind it and they know what applications are on it?

**Karen Kaushansky** [00:18:23] That is a very hard question. I mean I'm not Swiss. I can think on one hand that that's a good hypothesis because it feels more familiar. Imagine it was in Swiss German. I definitely like speaking a language like that will definitely build trust but I have to wonder of the value again. Like are we saying that a company is going to develop something that has so many possibilities to add value to a user like it's going to support whether and a timer that is going to do all of this? Sure, maybe, but again it would have to go pretty broad and do a lot of things in order to be valuable. And I think if you ask the Swiss person like yes I would definitely say that because they brought it home. You know I don't think that either builds a relationship builds a connection to his or if it's the right company and the right language but it doesn't do enough. I don't know.

**Livia Mosberger** [00:19:36] I have read that people have to trust in Amazon and they have to trust in the voice application above it. And so if they do not trust in Amazon it's very hard for a company to create like an Alexa skill because people do not trust the hardware from the first place. So that's kind of hard for some companies who want to get into these voice applications and create their own voice app.

**Karen Kaushansky** [00:20:08] Well but I don't think users distinguish like "oh I really want to you know talk to the Starbucks application but I don't want to talk to Amazon". Yes I want to order my coffee before I leave the house so you know it's there when I get to the store but I'm not going to do it because I have to go through Amazon. I don't think people see it that way. I think it's like yes it's Starbucks on Amazon but they don't differentiate like no to Amazon yes to Starbucks. They're just going to say no.

**Livia Mosberger** [00:21:07] So am have a few questions to voice commerce. Do you think that voice commerce has a great potential and is like the big next step in e-commerce in general?

**Karen Kaushansky** [00:21:25] I would use it for certain things. Amazon started in a place that was really smart which is reordering. When you have products like the order has already been set up. A great example are diapers. I get them every four weeks. And I'm running out. I'm just going to reorder them. Something that is repeatable, I can trust, I have confidence. I'm browsing for a new pair of shoes. It's really something that I know that I want. So it's definitely direct like a direct commerce.

**Livia Mosberger** [00:22:26] What do you think are the main challenges voice commerce has to face?

**Karen Kaushansky** [00:22:37] There's commerce in buying something but there's also shopping for something. And so if you know what you want: Great. If you're browsing, if you need to see things, if you need to decide, maybe to compare across, like do I want this model or this model? It's hard to crack in a voice only world. If you now start to think about the Google Home hub or like things with a screen, I could see that being definitely easier to do and I'm sure it will catch on a little bit.

**Livia Mosberger** [00:23:17] But you would say that the main products bought about voice commerce are repeatable and products I know and I've maybe already owned and not completely new products?

**Karen Kaushansky** [00:23:30] Well or specific things like Amazon delivers to Switzerland and my book club just decided what book we're going to read next. I know exactly what book it is like there's no browsing and comparing, I know what I want. I would absolutely order it by voice instead of sitting in front of a computer. So it's like do you know specifically what what you want, I think that you can definitely order from that.

**Livia Mosberger** [00:25:13] Do you think the people have more trust issues in voice commerce than in other voice use cases like and just ask some questions, some easy questions. Do you think people are less likely to use it because it's commerce and they have to spend money on it?

**Karen Kaushansky** [00:25:40] I don't think that that's the issue. I don't know for sure but I don't think that's the issue. Like oh my credit card is involved. I mean it is anyway always like in all different ways. So I don't think it's because it's commerce that there's more trust issues. I think that it's just hard by voice only to order stuff. And if it doesn't add some efficiencies like it's quicker it's easier, if it doesn't add something, then people won't do it.

**Livia Mosberger** [00:26:33] So this means people who use a voice assistant for a shopping list they would use it for voice commerce also. And people who say I don't want to do anything with a voice assistant, they would never do the voice commerce in the first place.

**Karen Kaushansky** [00:26:53] Exactly.

**Livia Mosberger** [00:26:56] What conditions must be given for voice commerce to be accepted? I know it goes in the same topic but is there anything to improve these conditions for acceptance?

**Karen Kaushansky** [00:27:25] You know, the actual interactions have to be easy to use. If it doesn't let me kind of speak naturally and then it's people won't use it either. If it's too hard to use, if I always have to remember "I can't say it that way, I have to say it this way", that's a barrier. That's not going to let people want to use it.

**Livia Mosberger** [00:28:21] So you would say the better the applications are and the better the voice recognition is, the more trust is automatically given?

**Karen Kaushansky** [00:28:33] The experience and the technology, like how is it designed? What questions is it asking, what order, how much can I say, how much can I drive the dialogue versus the application drives the dialogue, like just how much friction is there to actually order stuff.

**Livia Mosberger** [00:29:04] So you would say voice assistance and smart speakers will succeed on a large scale here in Switzerland or in Europe as it did in the US?

**Karen Kaushansky** [00:29:26] I don't know a lot about Switzerland. We will see. The UK has had it for a long time. So UK France Spain Germany Sweden has started. So like we're seeing countries come online but some of these countries are a little bit more tech-savvy, a little bit more adventurous and want to try new technologies. I don't know if it will be the same as in the US, because US will always have a head start. But it's definitely gaining a momentum here. Maybe one thing we didn't talk about it yet is all the use cases throughout my day, can the voice assistant come with me during the day? I mean the Google Assistant is on my phone, it's in my home, it kinda knows me throughout the day, whereas Amazon is just in my home, I don't use it on my phone. So it's not about the particular use case, it's more all the use cases in one day together. So maybe I am also willing to trust more if it does more things for me during the day, and that is the value again.



# Interviewtranskript

Interviewpartner: Riccardo Lopetrone, Swisscom

Datum: 5. Juli 2019

Uhrzeit Beginn: 9 Uhr

Uhrzeit Ende: 9.40 Uhr

## Einleitung

- Vorstellung Fragestellerin
- Vorstellung des Themas der Bachelorarbeit
- Aufbau des Interviews
- Informationen zum Zeitaufwand, Aufnahme des Gesprächs, Transkription und Auswertung
- Offene Fragen soweit?

### **02:57 Bitte beschreibe dich und deine Funktion bei der Swisscom in einigen Sätzen.**

Ich bin Product Manager bei Swisscom TV & Entertainment, mit Fokus auf sprachgesteuerte Produktfeatures. Ich habe beispielsweise die jetzige Lösung entworfen, die wir jetzt auf der Fernbedienung haben, bei der man via Taste, via Push-to-talk einen Befehl in verschiedenen Sprachen, inklusive Schweizerdeutsche Dialekte, sagen kann und dann wird auf dem TV entsprechend etwas ausgelöst.

### **03:41 Hattest du die Idee zu dieser Funktion oder hast du vor allem die Entwicklung der Funktion vorwärts getrieben?**

Ich habe mehr den Scope definiert, welche Befehle vorkommen sollen, wie die Experience aussehen soll. Natürlich nicht alleine, mit allen Leuten vom TV-Team zusammen. Die Entwicklung wurde intern gemacht in Zusammenarbeit mit der EPFL in Lausanne. Sie haben die Spracherkennung für das Schweizerdeutsche gebaut.

### **04:17 Wie bist du zu dieser Funktion gekommen? Was hast du vorher gemacht, dass du schlussendlich in diese Position gekommen bist?**

Ich bin seit über 10 Jahren im Produktmanagement von TV und bin dann vor ca. 3-4 Jahren mit dem Thema Spracherkennung und Sprachsteuerung in Berührung gekommen. Ich habe das Thema sehr spannend gefunden und habe mich dann auf das spezialisiert und fokussiert.

### **04:54 Du hast es gerade angesprochen, du bist also vor 3-4 Jahren das erste Mal in Kontakt gekommen mit Voice Assistants?**

In Kontakt gekommen bin ich schon vorher, mit Siri. Und meine Aufgabe war es schon immer, zu schauen, welche neuen Features innerhalb von TV wir hereinbringen könnten, um das Produkt attraktiver zu machen und um innovativ zu sein. Darum sind wir dann auf die Sprachsteuerung gekommen.

### **05:23 Beim letzten Mal hast du uns ja gezeigt, welche Assistenten du benutzt. Nutzt du diese auch privat und wenn ja, wofür?**

Ja, ich habe eine Amazon Alexa, ich habe Apple Home Pod, den brauche ich am wenigsten, er ist aber am besten, um Musik zu hören, aber auch am teuersten. Und Google Home nutze ich auch. Aber am

meisten, also z.B. das ganze Smart Home, das ich eingerichtet habe, nutze ich über Alexa und über verschiedene Echo Devices, ich habe sicher ca. acht Echo Devices in der ganzen Wohnung. Die ganzen Lichter sind von Philipps Hue und die ganze Lichtsteuerung mache ich via Voice, ich habe gewisse Routinen eingerichtet und auch die Musik steuer ich via Voice.

**06:29 Und bist zu generell zufrieden? Funktioniert es gut oder gibt es auch Sachen, die du nicht mehr benutzt, weil es nicht funktioniert hat?**

Ich glaube, man lernt mit der Zeit, sich so zu verhalten und so zu sprechen, dass man verstanden wird. Man ist bei den Sprachassistenten noch nicht so weit, dass man völlig natürlich sprechen kann. Wenn man ein bisschen etwas von der Technik versteht, die dahinter ist, kann man sich entsprechend anpassen und die Sätze so formulieren, dass man weiss, das man gut verstanden wird. Bei mir funktioniert es zu 99% gut.

**07:15 Die Swisscom befasst sich im TV-Bereich mit Voice und bietet eine Voice-Funktion zur TV-Steuerung an. Wie ist es dazu gekommen?**

Wir haben, schon bevor wir die Sprachsteuerung auf der Fernbedienung eingeführt haben, einen Sprachwissenschaftler gehabt, der bei uns in der Entwicklung arbeitet. Zusammen mit ihm haben wir entschieden, dass das doch etwas wäre, was wir auch bei uns im TV-Produkt einführen könnten als neue Innovation. Alle zwei Jahre haben wir eine grössere Innovation, die wir bei TV einführen möchten. Und zum auch zeigen, dass wir in der Schweiz das beste TV-Produkt haben, momentan haben wir in den meisten Haushalte in der ganzen Schweiz, vor einigen Jahren haben wir UPC überholt. Und nun versuchen wir auch immer, innovativ zu bleiben.

**08:30 Und das ist also eine dieser innovativen Funktionen?**

Ja genau.

**08:35 Was waren die Herausforderungen bei der Einführung der Funktion?**

Speziell war sicher, dass wir als erste den Schweizerdeutschen Dialekt trainieren wollten. Der Aufwand ist dort relativ gross, man muss Sprachaufnahmen in der ganzen Schweiz machen, in allen verschiedenen Regionen. Die Leute rekrutieren und die Aufnahmen durchführen war eine grosse Herausforderung, in der Logistik, aber auch technisch.

**09:11 Was lief gut bei der Einführung der Funktion? Ist diese gut angekommen?**

Es ist gut angekommen. Wir haben es auch in der Werbung aufgenommen als einer der neuen Punkte, was Swisscom TV kann. Die Nutzung war eigentlich recht gut wenn man so vergleicht mit dem Ausland. Zahlen kann ich dazu nicht geben, aber wir sind über dem Durchschnitt wie sprachgesteuerte TV-Geräte auf der ganzen Welt genutzt werden.

**09:58 Du hast vorher die EPFL angesprochen. Wurde diese Funktion gemeinsam entwickelt?**

Es war ein Spin-off der EPFL, die Firma Recapp.

**10:16 Und diese Firma hat das in Zusammenarbeit mit der Swisscom gemacht?**

Ja, genau.

**10:24 Wie ist das Kundenfeedback auf diese Funktion? Wie habt ihr das erfasst?**

Wir messen regelmässig die Kundenzufriedenheit durch quantitative Umfragen, die wir bei Kunden machen. Und dann fragen wir sie einerseits allgemein über das TV-Produkt an sich, also z.B. der Wiederempfehlungswert, das ist der MPS, und dann fragen wir auch zu einzelnen Funktionen. Wie

zufrieden ist man mit der Suche, mit der Bildqualität, mit der Senderauswahl, und Sprachsteuerung ist dann auch dazugekommen.

**11:01 Dann fragt ihr auch, ob es genutzt wird?**

Die Nutzung sehen wir direkt über das Backend, im Analytics. Wir sehen, wer welche Sprache wie nutzt, welches sind die Use-Cases, die am meisten genutzt werden.

**11:21 Das Feedback allgemein war aber gut?**

Es gab beides. Es waren nicht alle Sprachen gleich gut, aufgrund der Trainingsdaten. Bei Deutsch haben wir mehr investiert, weil der grösste Anteil der Kunden Deutsch spricht. Italienisch war etwas weniger gut als Deutsch. Das versuchen wir jetzt ständig zu optimieren. Eine grosse Herausforderung war auch im Zusammenhang mit TV: wir unterstützen in der Regel nur ein Sprachmodell. Der Kunde wählt am Anfang, wenn er die Box installiert, dass er alles auf Deutsch, Französisch, Englisch oder Italienisch möchte, und dann funktioniert die Spracherkennung auch nur in dieser Sprache, die er im Menu eingestellt hat.

**12:20 Also nicht wirklich etwas für zweisprachige Haushalte?**

Genau, das wird dann schwierig. Aber man muss trotzdem gewisse Begriffe in anderen Sprachen unterstützen wenn man das Deutsche Modell auswählt. Z. B. Namen von Schauspielern oder Filmtitel, die teilweise in Englisch sind. Z.B. ich möchte einen Film schauen mit Gerard Depardieu, dann muss das Französisch natürlich auch erkannt werden innerhalb des Deutschen Sprachprogramms. Das muss man extra trainieren, dass das funktioniert.

**12:54 Wisst ihr spezifisch, welche Anfragen gemacht werden? Welche Sprachfunktionen die Leute nutzen beispielsweise?**

Ja, wir wissen, welche Use-Cases die meistgenutzten sind, ich kann da ein paar nennen. Zum Beispiel Senderwechsel ist sehr beliebt. Wir haben über 600 Sender und zum Teil wissen die Leute nicht mehr, auf welcher Nummer der Sender jetzt genau ist. Die ersten 10 weiss man mehr oder weniger, aber wenn doch mal etwas auf Eurosport kommt oder auch sonst auf einem exotischen Sender, dann müsste man im TV-Guide suchen, und via Sprache muss man nur den Namen sagen und der TV springt dann direkt dorthin. Oder auch das Öffnen von Apps zum Beispiel. Man kann sagen "ich möchte Netflix schauen" und dann öffnet sich die Netflix-App. Und sonst müsste ich im TV zuerst in den Appstore gehen, Netflix suchen, anklicken. Es gibt gewisse Use-Cases, bei denen man schneller ist mit Voice, und andere, bei denen es immer noch Sinn macht, das mit der Fernbedienung zu machen.

**14:04 Gibt es auch etwas, was euch überrascht hat, was die Leute für Use-Cases haben, die ihr gar nicht bedacht habt?**

Ja, wir haben eine Feedbackmöglichkeit den Kunden gegeben. Wenn man in die manuelle Suche reingeht haben wir dort eine URL, bei der der Kunde noch mehr Infos findet, welche Use-Cases unterstützt werden. Das heisst [www.swisscom.ch/smartremote](http://www.swisscom.ch/smartremote), das ist der Marketingname dieser Lösung, und dort drauf findet man einerseits, was man alles sagen kann, aber es gibt auch ein Eingabefeld, bei dem die Kunden eingeben können, was sie vermissen. Und dort erhalten wir tausende von Feedbacks von Kunden, was sie sich wünschen würden. Es wird rege genutzt. Gewisse Sachen lassen wir auch einfließen, anderes ist zu kompliziert mit der Version von Sprachsteuerung, die wir heute haben, was wir vielleicht erst zu einem späteren Zeitpunkt umsetzen könnten.

**15:04 Gibt es weitere eingeführte Projekte im Voice-Bereich, nicht nur auf den TV-Bereich bezogen?**

Ja, an verschiedenen Orten. Der nennenswerteste ist wohl bei Customer-Care, im Zusammenhang mit dem Kundensupport, sogenannt IVR, das ist die automatische Abnahme der Anrufe. Dort gibt es auch eine gewisse Evolution. Ganz am Anfang kam eine Meldung: Warum rufen an? Haben sie ein Problem mit Administration? Oder technisch usw. Dann muss man eine Taste drücken auf dem Telefon. Und nun wollen wir eher in die Richtung gehen, auch mit Spracherkennung zu arbeiten. Der Kunde sagt beispielsweise "Ich habe ein Problem mit meiner Rechnung" und dann wird vom System erkannt, dass es um die Rechnung geht, dann wird das Telefon automatisch an die Administration weitergeleitet. Oder man kann sogar ohne Agent am Telefon irgendwann mal die Anfrage direkt automatisch beantworten, beispielsweise bei Standardanfragen

**16:50 Wohin möchte sich die Swisscom langfristig auf dem Gebiet Voice entwickeln?**

Das ist eine strategische Frage, die ich nicht beantworten kann. Wir beobachten, was im Umfeld passiert und schauen die Lösungen der grossen Player an, wie Amazon, Google usw. Was sicher ist: Wir werden nicht einen Assistenten bauen wie sie ihn haben, der mit ihnen konkurrenzieren kann. Für das ist Swisscom zu klein, nicht international tätig, nicht die gleichen Ressourcenmöglichkeiten wie Google oder Amazon. Das wird sicher nie der Anspruch sein, dass wir auf dem Level konkurrenzieren können. Aber es gibt sicher in der Schweiz gewisse Nischen oder Bereiche, die sehr schweizerisch, sehr lokal sind und evtl. gibt es da Möglichkeiten, etwas besser zu machen als Google oder Amazon.

**17:53 Wie steht die Swisscom allgemein zu den Plattformen von Amazon und Google? War es mal ein Thema, beispielsweise einen Alexa Skill zu entwickeln?**

Es war einmal ein Thema, wir hatten auch bereits einmal einen Proof-of-Concept gemacht und geschaut, wie das aussehen könnte, um vor allem den Mehrwert herauszufinden und und herauszufinden, wie Alexa funktioniert. Aber wir haben keine konkreten Pläne momentan.

**18:29 Würde es für die Swisscom einen Unterschied machen, wenn Amazon oder Google ihre Smart Speaker in der Schweiz offiziell launchen würden? Würde das etwas beeinflussen oder ändern? Oder betrifft euch das nicht?**

Doch, das betrifft uns auf jeden Fall. Wenn natürlich sie in der Schweiz offiziell ankommen, man kann sie ja heute schon kaufen, diese Geräte, einfach über einen Graumarkt, aber es ist noch schwierig, sie zu installieren, man muss etwas Work-Hands machen, aber wenn sie dann offiziell in der Schweiz sind denke ich, dass das die ganze Sprachsteuerung und Sprachassistenten einen Boom erleben werden wie im Ausland, wie in Amerika beispielsweise, wo jeder 3. Amerikaner schon einen Sprachassistenten zuhause hat. Und in Deutschland sind es unterdessen auch schon jeder 7. Und in der Schweiz denke ich, dass es eine ähnliche Entwicklung geben wird. Vielleicht etwas weniger schnell als dort, weil hier das Thema Datenschutz sehr präsent ist in der Schweiz. Man merkt es auch an verschiedenen Diskussionen, die jetzt in den Medien und in der Politik stattfinden. Darum könnte es sich in der Schweiz vielleicht nicht gleich schnell verbreiten. Aber wenn die Geräte eine gewisse Marktpenetration erreichen, dann muss man sich sicher auch als Swisscom Gedanken darüber machen, ob mir mit irgendwelchen Skills präsent sein wollen.

**20:03 Wie betreibt die Swisscom Research auf dem Gebiet Voice?**

Wir haben Kooperationen mit Universitäten, z.B. mit der Uni Luzern, die auch in diesem Bereich Voice forscht. Wir beteiligen uns ab nächstem Jahr auch an der Voice-First-Studie, die sie gemacht haben in diesem Jahr. Dort werden wir auch unsere Inputs einfließen lassen. Dann sind wir auch in Kontakt mit grossen Firmen, die in diesem Bereich tätig sind, und auch mit Technologiefirmen, die in diesem Bereich ein grosses Know-How haben. Wir überlegen uns auch immer, was Swisscom da für eine Rolle spielen könnte.

## Vertrauen und Nutzerakzeptanz in der Schweiz

### 21:06 **Wie versucht die Swisscom, die Nutzerakzeptanz zu erhöhen? Gibt es da überhaupt eine Möglichkeit?**

Ich glaube, es ist ganz wichtig, dass man zuerst einmal die Ängste abfängt, die die Kunden haben könnten, im Zusammenhang auch mit dem Datenschutz, da das ein heikles Thema ist. Da müssen wir transparent kommunizieren, was mir mit den Daten machen, also mit den Sprachdaten oder auch sonst mit Nutzungsdaten. Dann versuchen wir, über verschiedene Kanäle, sei das Werbung oder im Kundendienst oder auf der Webseite, möglichst offen zu kommunizieren, was das Produkt kann und was es noch nicht kann. Man muss natürlich den Mehrwert möglichst aufzeigen, was Sprache mit sich bringt. Man muss einerseits die negativen Seiten abfedern und abschwächen, und zum anderen die positiven Seiten stark hervorheben, was der Kunde für einen Mehrwert dadurch erreicht wenn er Befehle via Sprache abgibt.

### 23:00 **Denkst du, dass die Personen in der Schweiz weniger aufgeschlossen gegenüber neuen Technologien sind? Oder ist einfach der Datenschutz höher gewichtet als in anderen Ländern?**

Ich glaube, es ist sicher höher gewichtet als in gewissen Ländern. Ich sage jetzt mal gegenüber Amerika haben wir da sicher strengere Vorschriften. Aber ich würde nicht zwingend sagen, dass Schweizer allgemein gegenüber Neuerungen nicht aufgeschlossen sind. Es ist sicher sehr abhängig vom Alter des Kundensegments, die Jungen sind sicher offener für neue Technologien als ältere Leute, aber das schliesst nicht aus, das auch Kunden in dem Segment das nutzen. Wir sehen auch bei uns, bei uns gibt es durchaus pensionierte Leute oder Leute im pensionierten Alter, die sehr offen sind für neue Technologien und gutes Feedback dazu geben.

### 24:13 **Denkst du, die Personen in der Schweiz sind gegenüber Voice nur misstrauisch, weil sie es noch kaum kennen? Und sobald sie zb. Alexa offiziell kaufen können, verfliegt das Misstrauen?**

Ich bin überzeugt, sobald es auch in der Schweiz ein massentaugliches Produkt wird, dass diese Ängste kleiner werden. Vor allem ist wichtig: Leute gewichten den Datenschutz nur so lange hoch, bis der Nutzen höher ist, als die Bedenken. Thema Whatsapp, Facebook usw., das sind die besten Beispiele für einen sehr invasiven Umgang mit den Daten der Leute, eigentlich die schlechtesten Beispiele überhaupt im Moment. Und trotzdem gibt es keinen Rückgang bei den Nutzern, weil halt jeder Whatsapp nutzt, weil fast jeder auf Facebook ist. Da möchte man trotzdem nicht weggehen, weil man sonst etwas verpassen könnte. Und dann nimmt man das halt auch in Kauf. Und wenn wir auch mit unseren Produkten schaffen, dass die Leute einen so grossen Nutzen davontragen, dann ist die Angst automatisch kleiner im Bereich Datenschutz und Datennutzung. Vielleicht nicht die Angst, sondern es ist einem dann einfach gleichgültig. Die Angst ist dann vielleicht trotzdem da, doch man nimmt es halt in Kauf.

### 26:14 **Denkst du, dass Transparenz der Schlüssel zum Erfolg ist? Je transparenter eine Anwendung, desto besser?**

Je transparenter die Kommunikation ist zum Produkt, ja. Also vor allem im Zusammenhang mit Daten ist sicher extrem wichtig, dass man offen kommuniziert, dass man auch sehr transparent die Leute abholt. In der Regel hat man immer irgendwelche besondere Bestimmungen oder AGB's, die man akzeptieren muss, bevor man ein Produkt nutzt. Aber die Texte sind so lang, und ich weiss aus Erfahrung, dass die wenigsten das durchlesen. Und viele Firmen neigen dann dazu, die ganzen

Informationen dort drin zu verstecken, und wenn dann jemand reklamiert sagt man einfach "steht in den AGB's drin". Und ich denke, dass man noch einen Schritt weitergehen sollte und noch offener gewisse heikle Punkte kommunizieren sollte. Also nichts vor den Kunden verstecken, sondern hervorheben und proaktiv auf die Leute zugehen mit den Informationen. So sollen alle Leute, die es nutzen, sich auch bewusst sein, was es heisst.

**27:45 Denkst du, dass die Schweizer Kunden einem Schweizer Unternehmen wie der Swisscom automatisch mehr vertrauen, als den amerikanischen Hardware-Herstellern? Weil man Swisscom halt scho kennt?**

Wir haben natürlich als Swisscom eine gewisse Vertrauensbasis aufgebaut, eine gewisse Swissness. Man hat auch gewisse Erwartungen gegenüber einer Swisscom oder Vorstellungen. Ich denke durchaus, dass wir einen gewissen Vorteil haben in dem Bereich. Gemessen oder Leute befragt haben wir noch nie, das wäre eine spannende Studie, ob das wirklich so ist. Man weiss aber auch aus Erfahrung, dass Swisscom auch noch gerne als Thema in den Medien aufgenommen wird, vor allem auch auf den negativen Seiten wenn etwas passiert, weil es halt auch ein halbstaatlicher Betrieb ist und dadurch sehr stark unter die Lupe genommen wird, was da mit den Steuergeldern teilweise gemacht wird. Wir müssen noch eine Stufe mehr aufpassen als andere.

**29:06 Ihr könnt euch also nicht alles erlauben und einfach mal ausprobieren?**

Nein, es hat zum Teil auch politische Vorgaben.

**29:11 Was ist deiner Meinung nach ausschlaggebend, dass Voice-Anwendungen generell in der Schweiz Erfolg haben können? Was ist das Wichtigste, was passieren muss?**

Das eine ist sicher das mit dem Datenschutz, dass die Kunden das Gefühl haben, dass ihre Daten mit Vertrauen behandelt werden und nicht irgendwie zweckmissbraucht, was die Kunden nicht wollen. Und das andere ist die Spracherkennung möglichst perfekt funktionieren muss, da sonst das Projekt einfach nicht mehr genutzt wird. Das sehen wir auch bei vielen Umfragen, wo wir die Kunden fragen: "Was ist ihnen das wichtigste bei einem Sprachassistent?". Dann ist immer der erste Punkt die Spracherkennung, dass der Assistent mich versteht.

**30:12 Denkst du, dass das Misstrauen bei Voice Commerce höher ist als bei anderen Voice-Anwendungen?**

Ich glaube, Voice hat in vielen Bereichen Vorteile gegenüber anderen User Interfaces, weil es schnell geht von der Spracherkennung her, aber es gibt auch gewisse Nachteile je nach Use Case, weil ich halt kein visuelles Feedback habe. Es gibt Use-Cases, bei denen Voice alleine ein gewisses Vertrauen da ist, aber das was du ansprichst, da erwarte ich möglichst visuell ein Feedback, was ich da genau mache. Sobald es um Geld geht, denke ich, dass da die multimodale Experience eine grosse Rolle spielt, wo ich Voice mit einem Bildschirm verknüpfe. Wie zum Beispiel der Echo Show von Amazon oder vielleicht im Zusammenhang mit sonst einem Bildschirm, der dann auch ein visuelles Feedback gibt, was jetzt gerade passiert.

**31:58 Aber grundsätzlich würdest du nicht sagen, dass die Leute bei Voice Commerce vorsichtiger eingestellt sind.**

Doch ich denke schon. Also ich habe auch von der Uni Luzern gerade diese Woche eine Umfrage gezeigt bekommen, die sie gemacht haben, und dort sieht man, dass bei gewissen Use Cases der Kunde weniger Vertrauen hat via Voice. Und Voice Commerce ist einer davon.

**32:40 Denkst du, dass sich Voice Assistenten und Smart Speaker in der Schweiz im grossen Stil durchsetzen werden? Und wenn ja, in welchem Zeitraum?**

Längerfristig bin ich überzeugt, ja. Aber nicht in einem sehr kurzfristigen Zeitraum, ich denke dass so in ca. 5 Jahren Werte erreicht werden, die wir heute beispielsweise in Deutschland sehen.

**33:22 Und dann auch abhängig vom Markteintritt von Google und Amazon?**

Genau, abhängig davon wenn sie überhaupt in die Schweiz kommen, die Schweiz ist ja nicht so ein attraktiver Markt, durch die Mehrsprachigkeit, es ist sehr klein, wir haben verschiedene Dialekte.

## Interviewauswertung

Interviewauswertung			
Kategorie	U. Bucher	K. Kaushansky	R. Lopetrone
Private Nutzung	Spannend, dass VA in einer ganzen Reihe von Produkten zu finden sind, z.B. Autos	Timers, Alarms, allgemeine Fragen, Übersetzungen, spezifische eigene Skills in der Familie	Smart Home via Alexa, Musik abspielen, Einkaufsliste erstellen
	Lokale Anwendungen gehen nicht. Nutze VA für News, Google Maps, telefonieren, Spotify	Einkaufsliste via Voice, wäre ich noch in den USA würde ich auch direkt per Voice einkaufen	
Projekte	Workshop mit Tamedia zu Voice im Medienbereich	beim Google Assistant Team für Anwendungen zuständig, die die Produktivität betreffen	Sprachsteuerung des TV mit der Fernbedienung
	Studie Farmer und Uni Luzern, am Rande beteiligt		
Nutzung VA in der Schweiz	viel mehr Leute als angenommen nutzen in der Schweiz bereits VA		Es gibt in der Schweiz Nischen, die sehr lokal und schweizerisch sind. Das eröffnet Möglichkeiten.
	Altersunterschied zwischen Nutzern und Nichtnutzern gibt es nicht		In der Schweiz wird es die gleiche Entwicklung geben wie in den USA oder Deutschland
	Nutzung entspricht der Nutzung in Deutschland		VA werden sich längerfristig etablieren, in 5 Jahren sind wir vielleicht auf dem Stand wie Deutschland
Nutzerakzeptanz in CH	sobald es um Geld geht braucht man maximale Sicherheit		Entwicklung ist langsamer als im Ausland, weil in der Schweiz Datenschutz sehr präsent ist.
	man tastet sich schrittweise vor und baut nach und nach Vertrauen auf		man muss die Ängste der Kunden abfangen
Voice Commerce in CH	wer technologieaffin ist und Online-Shopping nutzt, wird auch schnell Vertrauen in Voice Commerce fassen		die Schweiz ist nicht so ein attraktiver Markt für Google und Amazon
	Voice Commerce in der Schweiz ist ein Huhn-Ei-Problem		
Begünstigung Akzeptanz	Grosse, bekannte Player, denen die Kunden bereits vertrauen, müssen vorausgehen	Anwendung muss erfolgreich ausgeführt werden können	
Information und Kommunikation	Medienhäuser müssen Informationen bereitstellen	Die Nutzer brauchen klare Informationen, was mit den Daten passiert, und sie müssen das auch verstehen können	transparent kommunizieren, was mit den Daten passiert
	Verlässlichkeit und Einfachheit müssen gegeben sein	Den Nutzern muss klar sein, was die Anwendung kann und was nicht	man muss den Kunden klar machen, was das Produkt kann und was nicht
		Nutzer müssen ausprobieren können, nur durch den Release gibt es noch keine Akzeptanz	sobald VA in der Schweiz massentauglich sind, werden die Ängste kleiner
		In Zukunft könnte es VA geben, die alles lokal verarbeiten und keine Daten nach aussen senden	
Nutzen und Wert		Ist der Nutzen und der Wert genug gross, werden die Datenschutzbedenken in den Hintergrund gestellt	Es gibt Use Cases, bei denen man mit Voice schneller ist, und solche, bei denen Voice keinen Sinn macht
		Anwendung muss einen möglichst grossen Nutzen stiften und Wert generieren	Leute gewichten Datenschutz so lange hoch, bis der Nutzen höher ist als die Bedenken.



			Wenn der Nutzen genug gross ist, sinkt die Angst um den Datenschutz zwar eher nicht, aber es wird eher gleichgültig und man nimmt es in Kauf.
Schwierigkeiten Akzeptanz	Unsicherheit bezüglich der Technologie erschwert Akzeptanz	wie sollen die Informationen bereitgestellt werden? Niemand möchte vom VA die AGB's vergelesen bekommen	Firmen neigen dazu, wichtige Informationen in den AGB's zu "verstecken"
	keine grossen und bekannten Anbieter in der Schweiz		Spracherkennung muss noch besser werden und die Dialekte besser verstehen
	kein offizieller Launch von Google oder Amazon, sieht auch nicht danach aus im Moment		Kunden dürfen nicht das Gefühl haben, dass ihre Daten missbraucht werden
	negative Schlagzeilen zur Stimmerkennung schaden der ganzen Technologie	Jeder kennt negative Beispiele aus den Nachrichten	
	Man ist gegen etwas, weil man es noch gar nicht kennt.		
	Man gewinnt in der Masse Vertrauen		
Vergleich Akzeptanz zum Ausland	In Deutschland scheint es gemäss der Farmer-Studie ähnlich zu sein wie in der Schweiz	Datenschutz ist in Europa höher gewichtet. In den USA ist die Mentalität eher Opt-out, falls es nicht passt. Europäer sind eher Opt-in und müssen sich zuerst ganz sicher sein	Datenschutz ist höher gewichtet als in anderen Ländern, wir haben strengere Vorschriften als z.B. die USA
	Im Vergleich zu USA haben wir in der Schweiz grosse Unterschiede bezüglich Sicherheit	Grosser Unterschied zu den USA. In Europa ist die Technologie noch ganz neu, die Leute haben den Nutzen noch gar nicht erkannt	Schweizer sind gegenüber Neuerungen aufgeschlossener, unabhängig vom Alter
Potenzial von Voice Commerce	Für 0815-Artikel mit tiefem Preis kann es funktionieren	Voice Commerce funktioniert für gewisse Produkte, vor allem für Nachbestellungen	
		Produkte, die ich genau kenne, sehr spezifisch, z.B. ein Buch	
		mit einem Screen kann das Einkaufen, aber auch das Vergleichen definitiv einfacher werden	
Herausforderungen Voice Commerce	mit mehr Risiko und höheren Beträgen nimmt die Akzeptanz rapide ab	vergleichen und dann entscheiden ist schwierig	Leute sind kritischer bei Commerce, es ist ein Use Case, dem noch wenig vertraut wird.
		Einfache Benutzung und möglichst natürliche Interaktion	
Kritischer Punkt beim Kaufprozess	Schluss des Kaufprozesses muss möglichst einfach sein, möglichst schnelle und einfache Bestätigung	wenn der Nutzen gross genug ist, gibt es keinen kritischen Punkt mehr. Wer bereits die Einkaufsliste via Voice macht, wird auch damit einkaufen	
	Es braucht einen Medienbruch, man muss sehen, dass es funktioniert hat		es braucht ein visuelles Feedback, vor allem sobald es um Geld geht
Masse an Anwendungen	Wenn Voice überall enthalten ist, ist die Wahrscheinlichkeit auch gross, dass es genutzt wird	je mehr man mit dem Assistent machen kann, desto mehr Vertrauen gibt es	
Voice auf Schweizerdeutsch	Schweizerdeutsch kann gut erkannt werden, auch die Dialekte. Wird z.B. in Callcentern bereits eingesetzt		
Plattformübergreifend		VA soll durch den ganzen Tag hindurch begleiten, nicht nur zuhause, sondern auch unterwegs, und ein einheitliches Erlebnis bieten. So steigt der Wert an.	



---

## Bisher erschienene Schriften

Ergebnisse von Forschungsprojekten erscheinen jeweils in Form von Arbeitsberichten in Reihen.  
Sonstige Publikationen erscheinen in Form von alleinstehenden Schriften.

Derzeit gibt es in den Churer Schriften zur Informationswissenschaft folgende Reihen:  
Reihe Berufsmarktforschung

### Weitere Publikationen

Churer Schriften zur Informationswissenschaft – Schrift 107  
Herausgegeben von Wolfgang Semar  
Debora Greter  
Wissensmanagement in der Lebensmittelindustrie  
Chur, 2020  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 108  
Herausgegeben von Wolfgang Semar  
Reto Siegenthaler  
Entwicklung eines kollaborativen Wissensmanagement im Krisenmanagementprozess  
Chur, 2020  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 109  
Herausgegeben von Wolfgang Semar  
Julia Knuchel  
Semantische Technologien – Nutzung, Bedürfnisse und Probleme in Forschungsprojekten  
Ein Beitrag zur Neuausrichtung der Forschungsunterstützung als Dienstleistung wissenschaftlicher  
Bibliotheken  
Chur, 2020  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 110  
Herausgegeben von Wolfgang Semar  
Christina Fischer  
Augmented Reality, der Designprozess und Prototyping: State of the Art  
Chur, 2020  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 111  
Herausgegeben von Wolfgang Semar  
Pascale Marder  
Gemeindearchive  
Zwischen Selbstverwaltung und Mandatsverhältnis  
Chur, 2020  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 112  
Herausgegeben von Wolfgang Semar  
Jin Chei  
Digitalisierungslandschaft in Schweizer Archiven  
Explorative Studie bezüglich Situation, Herausforderungen und gewünschter Unterstützung  
Chur, 2020  
ISSN 1660-945X

Churer Schriften zur Informationswissenschaft – Schrift 113  
Herausgegeben von Wolfgang Semar  
Raphael Jung  
Agile Marketing für Digital Marketing Agenturen  
Analyse von Projektmanagement Methoden und Definition von konkreten Handlungsempfehlungen  
Chur, 2020  
ISSN 1660-945X

---

## Über die Informationswissenschaft der Fachhochschule Graubünden

Die Informationswissenschaft ist in der Schweiz noch ein relativ junger Lehr- und Forschungsbereich. International weist diese Disziplin aber vor allem im anglo-amerikanischen Bereich eine jahrzehntelange Tradition auf. Die klassischen Bezeichnungen dort sind Information Science, Library Science oder Information Studies. Die Grundfragestellung der Informationswissenschaft liegt in der Betrachtung der Rolle und des Umgangs mit Information in allen ihren Ausprägungen und Medien sowohl in Wirtschaft und Gesellschaft. Die Informationswissenschaft wird in Chur integriert betrachtet.

Diese Sicht umfasst nicht nur die Teildisziplinen Bibliothekswissenschaft, Archivwissenschaft und Dokumentationswissenschaft. Auch neue Entwicklungen im Bereich Medienwirtschaft, Informations- und Wissensmanagement und Big Data werden gezielt aufgegriffen und im Lehr- und Forschungsprogramm berücksichtigt.

Der Studiengang Informationswissenschaft wird seit 1998 als Vollzeitstudiengang in Chur angeboten und seit 2002 als Teilzeit-Studiengang in Zürich. Seit 2010 rundet der Master of Science in Business Administration das Lehrangebot ab.

Der Arbeitsbereich Informationswissenschaft vereinigt Cluster von Forschungs-, Entwicklungs- und Dienstleistungspotenzialen in unterschiedlichen Kompetenzzentren:

- Information Management & Competitive Intelligence
- Collaborative Knowledge Management
- Information and Data Management
- Records Management
- Library Consulting
- Information Laboratory
- Digital Education

Diese Kompetenzzentren werden im Swiss Institute for Information Research zusammengefasst.



---

## Impressum

### Impressum

FHGR - Fachhochschule  
Graubünden  
Information Science  
Pulvermühlestrasse 57  
CH-7000 Chur

[www.informationsscience.ch](http://www.informationsscience.ch)

[www.fhgr.ch](http://www.fhgr.ch)

**ISSN 1660-945X**

### Institutsleitung

Prof. Dr. Ingo Barkow

Telefon: +41 81 286 24 61

Email: [ingo.barkow@fhgr.ch](mailto:ingo.barkow@fhgr.ch)

### Sekretariat

Telefon: +41 81 286 24 24

Fax: +41 81 286 24 00

Email: [clarita.decurtins@fhgr.ch](mailto:clarita.decurtins@fhgr.ch)