



# Voice Commerce

Ein Leitfaden für sprachgesteuerten digitalen Handel

# Voice Commerce

## Ein Leitfaden für sprachgesteuerten digitalen Handel

<b>Vorwort</b>	<b>2</b>
<b>Warum eigentlich Voice Commerce?</b>	<b>3</b>
Voice: Eine natürliche Art der Kommunikation. Einfach menschlich.	3
Voice: Die Technologie ist ausgereift. Die Zeit ist reif für Voice.	3
Voice: Next Level – Voice UX Design und SSML	4
Voice: Der holistische Ansatz für die Betrachtung macht den Unterschied.	5
Voice: Der Voice-Check.	5
<b>Wo in der Customer Journey ist Platz für Voice Commerce?</b>	<b>7</b>
Was ist eigentlich Voice Commerce?	7
Die unendliche Customer Journey verstehen	7
Voice Commerce in Phase 1: Awareness	8
Voice Commerce in Phase 2: Consideration	9
Voice Commerce in Phase 3: Conversion	9
Voice Commerce in Phase 4: Retention	10
Voice Commerce in Phase 5: Advocacy	10
Wie funktioniert der Infinity Loop der Customer Journey mit Voice?	10
<b>Voice Commerce starten, aber wie?</b>	<b>11</b>
Ideen für eine erfolgreiche Zukunft mit Voice Commerce.	11
Voice Commerce starten mit der Angebotsauswahl	11
Voice Commerce funktioniert wie gesprochene Sprache	12
Voice Commerce hat Vorbilder zum Lernen	13
<b>Voice Commerce: Inspirationen durch Beispiele</b>	<b>13</b>
Case Study 1: Mit Sprache vom Briefkasten zur Lieferung	14
Case Study 2: Der Inhouse-Versicherungsvertreter	15
Case Study 3: Interaktive Produktbegleitung von Fast-Moving Consumer Goods (FMCG)	16
Case Study 4: Die Mittagstheke der Metzgerei Kaufmann	17
<b>Voice Commerce. What's Next?</b>	<b>20</b>
<b>Autoren und Autorinnen</b>	<b>21</b>
<b>Anhang</b>	<b>22</b>
Glossar	22
Quellen und weiterführende Literatur	23
<b>Über uns</b>	<b>24</b>
Impressum	25

## Vorwort

Dialoge sind die logische nächste Entwicklungsstufe im Digital Commerce. Mit Voice Commerce lässt sich der Handel weiter entwickeln. Es kann der nächste Game Changer werden.

Voice Commerce sind – für die Autorinnen und Autoren dieses Leitfadens und mich – geschäftsunterstützende Dialoge zwischen Kunden und Unternehmen über digitale Schnittstellen, zum Beispiel mittels eines Smart Speakers, über Wearables oder Smartphones. Voice Commerce erweitert die asynchrone Kommunikation mit Inhalten auf Webseiten, in Apps oder in Chats um eine digitale, synchrone Sprachkommunikation.

Unternehmen können so Kunden und Interessenten auf einfache, schnelle und sehr alltagstaugliche Art mit Informationen versorgen. 1-to-1-Marketing wird im Kundenerlebnis real. In Voice-Commerce-Dialogen wird individuell auf die Fragen und Anliegen der Kunden eingegangen. Es werden die Informationen ausgetauscht, die von realem, situativem Interesse sind. Einfach. Interaktiv. Mit Sprache.

Gerade im aktuellen Lebensumfeld, das von massiven Veränderungen des gewohnten Lebens geprägt ist, sind neue Impulse für hybride Lebensweisen wichtig. Die Digitalisierung mit Voice Commerce zeigt über alle wesentlichen Bereiche der Customer Journey klare Nutzenpotenziale für Handel und Kunden. Und dies nicht nur theoretisch, sondern in inspirierenden Fallstudien.

Dieser Leitfaden bietet mit seinen eingängigen Beispielen konkrete Ansätze zur Gestaltung hybrider Käuferlebnisse: Einfach mal das Angebot für den Mittagstisch beim lokalen Metzger checken oder Schuhe zum Anprobieren und Abholen reservieren. Die Case Studies beleuchten Chancen und die Möglichkeiten zur Umsetzung von Voice Commerce. Fokussiert werden dabei Kundenerlebnisse und Geschäftspotenziale sowie die Komplexität der technischen Realisierbarkeit.

Insgesamt soll der vorliegende Leitfaden dazu beitragen, verschiedene Ansätze des Handels weiterzudenken und Verbindungen zwischen Lösungen vor Ort und dem digitalen Handel zu schaffen. Wird Voice Commerce so zum „Missing Link“ des Connected Commerce?

Meine Empfehlung für Sie: Einfach jetzt den ersten Schritt machen, den vorliegenden BVDW-Leitfaden lesen, sich inspirieren lassen ... und dann einfach mal machen.

**Oliver Bohl**

Vorsitzender der Fokusgruppe Digital Commerce

## Warum eigentlich Voice Commerce?

**Voice: Eine natürliche Art der Kommunikation. Einfach menschlich.**

***Welche Kommunikationsform ist menschlicher als Sprache? Gesprochene Sprache.  
Was sollte die natürliche Mensch-Maschinen-Schnittstelle sein? Sprache. Voice.  
Conversational Commerce.***

Wie zwischen Menschen führt Austausch über gesprochene Sprache auch zwischen Mensch und Maschine zur Verbesserung von Information, Beratung, Verkauf, Service und Empfehlungen.

In der alltäglichen Kommunikation zwischen Menschen nutzen wir primär die gesprochene Sprache. Und das, seit es Menschen gibt.

Anders war das bisher bei der Techniknutzung. Die Entwicklung von Maschinen hat technische Steuerungselemente in Arbeit und Alltag integriert. Darum haben technische Schnittstellen ihren Raum im Leben erobert. Einige haben wir wirklich lieb gewonnen, wie Schalter, Drehknöpfe oder Computer-Mäuse.

Historisch sind die Bedienelemente für Maschinen – die Mensch-Maschine-Schnittstellen – von Hebeln und (Lenk-)Rädern geprägt. Mit fortschreitender Entwicklung wurden diese Elemente durch Schalter, Tasten, Regler und andere Bedienelemente ergänzt. Aktuell sind berührungsempfindliche Oberflächen (Touchscreens) weit verbreitet. Sie haben viele Tasten und Schalter verdrängt bzw. diese virtualisiert.

Zusätzlich haben sich in den letzten Jahren sprachgesteuerte Schnittstellen etabliert. In verschiedenen Nutzungskontexten findet damit die natürlichste aller Kommunikationsformen des Menschen immer stärkeren Einzug im sprachbasierten Austausch mit Maschinen.

Menschen nutzen alltäglich verschiedene Dinge, z. B. Haushaltsgeräte, Computer, mobile Endgeräte und smarte Armbanduhren (Wearables).

Die Herausforderung der Interaktion liegt dabei häufig im Nutzungskontext. Wenn z. B. die Hände zur Steuerung nicht eingesetzt werden können oder es einfach bequemer ist, schnell etwas zu sagen, statt nach einem Gerät zu greifen, schlägt die Stunde der Sprache. Neben der Einfachheit der Sprache als Schnittstelle zu Maschinen spielt also auch die Situation der Nutzung eine wichtige Rolle.

Eine echte Herausforderung ist dabei das Design des natürlichen Sprachdialogs. („Wie erkenne ich die Absicht – den Intent – der Nutzer?“ ... „Wie kann ich das passende Antworten erlernen?“).

**Voice: Die Technologie ist ausgereift. Die Zeit ist reif für Voice.**

Voice works. Es funktioniert. Und macht Spaß. Damit liegt Voice voll im Trend. Die Anzahl der Sprachanwendungen sowie der sprachgesteuerten Geräte ist in den letzten Jahren nahezu exponentiell gestiegen.

Bereits Ende 2019 gab es weltweit rund zwei Milliarden Geräte mit Sprachsteuerung. Schon Ende 2020 hatte sich die Anzahl mehr als verdoppelt auf viereinhalb Milliarden. Dank Alexa, Siri und Co. sind sprachgesteuerte Anwendungen omnipräsent und für viele Menschen nicht mehr aus ihrem Alltag wegzudenken.

Die Gründe für das Wachstum liegen auf der Hand: Die Technologien zur Verarbeitung natürlicher Sprache haben sich enorm verbessert. Hier zählt sich der jahrelange Entwicklungseinsatz der führenden Unternehmen wie Amazon, Apple oder Google aus. Was gut funktioniert, fühlt sich bei der Nutzung gut an. Die Nutzungserlebnisse erfüllen die Erwartungen an die Anwendung.

Menschen entwickeln neue Interaktionsgewohnheiten und setzen – zu Hause oder unterwegs – immer häufiger Sprache zur Interaktion mit Anwendungen ein. Das funktioniert oft schon ganz intuitiv. Der Nutzer braucht schließlich keine Anleitung für das Sprechen. Allerdings erwarten einige Anwendungen spezielle „Kommandos“ oder präzise Anweisungen, etwa eine vollständige Adresse mit Stadt, Straße und Hausnummer, und nicht: „Ich möchte zu ... fahren“.

Einfach miteinander reden. – Leider gar nicht so einfach. Menschliche Sprachmuster weisen nahezu unendliche Variationen in Syntax, Vokabular, Tonalität und Semantik auf. Hinzu kommen die Untertöne – von Witz oder Ironie gar nicht erst zu reden. Für die erfolgreiche Interpretation von Sprache gilt es deshalb zusätzlich übergeordnete Faktoren wie Kontext oder Ironie und Sarkasmus zu berücksichtigen, um nur einige zu nennen.

Die Bedeutung kann leicht das Gegenteil der Summe aller Wortbedeutungen aussagen, wie: „Was für ein faszinierendes Kleid! Ich glaube nicht, dass ich so etwas schon einmal gesehen habe.“ Wie klingt das? Schmeichelhaft – oder eher herablassend?

Neben leistungsstarken Machine-Learning-Modellen für die Spracherkennung tragen Technologien wie Sentimentanalyse<sup>1</sup>, semantische Netzwerke<sup>2</sup> oder Ontologien<sup>3</sup> dazu bei, dass Anwendungen die menschliche Sprache besser verstehen.

### Voice: Next Level – Voice UX Design und SSML

Einfach ist erst mal gar nicht so einfach, sondern schwierig. Wir müssen bei der Umsetzung von Sprachschnittstellen in Marketing und Unternehmen loslassen lernen. Loslassen von den gelernten Strukturen. Von den alten Schnittstellen.

Die Entwicklung von Mensch-Maschine-Schnittstellen hat Notwendigkeiten unserer Maschinen in Abläufe übersetzt, die Eingabe der maschinellen Funktion unterstellt. Unsere Computertastaturen haben ihren Ursprung noch in der Ära der Schreibmaschinen. Der visuelle Content im Web hat Website-Strukturen nach dem Vorbild von Kompendien erzeugt. Ebenso ist eine Maus, die historisch etwas später hinzukam, um etwas auf Bildschirmen zu nutzen, ein ganz anderes Nutzungsprinzip als Voice.

Ein Dialog ist keine erfragbare Bibliothek. Diese Analogien gilt es bei Sprache einfach zu vergessen. Verschachtelte Navigationsmenüs haben im reinen Sprachdialog keinen Platz. Kompetentes Voice UX Design lässt sich von Intuition, Nutzungskontext und einfachstem Sprachdialog leiten.

---

<sup>1</sup> **Sentiment Analysis** ist ein Untergebiet des Text Mining und bezeichnet die automatische Auswertung von Texten mit dem Ziel, eine geäußerte Haltung als positiv oder negativ zu erkennen.

<sup>2</sup> Ein **semantisches Netzwerk** ist ein formales Modell von Begriffen und ihren Beziehungen (Relationen). Es wird in der Informatik im Bereich der künstlichen Intelligenz zur Wissensrepräsentation genutzt. Gelegentlich spricht man auch von einem Wissensnetz.

<sup>3</sup> Die **Ontologie** ist eine Disziplin der Philosophie, die sich mit der Einteilung des Seienden und den Grundstrukturen der Wirklichkeit befasst, z. B. mit Begriffen wie Existenz, Sein, Werden und Realität.

Das ist aber noch nicht alles: Designsprachen wie SSML (Synthetic Speech Markup Language) ermöglichen es Designern, Sprachausgaben zu generieren, die die feinsten Nuancen der menschlichen Sprache effektiv erzeugen können.

Selbst Emotionen, Zwischenrufe, Akzente oder Änderungen in Tonhöhe, Lautstärke und Geschwindigkeit lassen sich integrieren. Die Sprachausgabe der neuesten Sprachassistenten-Generation klingt natürlich wie noch nie und ist von einer „realen“ menschlichen Stimme kaum noch zu unterscheiden.

### **Voice: Der holistische Ansatz für die Betrachtung macht den Unterschied.**

Voice Commerce ist nicht als punktueller Ansatz erfolgreich, sondern der Einstieg in Voice erfordert einen ganzheitlichen Ansatz, welcher Kundenorientierung mit Marketing 4.0 vereint – mit einem Fokus auf wertschöpfendem Einsatz von Voice. Und hier insbesondere mit dem Fokus auf gute Voice User Experience als Grundlage.

Voice wird somit zum nächsten Schritt der Digitalisierungs-Journey von Unternehmen.

Positiv gesehen: Voice kann die Markenbotschaft und -konsistenz stärken und wirtschaftliche Vorteile durch Effizienz sowie Kundenzufriedenheit mit sich bringen.

Lassen wir uns doch einfach auf das Gedankenexperiment ein. Voice und wir, aber holistisch, bitte.

Der Startpunkt für den holistischen Ansatz ist eine detaillierte Top-down-Betrachtung und Analyse der Beziehungen zwischen Kunden und Unternehmen sowie der gesamten digitalen Landschaft, zum Beispiel durch eine Agentur oder eine Beratungsgesellschaft.

### **Voice: Der Voice-Check.**

Das Ziel für den Voice-Check ist die Aufnahme der Konversations-Interaktionen (der Kunden-Unternehmens-Beziehungen auf der Customer Journey) und deren Integration in eine konsistente strategische Analyse.

Um einen ersten Eindruck zu vermitteln, haben wir einen Quick-Voice-Check als kleinen Fragebogen für Sie entworfen:

#### **Voice-Check:**

##### **Phase 1 / „Touchpoint Recherche“**

Identifizieren der konversationsrelevanten Touchpoints und Interaktionen mittels Recherche, Beobachtung und Datenanalyse:

- Standard-Touchpoints erfassen
- (gegebenenfalls) ergänzen um
  - individuelle Touchpoints – jeweils pro Kundengruppe
    - Interessentinnen und Interessenten
    - Kundinnen und Kunden
    - Mitarbeitende
    - weitere
  - die eingesetzten digitalen Anwendungen
  - Prozessbeschreibungen

### Voice-Check:

#### Phase 2 / Entwicklung „Conversational Landscape Map“

Bestimmen mehrwertbringender Anwendungsfälle für den Einsatz von Sprache in den Touchpoints und Prozessen aus Phase 1.

Beschreibung dieser Fälle mit einer ersten Bewertung von Aufwand und GuV-Einfluss sowie Kundennutzen. Darstellung aller Komponenten in einer „Conversational Landscape Map“.

- Customer-Journey-basierte Analyse von Interessentenkontakten entlang den Stationen Verkauf, After-Sales-Service und Empfehlungsmanagement
- Markierung aller Fälle mit „Hands-free“-Vorteilen
- Markierung aller Fälle mit „Privacy“-Notwendigkeiten
- Markierung der einfach zu integrierenden Funktionen an bestehenden Schnittstellen, z. B. an CRM-Systemen
- Markierung von Optionen für neue Digitalisierungen / Automatisierungen
- weitere

### Voice-Check:

#### Phase 3: Ableitung des „Holistic Voice Branding“

Ableitung einer Voice-Markentonalität mit eigener Persona sowie allen notwendigen Komponenten für die Umsetzung in eine Corporate Voice. Priorisierung der Anwendungsfälle aus der Analyse unter Verwendung der Einordnung auf der Conversational Landscape Map. Gestaltung der notwendigen Komponenten.

- Conversational Design
- Voice UX Studies („Wizard of Oz“-Tests), High Fidelity Prototyping
- Umsetzung Sonic Branding
- IT-Architecture Check (APIs, ...)

### Voice-Check:

#### Phase 4: Technische Umsetzung, Launch-Unterstützung

Konzipieren, Programmieren, Testen und Veröffentlichen aller Komponenten für die Voice-Interface-Anwendungen:

- Umsetzung der Voice-Anwendung, inklusive der Programmierung aller notwendigen Intents
- Umsetzung der IT-Integration, z. B. Verbindung der Voice Interfaces mit APIs
- Test und Abnahme des Gesamtsystems
- Soft-Launch mit Friends-&-Family-Programm
- Optimierungsphase mit den Erkenntnissen des Soft Launch
- Launch
- UX-Testing mit realen Nutzerinnen und Nutzern und realer IT-Umsetzung
- weitere Optimierung

## Wo in der Customer Journey ist Platz für Voice Commerce?

### Was ist eigentlich Voice Commerce?

Wir verstehen Voice Commerce nicht nur als Voice-Vertrieb, sondern in einem weiteren Sinne, der auch alle die Transaktion vorbereitenden und unterstützenden Aktivitäten beinhaltet. Orientiert an einer „unendlichen“ Customer Journey verstehen wir den gesamten Kontext von Information bis zur Empfehlung – und nicht die ausschließliche Fokussierung auf den Kauf.

Diese Customer Journey ist es, auf der Unternehmen ihre Kundinnen und Kunden begleiten und die eine Einordnung möglicher Voice-Touchpoints in eine strukturierte und erfolgversprechende Systematik ermöglicht.

### Die unendliche Customer Journey verstehen

Üblicherweise wird die Customer Journey in mehrere Phasen eingeteilt, die den jeweiligen Bedarf der Kunden an verschiedenen Punkten ihrer Kaufentscheidung widerspiegeln. Dabei orientiert sich die Customer Journey an den wissenschaftlichen Kaufprozess-Modellen, insbesondere am 4-Phasen-Modell AIDA.

Aktuellere Versionen ergänzen weitere Bereiche der Journey: „Retention“ und „Advocacy“. Weitere Veränderungen in der Struktur der Reise sind die Zusammenfassung der Phasen „Interest“ und „Desire“ zu „Consideration“ sowie die Aktualisierung von „Action“ zu „Conversion“ (vgl. Abbildung „Marketing in einer hybriden Welt“).



Oft wird die Customer Journey als linearer Ablauf und damit als endliches (nach dem Kauf bzw. einer Empfehlung endendes) Phasen-Modell betrachtet. Bezieht man aber die Frage der Kundenbindung mit ein, wird aus dem einmaligen Weg eine lebenslange Reise (Stichwort: Customer Lifetime Value).



Es ergibt sich eine unendliche Schleife aus Awareness (Aufmerksamkeit), Consideration (Information), Conversion (Entscheidung) sowie Retention (Nutzung/Bindung) und Advocacy (Empfehlung). Voice Marketing und Kundenbindung mit Sprachassistenten lassen sich gut in diese unendliche Kundenreise eingliedern. Voice Apps haben – wie viele andere Marketinginstrumente – nicht zwangsläufig singuläre Rollen im Rahmen der Customer Journey. Sie begleiten die Reise mit verschiedenen Funktionen über verschiedene Phasen.

Die „liegende 8“ (oder das mathematische Unendlich-Zeichen) sind aus diesem Grund auch die ideale Darstellungsmöglichkeit, um einen Rahmen für Voice Marketing aufzuspannen, wie die folgende Abbildung zeigt.



Die nach einer Betrachtung zentraler Dos und Don'ts für Sprach-Dialoge vorgestellten Case Studies werden deshalb immer in diesen Customer Life Infinity Loop eingeordnet, wobei die Bedeutung der Voice Apps innerhalb dieser Phasen aufgezeigt wird.

### Voice Commerce in Phase 1: Awareness

Der erste Schritt für den Erfolg: bemerkt werden. Wir brauchen Aufmerksamkeit. Die Aufmerksamkeit für ein Produkt oder eine Marke kann prinzipiell auf zwei Wegen erreicht werden:

Zum einen durch **Content Marketing**, bei dem nicht direkt werbliche, aber für die Kunden wertvolle Inhalte angeboten werden. Das können beispielsweise per Voice App abfragbare Informationsangebote oder interaktive Hilfestellungen sein, die die Kunden wegen ihrer für sie hilfreichen oder unterhaltenden Funktion nutzen. So werden sie auf Produkt und Marke aufmerksam.

Zum anderen bietet sich als Alternative oder Ergänzung **Werbung** an. **Interaktive Audio-Ads** erregen bei Streaming-Anbietern wie Amazon oder Spotify nicht nur die Aufmerksamkeit der Nutzer, sondern ermöglichen es auch, unmittelbar einen Dialog zu beginnen.

Beispielsweise kann auf ein zum Content passendes Produkt hingewiesen werden. Bei Interesse kann direkt über einen Beschreibungsdialog ein „Zurücklegen“ in einer Wunschliste oder ein Direktkauf erfolgen.

Allein das Entdecken einer Sprachanwendung, die um Bedürfnisse, Erlebnisse oder Fragen konzipiert ist, kann ein Bewusstsein für spätere Produktnachfrage wecken. Aufmerksamkeit geweckt. Check. – Weiter zur nächsten Phase.

### Voice Commerce in Phase 2: Consideration

Jetzt wird es ernst. Die zukünftigen Kundinnen und Kunden beschäftigen sich mit unserem Produkt oder unserer Dienstleistung. Hier kann das Angebot von Voice den entscheidenden Unterschied machen. Wer an der Kundenschnittstelle über Sprachassistenten bzw. Voice Apps nicht vorhanden ist, kann nur schlecht „in Touch“ bleiben. Wobei es nicht reicht, im Sinne des olympischen Gedankens einfach nur dabei zu sein.

Eine gute Aufarbeitung von Fragen und Antworten um die Intents der Interessierten macht den Unterschied. Die Aufmerksamkeit haben wir. Jetzt gilt es, das allgemeine Interesse auf die Vorteile unseres Angebotes zu lenken. – Aber nicht zu platt, bitte. Hier sind die Skills einer vertrauensvollen persönlichen Beratung gefragt, nicht die Qualitäten von Marktschreibern.

Die bereits verfügbaren Informationen zu den Produkten oder Dienstleistungen (Produktbeschreibungen, Auswahlhilfen, USPs, FAQs etc., aus dem Web oder anderen Medien) kann hier gut wiederverwendet werden. Diese Art der Content-Wiederverwendung stellt zusätzlich die notwendige Konsistenz sicher.

Kurz: Es geht darum, den Kundinnen und Kunden rund um die Uhr zur Verfügung zu stehen, wenn sie sich für unsere Leistungen interessieren und nähere Informationen dazu benötigen.

Und: Wer in dieser Phase schon als beratender Dialogpartner erlebt wird, schafft Vertrauen. Und das ist wichtig für Entscheidung (und mögliche kognitive Dissonanzen) sowie Erwartungen für After-Sales-Service.

### Voice Commerce in Phase 3: Conversion

Conversion. Zentraler, aber nicht einziger Punkt für Conversational Commerce. Gerade beim Kauf können dialogbasierte Ansätze die Kunden überall unterstützen. Zu Hause, unterwegs oder direkt am Point of Sale. Jetzt nicht lockerlassen.

Ein echter Voice-Vorteil: Der digitale Übergang zwischen Information, Beratung und Kaufentscheidung kann im Dialog leicht modelliert – und überall eingesetzt – werden.

Ein Kauf im Dialog ist so convenient wie mit einem Klick. „Wollen Sie das gewählte Produkt jetzt kaufen? ... Ja ... Soll ich es wieder an die Adresse wie beim letzten Einkauf liefern lassen? ... Ja.

Zukunftsvision? Nein. Käufe per Sprachbefehl sind gerade für digitale Services (Leihen oder Kaufen von Filmen) und für Verbrauchsgüter (zum Beispiel Waschmittel) sehr leicht und rechtssicher umsetzbar.

Ganz besonders bietet sich Voice Commerce für den Wieder- bzw. Nachkauf an. Anbietende und Nachfragende kennen in diesem Fall bereits das Kaufverhalten: „Darf es wieder das Waschmittel aus dem letzten Einkauf sein?“ Auf Grundlage des Verhaltens anderer Kunden eröffnet sich zudem die Gelegenheit zu Cross-Selling-Angeboten: „Wollen Sie zu Ihrem Waschmittel auch die Variante für Buntwäsche bestellen?“

Insbesondere Low-Involvement-Produkte (Produkte des alltäglichen Lebens) sind für den Einsatz von Voice Commerce prädestiniert.

Die eigentliche Abwicklung des Kaufs kann entweder mit einem kurzen Voice-Dialog einfach und schnell erledigt werden („Ihr Waschmittel ist bestellt. Ich melde mich wieder, wenn es verschickt worden ist“) oder es wird ein Say-and-Collect (in den Warenkorb einer App oder eines Onlineshops legen) realisiert.

### Voice Commerce in Phase 4: Retention

Auf den Kauf folgt die Übergabe des Produktes – durch einen Lieferdienst oder einen Mitarbeiter, beim Kunden an der Haustür oder vor Ort zur Abholung. Jetzt noch auspacken und – hoffentlich – zufrieden sein. Im Idealfall genießen. Bei bestelltem Essen ganz im Wortsinn. Aber auch bei der Lieferung eines Waschmittels sollte das gute Gefühl entstehen, alles richtig gemacht zu haben, also das richtige Produkt in der richtigen Menge zur richtigen Zeit gekauft zu haben.

Außerhalb von Nachkäufen/Wiederkäufen von Commodities bestehen mannigfaltige weitere Optionen zur Interaktion zwischen Anbietenden und Nachfragenden. Geht es zum Beispiel beim Kauf um ein hochwertiges Kochgerät, lassen sich Rezepte, Einkaufslisten, Tutorials anbieten, damit das Ausprobieren des neuen Gerätes auch zu möglichst überzeugenden Ergebnissen führt. Die Nutzererfahrung geht also über „Da habe ich etwas Schönes gekauft“ hinaus hin zu: „Mit meinem neuen Gerät habe ich aber etwas wirklich Leckeres gekocht.“ Das Angebot des dauerhaften Dialogs zur weiteren Verwendung stärkt die Kundenbeziehung weiter, im genannten Beispiel etwa mit einer Voice-Unterstützung beim Kochen durch den Kochtopf-Hersteller.

So kann aus einem rein materiellen Stolz des Besitzes ein kreativer Stolz des Könnens werden – gute Voraussetzung für eine nachhaltige Kundenbindung.

Die Anzahl möglicher Ansätze für Voice in der Nachkaufphase und der Kundenbindung ist schier unbegrenzt. Beispielhaft seien regelmäßige Reinigungen als Einsatzmöglichkeiten für Voice genannt. Im Dialog kann die Reinigung Schritt für Schritt besprochen und durchgeführt werden. Kunden erleben, wie gut sie mit ihrem Gerät umgehen können, und haben das befriedigende Gefühl, es optimal zu pflegen. Und mit anderen Ansätzen können auch Mitarbeitende in Kundenservicecentern entlastet werden. So führt Voice Commerce zur Win-win-Situation einer engeren Kundenbindung bei gleichzeitiger Kostenreduzierung.

### Voice Commerce in Phase 5: Advocacy

Empfehlungsmarketing hat eine lange Historie und ist älter als die Betriebswirtschaftslehre. Eine historische Art der Empfehlung ist „by Appointment to Her Majesty“. Das englische Königshaus pflegt noch heute die Liste der Royal Warrant of Appointment. Aktuell stehen eher Sterne-Bewertungen auf Online-Portalen und Kommentare aus den Nutzer-Communities sowie Meinungen von Influencern und Influencerinnen hoch im Kurs.

Verstärkt werden kann die Empfehlungsbereitschaft durch Dialog: „Sind Sie mit Ihrem Einkauf bei uns zufrieden? Wie würden Sie Ihren Einkauf bewerten?“ Werden zusätzliche Services genutzt, können auch diese additiven Nutzungserlebnisse Grundlage für Empfehlungen werden. „Würden Sie die Dialog-Unterstützung bei der Reinigung weiterempfehlen?“ – Oder: „Würden Sie das Rezept weiterempfehlen?“ Werden dann Rezepte getauscht, Dinner-Abende organisiert und laufend über die superleckeren Ergebnisse und die schönen Events kommuniziert, ist die Advocacy-Community perfekt.

### Wie funktioniert der Infinity Loop der Customer Journey mit Voice?

Voice Commerce kann dazu beitragen, dass die Phasen der Customer Journey immer wieder neu angestoßen und durchlaufen werden. Zum einen kann der Loop zu Cross- oder Up-Selling führen. Aus dem Dialog zu einem Produkt können der Dialog und die Unterstützung während dessen Nutzung werden und schließlich der Anstoß, sich für weitere, ergänzende Produkte oder Dienstleistungen zu interessieren. Zum anderen können Kunden über die Produktlebenszeit so stark an ihr Produkt gebunden sein, dass bei einer Ersatzbeschaffung nur das gleiche Produkt oder sein direkter Nachfolger infrage kommt. Die Reise startet dann erneut.

Gerade die Ersatzbeschaffung zeigt die Vorteile einer Bindung über einen laufenden Dialog. Im klassischen Produktnutzungskontext gibt es keine Beziehung zu den Nutzern. Dauer der Produktnutzung oder Fragen zu Reinigung und Service können im zukünftigen Kontext aber verfügbar sein. Daraus erwächst die Möglichkeit, Marken- und Produktbindung im Voice-Dialog aktiv zu unterstützen: „Ihre Waschmaschine hält schon länger als vergleichbare Modelle. – Die aktuelle Nachfolgemaschine hat neben weiteren Komfortfunktionen einen wesentlichen niedrigeren Energieverbrauch. – Darf ich Sie informieren, wenn wir die Maschine zu einem besonders günstigen Preis anbieten? – Ihre alte Maschine nehmen wir bei Lieferung einer neuen gern zurück.“ – Einfach mal drüber reden hat selten geschadet. Und gute Berater, gute Verkäufer, die das Vertrauen der Kunden verdient haben, können auf weiterhin gute Geschäfte in der Zukunft hoffen. – Eine treue, empfehlungsbereite Kundschaft haben sie schon. Auf die Zukunft. Mit Voice Commerce.

## Voice Commerce starten, aber wie?

### Ideen für eine erfolgreiche Zukunft mit Voice Commerce.

**„Hey, Sprachassistent: Was sagst du zu Visionen?“ –  
„Nicht jeder, der Visionen hat, muss zum Arzt gehen. Eine neue Geschäftsidee sollte aber auf jeden Fall eine Zielvision haben.“**

Pauschal ein „Kochbuch“ für Voice Commerce in allen Unternehmen bieten? Hoffentlich haben Sie das nicht erwartet. One-size Sprachdialog fits all? Nein. – Das haben wir nicht im Angebot. Aber ...

Ein paar – aus unserer Sicht – erfolgversprechende Ideen zum Start mit Voice Commerce sowie einige Ansätze für die Produktauswahl und für den Nutzungskontext der Dialoge haben wir aber schon.

Ein paar allgemeine Tipps vorab: Besser gut funktionierende Beispiele adaptieren als alles neu erfinden. Das macht Risiken leichter kalkulierbar. Also warum nicht von den Besten lernen und analysieren, was erfolgreiche Voice-Commerce-Anwendungen ausmacht?

Man sollte akzeptieren, dass es für viele Gestaltungsprozesse Menschen gibt, die ihre Arbeit professionell ausführen, zum Beispiel spezialisierte Agenturen mit Dialog-Designern und Voice-Architekten.

### Voice Commerce starten mit der Angebotsauswahl

**„Hey, Sprachassistent. – Wie finde ich ganz einfach den für mich und mein Unternehmen besten Einstieg in Voice Commerce?“ –  
„Nimm etwas, das für deine Kunden und Kundinnen ganz einfach zu verstehen ist. Weil einfach einfach einfach ist.“**

Produkte und Dienstleistungen, die einfach zu erklären sind, benötigen keine komplizierten Dialoge. Und je weniger Informationen übermittelt werden müssen, desto schneller und einfacher kommt man bei seinem Gespräch zu guten Ergebnissen. Eine ideale Ergänzung zu einem einfachen Inhalt ist ein möglichst geringes Fehlkaufisiko. Hier helfen traditionell Umtauschgarantien und aufwandsarme Möglichkeiten zur Rückabwicklung. Gerade dieser „Sicherheitsaspekt“ kann für die Händler-Bindung ausschlaggebend sein.

Zum anderen ist die Bindung an Produkt, Marke und Händler ein wesentlicher Aspekt. Und gerade hier kann Voice Commerce einen wichtigen Beitrag leisten und ein guter Startpunkt für Unternehmen sein, z. B. im After-Sales-Service oder Voice-Content-Marketing. Ist hier eine Bindung hergestellt, ist im Nutzungskontext ein „Ja, ich will“ auch für weitere Produkte oder Dienstleistungen schnell hergestellt.

Für einen Sprachdialog-gestützten Kauf erscheinen uns Produkte oder Dienstleistungen mit geringem Erklärungsbedarf oder solche mit Nachbestellbedarf besonders geeignet. Grundsätzlich passend sind deshalb Produkte und Dienstleistungen, die folgende Eigenschaften erfüllen:

- Produkte mit geringen Qualitätsunterschieden in einer Angebotsgruppe wie bei Mehl, Zucker, Salz. Produkte mit niederschwelligem Interesse. „Was von allen Anbietenden für mich immer gleich funktioniert, muss mir keiner erklären.“
- Produkte, mit denen die Nachfragenden bereits persönliche Erfahrung haben. „Was ich kenne, muss mir keiner erklären.“ Ideal ist Vorerfahrung mit demselben Produkt. Bei Nach- oder Wiederkauf darf das Produkt auch komplex sein.

Für Dienstleistungen gilt Analoges. Die Abholung durch ein Taxi ist eher ein Commodity. – Die erneute Buchung eines Wellness-Wochenendes „Für das nächste Wochenende – alles wie beim letzten Mal“ ist ebenso vorstellbar.

### **Voice Commerce funktioniert wie gesprochene Sprache.**

*„Hey, Sprachassistent. Für einen Voice-Dialog vertone ich einfach den Content meiner Website, oder?“*

*„Der Content deiner Website ist bestimmt eine wertvolle Basis. Gesprochene Sprache wird von Menschen allerdings anders verarbeitet als gelesener Text. Deshalb ist ein Voice-Dialog kein vertonter Website-Content. Hören und lesen funktioniert. Aber unterschiedlich.“*

Menschen hören und sprechen in anderen Strukturen und Geschwindigkeiten, als sie lesen und schreiben.

- Menschen können in der Regel fast dreimal schneller identische Inhalte sprechen als über eine Tastatur eingeben. – Eine einfach funktionierende Diktier-Software ist echte Convenience. Vielleicht ein Grund für das Diktieren von Kurznachrichten am Smartphone.
- Menschen lesen in der Regel rund 30 Prozent schneller, als sie denselben Inhalt über ihr Gehör aufnehmen können. Weniger ist deshalb mehr. Und langsamer leichter verständlich.
- Der Satzaufbau ist häufig anders, teilweise im Gesprochenen nicht einmal mit grammatikalisch vollständigen Sätzen.

Diese Erkenntnisse stellen Anspruch und Herausforderung dar: Die Inhalte für einen Voice-Commerce-Dialog sollten komprimiert und einfach verständlich sein. Sie sollten sich an gesprochener Sprache orientieren, nicht an der Schriftsprache. Die Sprechgeschwindigkeit sollte moderat gehalten sein. Der Inhalt sollte relevant, kompetent und nicht einschläfernd sein. Ob einfach wirklich einfach ist?

### Voice Commerce hat Vorbilder zum Lernen

#### Lernen von Chatbots und natürlich-sprachlichen Systemen.

Nutzerverhalten ändert sich – Erfahrungen sind dabei prägend. Was als „Standard“ durch Marktführer erlebt wird, das prägt. GAFA oder GAFAM sind prägend für das, was Nutzende in der digitalen Welt erwarten.

Nutzer stellen deshalb Fragen an Bots im Volltext wie an einen menschlichen Dialogpartner, verwenden aber auch Ein-Wort-Sätze oder einfache Wortgruppen. Gern mit oder ohne fragende Betonung.

Das korrekte Erkennen der Kunden-Intents als eine Basisarbeit kann aus der Systematik der Entwicklung von Chatbot-Dialogen abgeleitet werden. Auch die Erfahrungen aus der Implementierung und Optimierung von Sprachdialogsystemen (natürlich sprachlich, bitte) können hilfreich sein.

## Voice Commerce: Inspirationen durch Beispiele

Aus unserer Sicht ist der Weg klar. Commerce wird immer digitaler. Next Level: Voice Commerce. Noch nicht überzeugt? Dann helfen vielleicht zusätzliche Anregungen für den Schritt zu Voice-Commerce-Projekten.

#### Unsere Case Studies:

- **Case Study 1:**  
Mit Sprache vom Briefkasten zur Lieferung.  
Von der Awareness zur Conversion – ein Schuhkauf.
- **Case Study 2:**  
Der Inhouse-Versicherungsvertreter.  
Von der Awareness zur Conversion – Voice auch für Dienstleistungen.
- **Case Study 3:**  
Interaktive Produktbegleitung von Fast-Moving Consumer Goods (FMCG):  
After-Sales, Nutzungserlebnis anreichern, Kundenbindung, mehr Markenbindung, mehr Sales.
- **Case Study 4:**  
Die Mittagstheke der Metzgerei Kaufmann. Voice auch für kleinere Unternehmen:  
Service und Bequemlichkeit, Kundenbindung, mehr Wiederkäufe.

Die Case Studies sind im Stil ihrer jeweiligen Autorinnen bzw. Autoren gehalten und weichen dadurch vom restlichen Dokument ab.

## Case Study 1: Mit Sprache vom Briefkasten zur Lieferung

### Kurzinfo zur Case Study: Fokus Handel, Cross-Channel: analog/Print und digital

Für den Handel zeigt diese Case Study, wie in einer Cross-Channel-Anwendung die Werbewirkung von Print-Medien durch Sprachassistenten gesteigert und zusätzliche Käufe generiert werden können.

### Beschreibung des Anwendungsfalls

Silke kommt gerade vom Briefkasten und hat eine Menge Papier in der Hand: mehrere Flugblätter von Händlern vor Ort, einige Direktmailings als Brief und eine Zeitschrift (in der natürlich auch Werbeanzeigen zu finden sind). Alles genau durchzulesen, dazu hat Silke zwar keine Lust, aber wie viele Menschen schaut auch sie zumindest mal kurz über alles drüber, ob nicht doch das eine oder andere Interessante für sie dabei ist. Und siehe da: In der Werbung eines Schuhgeschäfts fallen ihr ein paar Stiefeletten ins Auge, die ihr gefallen.

Es gibt aber ein Problem: Silke ist mit ihrem kleinen Kind zuhause und später kommt noch eine Freundin vorbei. Außerdem macht ihr Rücken heute mal wieder Probleme. Deshalb verwirft sie die Idee, spontan in den Schuhladen zu gehen. Das wäre ihr alles viel zu mühsam – noch dazu weiß sie ja gar nicht, ob die Stiefeletten überhaupt in ihrer Schuhgröße im Laden wären.

Das hätte es auch schon sein können mit dem (Nicht-)Umsatz für den Händler. Doch dann sieht Silke neben den Schuh-Angeboten einen Hinweis auf den Google Assistant. Da steht: „Sagen Sie einfach: Hey Google, rede mit Schuh Meier und frage, ob im Nürnberger Laden die Stiefelette ‚Sommerbrise‘ noch in deiner Größe vorrätig ist.“

Als eine von gut 50 Prozent aller Deutschen, die einen Sprachassistenten benutzen, ist Silke begeistert. Ohne die Werbung oder das Kind aus der Hand legen zu müssen, kann sie ihre Frage einfach aussprechen und der Sprachassistent antwortet sofort: „Ja, dieser Schuh ist in deiner Größe – Schuhgröße 40 – noch vorrätig. Soll ich ihn dir – natürlich unverbindlich – für 24 Stunden zurücklegen?“ „Ja klar, super“, antwortet Silke, und der Sprachassistent sagt noch: „Geht in Ordnung, die Stiefeletten Sommerbrise in Größe 40 sind bis morgen 13:45 Uhr für dich reserviert. Nenne im Laden einfach dein persönliches Reservierungs-Codewort, das ‚Vierzehn Sonnentage‘ lautet. Und wenn du das vergisst, kannst du mich jederzeit nochmal danach fragen.“

Am nächsten Morgen macht Silke mit dem Kinderwagen einen Spaziergang, geht nebenbei im Schuhgeschäft vorbei. Mit ihrem Codewort erhält sie die Stiefeletten einfach und schnell zur Anprobe. Der Schuh gefällt ihr sehr gut, er passt, wird gekauft und mitgenommen.

Übrigens: Hätte Silke bei diesem Anbieter bereits ein Kundenkonto gehabt und dieses mit der Voice App verbunden, hätte sie neben der Reservierung auch die Möglichkeit angeboten bekommen, den Schuh mit einem Lieferservice direkt zu ihr nach Hause bringen zu lassen – eine weitere große Erleichterung, nicht nur in Pandemie-Zeiten.

### Fachliche Einordnung dieser Sprachassistenten-Anwendung

#### Commerce- / Umsatz-Relevanz

Interesse und Kaufimpulse entstehen in diesem Beispiel erst einmal, wenn die klassischen Werbemittel angesehen werden. Oft führt dieser erste Impuls dann aber doch nicht zu einem Kauf, weil ein direkter „Rückkanal“ fehlt. Potenzielle Käufe gehen so verloren. Im Zeitraum bis

zum nächsten Ladenbesuch kann eine Alternative gefunden worden sein. Vielleicht wird der (ja erst mal sogar erfolgreiche) werbliche Impuls aber auch ganz einfach vergessen.

Mit Hilfe der Voice App eines Sprachassistenten (z. B. Google Assistant) steht das anbietende Unternehmen im Moment des Impulses zum Dialog bereit – und das einfach und komfortabel per Sprache. So wird der Weg im Conversion Funnel deutlich verkürzt und die Kaufwahrscheinlichkeit erhöht sich. Sprachassistenten können so die Werbewirkung von Printmedien steigern.

### Allgemeine Eignung für andere Unternehmen

Ein Say & Collect („Sagen und Abholen“) ist geeignet für Händler unterschiedlicher Branchen, die stationäre Filialen und/oder Lieferservices betreiben. Aber auch für Webshop-Betreiber kann es interessant sein, wenn Nutzende z. B. offline beworbene Produkte einfach per Sprache in den Warenkorb oder auf eine Merkliste zur späteren Bestellung legen können.

### Aufwand / Projektgröße

\*\*\* (3/5)

## Case Study 2: Der Inhouse-Versicherungsvertreter

### Kurzinfo zur Case Study

Im Kapitel 3 haben wir festgestellt, dass eher einfache Low-Involvement-Produkte für Voice Commerce geeignet sind. Und nicht alles, was auf den ersten Blick kompliziert erscheint, ist es auch.

In dieser Case Study wollen wir mit dem Vorurteil aufräumen, dass Versicherungsprodukte per se kompliziert sind. Auch bei diesen Produkten gibt es einige „Direct Sales“-fähige Varianten. Diese eignen sich gut für Voice Commerce. Beispiele für Kfz-Versicherungen sowie weitere Optionen zeigt diese Case Study.

### Beschreibung der Case Study

Wer über ein Kleinstfahrzeug wie einen Motorroller verfügt, weiß: Es muss jedes Jahr ein neues Versicherungskennzeichen gekauft und am Fahrzeug angebracht werden. Meist ist dazu kein weiterer Informationsaustausch notwendig. Physisch beim Vertreter, per E-Mail, postalisch oder online kann die jährliche Versicherung erneuert werden. Gerade dieses Szenario eignet sich perfekt für den Kauf per Sprachbefehl. Die Tarife sind nach Laufzeit gestaffelt und leicht zu berechnen. Des Weiteren sind die Kosten sehr überschaubar.

Viele Mütter und Väter wissen es: Wenn die Kinder im „Führerscheinalter“ sind, verbringt man oft Zeit mit Anrufen bei Callcenter-Mitarbeitenden. Nach anfänglicher Aufklärung über neue Versicherungsbedingungen oder Familientarife sind die Optionen unkompliziert. Auch in diesem Fall wäre eine kurze Beratung per Sprachassistent mit anschließender Tarifierfassung – ohne Umwege und rund um die Uhr – eine ausgesprochen bequeme Alternative.

Abgesehen von Fahrzeugversicherungen gibt es viele weitere Policen, die sich per Sprachbefehl kaufen lassen. Reiserücktrittsversicherungen sind kostengünstig und werfen wenig Fragen auf. Wer seine Flüge direkt bucht, vergisst diese aber oft. Mit einem Sprachassistent ließe sich eine Reiserücktrittsversicherung schnell abschließen.



Die oben genannte drei Beispiele haben eine weitere Besonderheit über ihre Eignung als Voice-Commerce-Produkte hinaus: Sie beschreiben allesamt Szenarien, die sich als Teil eines Customer-Journey-Loops wiederholen. Im Gegensatz zu einer Lebensversicherung, die fast immer nur einmalig und nach einer intensiven persönlichen Beratung abgeschlossen wird, kommen kleinere Anpassungen oder neue kleine Policen immer wieder hinzu.

### **Fachliche Einordnung dieser Sprachassistenten-Anwendung**

#### **Commerce- / Umsatz-Relevanz**

Zwei Faktoren sind in dem Kontext von Business-Value besonders relevant: das Potenzial erhöhter Umsätze, weil die Inanspruchnahme von Versicherungsprodukten durch Absenkung der Hürden erheblich vereinfacht wird. Man muss nicht mal zum Telefon greifen oder zum Laptop laufen. Eine Aussage und kurze Unterhaltung genügen. Des Weiteren würden die Kosten im Kundenservicebereich weiterhin sinken, weil viele der häufigen „niederschwelligen“ Anfragen automatisch abgearbeitet werden.

#### **Allgemeine Eignung für andere Unternehmen**

Im Grunde genommen eignet sich der Einsatz von Sprachassistenten als zusätzlicher Vertriebs- und Kommunikationskanal sehr gut für Versicherungsgesellschaften wie auch für Anbieter anderer Leistungen, deren Produkte vielfältig sind und (oft) über Mikrokommunikation in Form von kurzen Anrufen vertrieben werden.

#### **Aufwand / Projektgröße**

\*\*\* (3/5)

Da die involvierten Produkte vergleichsweise einfach strukturiert sind, können die Dialoge im Sprachassistenten mit wenig Aufwand gestaltet werden. Sicherlich ist dann die Einbindung in die Backend-Systeme, die die Vertriebsprozesse steuern, ein Kostentreiber. Ferner spielen – wie in der Finanzbranche üblich – Sicherheitsbedingungen und Datenschutz eine höhere Rolle als in normalen E-Commerce-Anwendungen.

### **Case Study 3: Interaktive Produktbegleitung von Fast-Moving Consumer Goods (FMCG)**

#### **Kurzinfo zur Case Study**

Hersteller von Konsumgütern stehen oft vor der Herausforderung, ihrer Marke ein Alleinstellungsmerkmal zu verschaffen und den Konsum mit mehr Emotionen aufzuladen. Wer eine emotionale Bindung zu einem Produkt verspürt, kauft häufig dasselbe nochmal. Diese Case Study zeigt, wie Voice Apps auf Sprachassistenten den Konsumenten Mehrwerte bieten, das Produkt emotionalisieren und so zur Bindung an die Marke beitragen können.

#### **Beschreibung der Case Study**

Ein gutes Beispiel für Produkte mit vergleichsweise geringer assoziativer Emotionalität sind Gesichtsmasken oder andere Pflege- und Wellnessprodukte. Zumeist wird ein guter halber Regalgang im Supermarkt mit Hunderten von solchen Produkten konfektioniert. Der Kauf des Produkts ist zweckgebunden: Es steht auf der Einkaufsliste, und wenn es auch auf der Packung steht, passt das schon. Die Marke ist oft zweitrangig.

Versuche, diese Gleichgültigkeit durch Begeisterung zu ersetzen, sieht man an der bunten Vielfalt im Regal – die oft eher verwirrt als hilft. Es stellt sich die Frage: Wie kann man einen Anreiz kreieren, um das Erlebnis mit solchen Produkten dermaßen anzureichern, dass Kunden künftig bei der Marke und in dem Sortiment der Marke bleiben?

Relevant ist Voice auch (schon). Eine aktuelle Studie (Quelle: Beyto Smart Speaker & Voice Studie 2021–2022) liefert hier die Informationen:

- über 50 % Voice-Nutzung in Deutschland (Smartspeaker, auf Smartphone, TV-Steuerung etc.) ... „zumindest gelegentlich“
- rund 66 % Voice-Nutzung bei 18- bis 34-Jährigen
- über 20 % Besitz eines Smart Speakers in Deutschland im Durchschnitt
- rund 34 % Besitz eines Smart Speakers in Deutschland bei 18- bis 34-Jährigen

Da die Anwendung von Pflegeprodukten – im Gegensatz zu anderen Hygieneprodukten, wie zum Beispiel Duschgel – eher als kontemplatives, entspannendes Erlebnis empfunden wird, wäre es durchaus denkbar, den Konsum durch entsprechende Audio-Inhalte, z.B. in Form von passenden Klängen und Melodien, deren Länge auf die empfehlende Anwendungs- oder Einwirkdauer des Produktes zugeschnitten ist, zu begleiten.

Die Idee ist, auf jeder Packung eine entsprechende „Anleitung“ zu platzieren, um mit dem Sprachassistent ein Hörerlebnis zu starten. Denn häufig ist es der Fall, dass die Anwendung von Pflegeprodukten ein gewisses Zeitkontingent von einigen Minuten in Anspruch nimmt, welches sonst anderweitig überbrückt werden müsste. Ein Befehl wie „Alexa, rede mit (Markenname) und starte das Wellness-Programm ‚Frühlingsluft!‘“ würde als Invokation die Sprachanwendung öffnen und das produktbegleitende Erlebnis initiieren. Die entsprechende Musik und Wellness-Tipps schaffen für den Konsumenten in dem Augenblick einen erheblichen Mehrwert – und geben der Produktreihe ein mächtiges Alleinstellungsmerkmal. Das stärkt die Kundenbindung und beschleunigt – begleitet von entsprechenden Werbemaßnahmen – das Anwerben neuer Kunden.

Da gerade bei Konsumgütern der Fokus eben auf Konsum liegt, sind Smart-Speaker besonders geeignet als Kommunikationskanal. Sie sind omnipräsent und bieten ein Audio-Erlebnis an. Man ist nicht gezwungen, in der Nähe vom Laptop oder Fernseher zu bleiben oder das Smartphone in die Hand zu nehmen.

### **Fachliche Einordnung dieser Sprachassistenten-Anwendung**

#### **Commerce- / Umsatz-Relevanz**

In diesem Case steht Umsatzsteigerung durch Sichern und Ausweiten der Marktanteile und Stärkung der Kundenbindung an die Marke an erster Stelle. Andere „klassische“ Szenarien für konversationelle Schnittstellen, wie Bots oder Informationsdienste im Callcenter, können eine Rolle spielen, aber Unterhaltung und Wellness sind hier präsenter als eine reine Anleitung zur Anwendung der Produkte. Besonders deutlich ist die Relevanz des Customer-Journey-Loops in diesem Case: Die Kaufentscheidung wird positiv durch Neugier auf das Erlebnis beeinflusst, die Bindung erfolgt durch das Erlebnis selbst.

#### **Allgemeine Eignung für andere Unternehmen**

Alle Hersteller und Vertreiber von Konsumgütern können theoretisch von diesem Ansatz profitieren. Besonders wichtig dabei: Die Storyline muss stimmen.

Da solche Konsumgüter durch ihre Bekanntheit das Vertrauen der Verbraucher genießen, eignen sie sich – über den eben beschriebenen Case des Markenerlebnisses und der Kundenbindung hinaus – auch sehr gut für den Direktverkauf per Sprachbefehl (siehe Case 2 – „der Inhouse-Versicherungsvertreter“). Ferner ist es – nach Akzeptanz dieses Werbe- und Vertriebskanals durch den Verbraucher – durchaus denkbar, exklusive Produkteinführungen zu lancieren, die nur (!) per Sprachbefehl zu erwerben wären.

### Aufwand / Projektgröße

\*\*\*\* (4/5)

Die notwendige Businesslogik und Dialog-Architektur für eine solche Anwendung ist nicht hochkomplex – man müsste sich aber zweier möglicher Kostentreiber bewusst sein. Zum einen, wenn man das reine Hörerlebnis durch Kaufmöglichkeiten ergänzen möchte, wäre es natürlich notwendig, ein Warenwirtschaftssystem im Backend zu haben, mit Zahlungsanbieter, Kunden-Login etc. Zum anderen ist eine solche „Infotainment-Anwendung“ von Natur aus Content-lastig. Oft wird der Aufwand – besonders für die Erstellung von multimedialen Inhalten – stark unterschätzt.

## Case Study 4: Die Mittagstheke der Metzgerei Kaufmann

### Kurzinfo zur Case Study

Diese Case Study zeigt, dass Voice Commerce und Sprachassistenten nicht nur für große Unternehmen und Marken sinnvolle Anwendungsfälle bieten, sondern auch kleinere und regionale Unternehmen von Voice Commerce profitieren können.

### Beschreibung des Anwendungsfalls

Mittagspause. Kollege Schmidt fragt, wer noch mitgeht, irgendwo eine Kleinigkeit zu essen. „Ich bin dabei“ kommt es aus verschiedenen Ecken des Büros. Fragt sich nur, wohin: zum Döner direkt ums Eck, zum Chinesen oder zu der etwas weiter entfernten Metzgerei, die an der Mittagstheke der Bistro-Ecke leckere und täglich wechselnde Mittagsgerichte bietet. Früher gab es dann immer die Diskussion, ob man den etwas weiteren Weg zu der Metzgerei wirklich auf sich nehmen wolle, wisse man doch gar nicht, was es aktuell als Tagesgerichte gibt und ob man darauf Lust hat. Und mit diesem Risiko der Unsicherheit des Angebots und des etwas weiteren Wegs erschien dann das Metzgerei-Bistro doch nicht so attraktiv und es gab wieder „den Standard-Döner direkt ums Eck“.

Das war einmal. Seit neuestem ist aber Alexa im Team. Und täglich heißt es seitdem: „Alexa, frage Metzgerei Kaufmann, was es heute als Tagesgericht gibt“. Niemand muss erst an einen Rechner gehen und einen Browser öffnen, eine URL eingeben und die richtige Information suchen. Und niemand greift erst aufwendig extra zum Telefon (und die Metzgerei Kaufmann muss auch keine Mitarbeitenden dafür bezahlen, Telefonanrufe entgegenzunehmen).

Die potenziellen Mittagskunden nutzen ganz einfach die Voice App (bei Amazon „Skill für Amazon Alexa“, genannt) der Metzgerei Kaufmann. Und deshalb muss nur irgendjemand im Büro die Frage in den Raum sprechen und schon weiß man, was es heute Leckerer an der Mittagstheke des Metzgerei-Bistros gibt.

Seit das so einfach ist, nehmen Herr Schmidt und seine Kollegen viel öfter den etwas weiteren Weg zur Metzgerei auf sich – haben sie doch jetzt die Sicherheit zu wissen, was es dort zu Mittag gibt. Und auch die eine oder andere Kollegin, die eigentlich nicht mitkommen wollte, entscheidet sich manchmal spontan doch dazu, wenn sie hört, dass es heute eines ihrer Lieblingsgerichte gibt.

Jetzt wünscht sich das Kollegenteam nur noch, dass man die Gerichte über Alexa auch noch vorbestellen bzw. reservieren könnte. Oder zumindest sollte Alexa ihnen sagen, ob noch ausreichend Portionen vorhanden sind, bevor sie loslaufen. Aber das könnte ja eine Erweiterung im nächsten Update der Voice App sein ...

### **Fachliche Einordnung dieser Sprachassistenten-Anwendung**

#### **Commerce- / Umsatz-Relevanz**

Hier wird der Sprachassistent als Vertriebsinstrument genutzt, insbesondere für die Informations- und Entscheidungsphase vor dem Kauf. Die Metzgerei ermöglicht es Kundinnen und Kunden, auf sehr einfache Art (nur Aussprechen einer Frage) Informationen zum aktuellen Angebot und dessen Verfügbarkeit zu erhalten. So können Anbieter die Kaufentscheidung im entscheidenden Moment – kurz vor der eigentlichen Entscheidung – zu ihren Gunsten beeinflussen.

Ein Reservierungs- oder gar Bestellservice mit Lieferung würde diese Anwendung noch um die Möglichkeit einer direkten Conversion (Kauf) sinnvoll erweitern.

#### **Allgemeine Eignung für andere Unternehmen**

Eine solche Anwendung ist analog geeignet für jedes Unternehmen, das regelmäßig wechselnde Angebote hat, seien es Tages- oder Wochenangebote von kleinen regionalen Unternehmen oder auch von großen Filialketten.

Darüberhinaus ist die Abfragbarkeit von für die Kaufentscheidung direkt relevanten Informationen jeder Art ein grundsätzlicher vertriebsorientierter Ansatz für Voice Apps auf Sprachassistenten.

#### **Aufwand / Projektgröße**

\* (1/5)

Es handelt sich um einen vergleichsweise einfach umsetzbaren Anwendungsfall. Die Anzahl der möglichen Dialoge ist klar begrenzt, das Voice Design klar auf einen Anwendungsfall fokussiert und die notwendige Programmierlogik eher simpel.

## Voice Commerce. What's Next? ...

### Kann eine unendliche Reise vorbei sein?

Die ersten Erlebnisse mit einem neuen Smart-Speaker sind oft sehr durch den Unterhaltungswert geprägt: Man freut sich über den Wetterbericht und hat Spaß an Chuck-Norris-Witzen.

Es ist gerade dieser Infotainment-Ansatz, der viele Entwickler und Unternehmen dazu bewegt hat, die erste Generation von Skills und Actions an den Markt zu bringen – natürlich beflügelt durch Smart-Home-Anwendungen und Spaß am Basteln. Chit-Chat ist Teil der „Sprich mit mir“-Welt.

Low-Code-Entwicklungsumgebungen und recht freie und gute Unterstützung durch die Marktführer trugen zusätzlich dazu bei, dass die Anzahl der Anwendungen in den letzten Jahren rasch weiter gewachsen ist. Es kann aber auch beobachtet werden, dass viele der Angebote – besonders von Großkonzernen – nach der Veröffentlichung wieder vom Markt genommen werden.

Grund dafür ist oft, dass die Nutzung – und natürlich der Business Value – zum Beispiel aufgrund mangelhafter Dialog-Architektur unter den Erwartungen blieb. Nun wird die zweite Generation vieler Anwendungen an den Markt gebracht, weil Konzerne festgestellt haben, dass mit positiven Nutzererlebnissen der Case besser wird. Die Sprachanwendung verhindert ein Verdrängen von der Sprach-Schnittstelle. Mit Voice Commerce werden zusätzliche, omnipräsente Vertriebskanäle eröffnet.

Anekdotisch dargestellt gibt es zwei Szenarien, die heutzutage sehr oft beobachtet werden: zum einen meist jüngere Handy-Nutzer, für die Smartphone mehr smarte Geräte sind als Telefone – und für die „Telefonieren“ nur noch eine der Anwendungen am Rande ist. Ihr Verhalten hat sich geändert.

Horizontal und weggerichtet vor dem Mund wird „telefoniert“. Oder diktiert: für den Versand von Nachrichten als Sprach- oder Textnachricht.

Oft ist dabei die Kommunikation inzwischen nicht mehr an einen Menschen gerichtet, sondern an einen Sprachassistenten. In dem Fall wird das Gesprochene verschriftlicht und dann weiterverarbeitet. Für „Post-Millennials“, die vielleicht heute noch nicht ihre volle Kaufkraft entwickelt hat, ist die Nutzung von Sprach-Schnittstellen selbstverständlich – Tippen wird zunehmend als lästig empfunden. Diese Generation erwartet, ihr künftiges Kaufverhalten per Sprachbefehl auszusteuern. Voice Commerce kommt. Wir empfehlen, dabei zu sein.

## Autoren und Autorinnen

### **Daniel Fitzpatrick**

Director of Conversational Solutions &  
Leader of the Reply Community of Practice for Voice Machine Interfaces,  
Triplesense Reply

### **Maria Müller**

Head of Voice UX Design,  
Triplesense Reply

### **Thomas Hörner**

Berater, Autor,  
Redner und Dozent

### **Michael Fischer**

Abteilungsleiter MDA3 Content & Channel Management,  
KfW Bankengruppe

### **Kontakt:**

### **Caroline Eichholz**

Referentin Digital Marketing,  
Bundesverband Digitale Wirtschaft (BVDW) e. V.

## Anhang

### Glossar

#### Actions

Actions werden beim Google Assistant (s. u.) die Voice Apps (s. u.) genannt, die dessen Funktionalität um eigene Dialoge erweitern.

#### Alexa

Alexa ist der Sprachassistent (s. u.) von Amazon. Er ist einer der bei Konsumenten am weitesten verbreiteten Sprachassistenten.

#### Conversational Commerce

Voice Commerce umfasst auf menschlicher Sprache basierende Informations-, Kommunikations- und Transaktionsprozesse in digitalen Netzen, die das Ziel haben, Waren oder Dienstleistungen anzubieten und abzusetzen. Insbesondere zählen dazu die Marketingkommunikation mit Messengern, Chatbots und Sprachassistenten (Letzteres wird unter dem Begriff Voice Commerce (s. u.) zusammengefasst).

#### GAFA oder GAFAM

GAFA ist das Akronym für Google, Apple, Facebook und Amazon. – Als GAFAM inklusive Microsoft. In dieser Version auch als Big Tech oder Big Five bekannt.

#### Google Assistant

Google Assistant heißt der Sprachassistent (s. u.) von Google. Er ist einer der bei Konsumenten am weitesten verbreiteten Sprachassistenten.

#### Skills

Skills werden bei Alexa (s. o.) die Voice Apps (s. u.) genannt, die deren Funktionalität um eigene Dialoge erweitern.

#### Smart Speaker

Smart Speaker sind Stand-alone-Geräte zum Zugriff auf Sprachassistenten (s. u.). Sie bestehen im Wesentlichen aus Lautsprecher, Mikrofon und eingebetteter Elektronik. Für ihre Funktion benötigen Sie einen aktiven Internetzugang für die Verbindung zur Sprachassistenten-Software. Weit verbreitete Smart Speaker sind – neben einer großen Vielfalt am Markt – Amazon Echo und Google Home bzw. Nest.

#### Sprachassistenten

Sprachassistenten sind Software-Systeme, die mittels Kommunikation in natürlicher, menschlicher Sprache Informationen abfragbar machen, Dialoge führen und Assistenzdienste der verschiedensten Art erbringen können. Sprachassistenten werden oft über Smart Speaker (s. o.) verwendet, sie können aber auch – je nach Sprachassistent – über Smartphone-Apps, in Kfz oder Haushaltsgeräte integriert oder als proprietäre Systeme z. B. durch Telefonanrufe genutzt werden. Weit verbreitete Sprachassistenten sind Amazons Alexa, Apples Siri oder der Google Assistant.

#### Sprach-Dialog-Systeme (SDS)

Sprach-Dialog-Systeme sind gekennzeichnet durch einen bidirektionalen sprachlichen Dialog zwischen Mensch und Maschine. Diese Systeme haben im Gegensatz zu Sprachassistenten (s. o.) meist einen starren, vorab festgelegten Dialogprozess und sind meist auf einen klar begrenzten Aufgabenbereich beschränkt. Beispiele sind die bekannten Telefon-Dialog-Systeme.

### Voice App

Eine Voice App ist ein Programm, mit dem man eigene Dialoge und Funktionalitäten in Sprachassistenten integrieren kann. Voice Apps können analog zu Smartphone-Apps verstanden werden (die das Smartphone-Betriebssystem um eigene Funktionalitäten erweitern). Bei Alexa (s. o.) heißen Voice Apps dann Skills (s. o.), beim Google Assistant (s. o.) heißen sie Actions (s. o.).

### Voice Commerce

Unter dem Begriff Voice Commerce werden Aktivitäten zusammengefasst, bei denen digitale Sprachschnittstellen (z. B. Sprachassistenten (s. o.)) zum Einsatz kommen, um reale oder elektronische Waren sowie Dienstleistungen anzubieten und abzusetzen. Zu unterscheiden ist der Begriff vom Conversational Commerce (s. o.), der mehr umfasst.

### Voice Marketing

Voice Marketing sind Marketing-Strategien und -Taktiken, bei denen für die Kommunikation mit der Zielgruppe im Wesentlichen die gesprochene Sprache das zentrale Kommunikationsmittel darstellt.

## BVDW-Publikation zum Thema Conversational Commerce

BVDW (2019): Conversational Commerce – Time to talk. Berlin: Bundesverband Digitale Wirtschaft, 2019. Kostenfreier Download unter:

<https://www.bvdw.org/der-bvdw/news/detail/artikel/neues-bvdw-kompendum-wie-smart-speaker-und-sprachassistenten-den-handel-zum-conversational-commerce/>

## Quellen und weiterführende Literatur

<https://www.funkschau.de/markt-trends/mit-sprachinterfaces-altershuerden-ueberwinden.178638.html>

<https://t3n.de/news/fuer-voice-technologie-dating-1258911/>

[https://www.wuv.de/specials/primetime\\_fuer\\_audio/wie\\_man\\_das\\_kundenerlebnis\\_ueber\\_die\\_stimme\\_steigert](https://www.wuv.de/specials/primetime_fuer_audio/wie_man_das_kundenerlebnis_ueber_die_stimme_steigert)

<https://www.funkschau.de/markt-trends/mehr-empathie-und-emotionen-bitte.172528.html>

<https://www.beyto.com/smart-speaker-voice-studie-2021-2022/>

<https://sprachassistenten.marketing/literatur/buch-marketing-mit-sprachassistenten/>



## **Bundesverband Digitale Wirtschaft (BVDW) e. V.**

Der Bundesverband Digitale Wirtschaft (BVDW) e. V. ist die Interessenvertretung für Unternehmen, die digitale Geschäftsmodelle betreiben oder deren Wertschöpfung auf dem Einsatz digitaler Technologien beruht. Als Impulsgeber, Wegweiser und Beschleuniger digitaler Geschäftsmodelle vertritt der BVDW die Interessen der digitalen Wirtschaft gegenüber Politik und Gesellschaft und setzt sich für die Schaffung von Markttransparenz und innovationsfreundlichen Rahmenbedingungen ein. Sein Netzwerk von Experten liefert mit Zahlen, Daten und Fakten Orientierung zu einem zentralen Zukunftsfeld. Neben der DMEXCO und dem Deutschen Digital Award richtet der BVDW eine Vielzahl von Fachveranstaltungen aus. Mit Mitgliedern aus verschiedensten Branchen ist der BVDW die Stimme der digitalen Wirtschaft.

### **Digital Commerce**

Die Fokusgruppe Digital Commerce im BVDW steht als Thought Leader für die Digitalisierung der Customer Journey. Nicht nur Händler und Brands, sondern auch Logistikunternehmen, Plattformbetreiber, Fintechs und Dienstleister gestalten die Debatte um die aktuellen Herausforderungen. Dabei gibt es eine Vielzahl an Playern und Stakeholdern, deren Interessen sich mit der fortschreitenden Digitalisierung und neuen Technologien wie Künstlicher Intelligenz und Voice Assistants verändern. Wir stehen für Transparenz, Aufklärung und sind Meinungsgeber für praxisnahe Strategien zur Umsetzung einer erfolgreichen Digitalisierung der Customer Journey.

**[www.bvdw.org](http://www.bvdw.org)**



DIGITAL COMMERCE  
FOKUSGRUPPE IM BVDW

Impressum

## Voice Commerce – Ein Leitfaden für sprachgesteuerten digitalen Handel

Erscheinungsort und -datum	Berlin, Juli 2022
Herausgeber	Bundesverband Digitale Wirtschaft (BVDW) e.V. Schumannstraße 2, 10117 Berlin, +49 30 2062186 - 0, info@bvdw.org, www.bvdw.org
Geschäftsführer	Sven Bornemann
Präsident	Dirk Freytag
Vizepräsidenten	Thomas Duhr, Anke Herbener, Corinna Hohenleitner, Dr. Moritz Holzgraefe, Alexander Kiock, Julian Simons
Kontakt	Caroline Eichholz, Referentin Digital Marketing, eichholz@bvdw.org
Vereinsregisternummer	Vereinsregister Düsseldorf VR 8358
Rechtshinweise	Alle in dieser Veröffentlichung enthaltenen Angaben und Informationen wurden vom Bundesverband Digitale Wirtschaft (BVDW) e.V. sorgfältig recherchiert und geprüft. Diese Informationen sind ein Service des Verbandes. Für Richtigkeit, Vollständigkeit und Aktualität können weder der Bundesverband Digitale Wirtschaft (BVDW) e.V. noch die an der Erstellung und Veröffentlichung dieses Werkes beteiligten Unternehmen die Haftung übernehmen. Die Inhalte dieser Veröffentlichung und / oder Verweise auf Inhalte Dritter sind urheberrechtlich geschützt. Jegliche Vervielfältigung von Informationen oder Daten, insbesondere die Verwendung von Texten, Textteilen, Bildmaterial oder sonstigen Inhalten, bedarf der vorherigen Zustimmung durch den Bundesverband Digitale Wirtschaft (BVDW) e.V. bzw. die Rechteinhaber (Dritte).
Ausgabe	Erstausgabe
Titelmotiv	© iStock / Rudzhan Nagiev / Macrovector