

Niels Dahnke

VOICE SEARCH – EIN HYPE, DER (NOCH) KEINER IST

Folgt man der Berichterstattung der letzten Jahre, müsste Voice Search bereits eine der wichtigsten Online-Technologien überhaupt sein. Doch die sprachbasierte Suche stößt bei Online-Marketern durchaus auf geteilte Meinungen. Woran liegt es, dass es bisher scheinbar keinen Konsens hinsichtlich der Potenziale der Sprachsuche gibt? Wie ist der derzeitige Stand der Technologie und wohin bewegt sich der Trend? All das wird Niels Dahnke, Head of SEO bei suxeedo, im Folgenden einmal näher betrachten.



DER AUTOR



Niels Dahnke ist Head of SEO bei der Content-Marketing-Agentur suxeedo. Er zählt zu den SEOs der ersten Stunde und ist seit 1993 als Entwickler im Internet unterwegs – sowohl für eigene Projekte als auch im Kundenauftrag. Technische Suchmaschinenoptimierung (mit einem Fokus auf den Themen Strategieentwicklung und Relaunch) ist sein Spezialgebiet.

Zunächst einmal ist das Thema sehr viel komplexer, als es auf den ersten Blick erscheint. Sprachsuchen werden in verschiedenen Kontexten getätigt, ob im Auto via Apple Carplay oder Android Auto, mobil mit dem Smartphone oder zu Hause per Smart Speaker. Dementsprechend unterscheiden sich auch das jeweilige Nutzerverhalten und die Such-Intentionen. Je nach Kontext und Nutzer-Intention stellt sich heraus, ob die Sprachsuche relevant für Unternehmen ist oder nicht. Dabei ist es entscheidend, genau zu differenzieren. Voice Search befindet sich noch immer in der Test- und Entwicklungsphase – hier wird noch viel passieren.

Amazon und Google sind die größten Player im Wettbewerb

Zugegebenermaßen ist der Hype um die Sprachsuche in den letzten Monaten ein wenig abgeflacht. Dennoch findet sich eine Vielzahl an Artikeln zum Thema. Kein Wunder, haben doch alle relevanten Player bereits ihre Produkte auf den Markt geworfen bzw. launchen gerade eine neue Generation ihrer Smart Devices. In der westlichen Welt sind Amazon und Google die beiden größten Wettbewerber. Zwar sinken seit 2016 die kumulierten Marktanteile von Amazon und Google, allerdings kann dies vor allem auf die gestiegene Anzahl der Wettbewerber zurückgeführt werden. Dies wird besonders deutlich,

Foto: mrgao / gettyimages.de

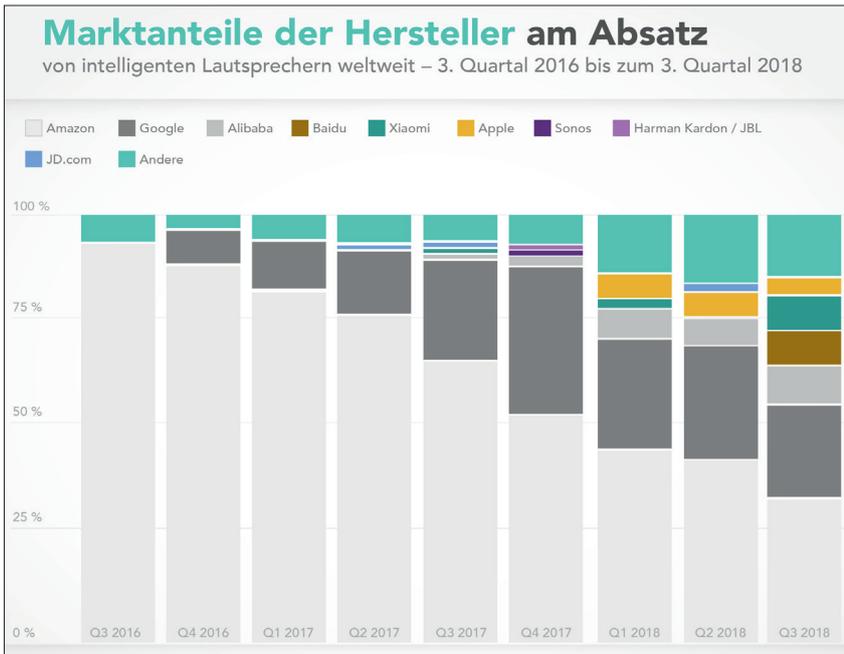


Abb. 1: Marktanteile der größten Smart-Speaker-Hersteller weltweit 2016-2018 (Quelle: Statista; <https://einfach.st/statista4>)

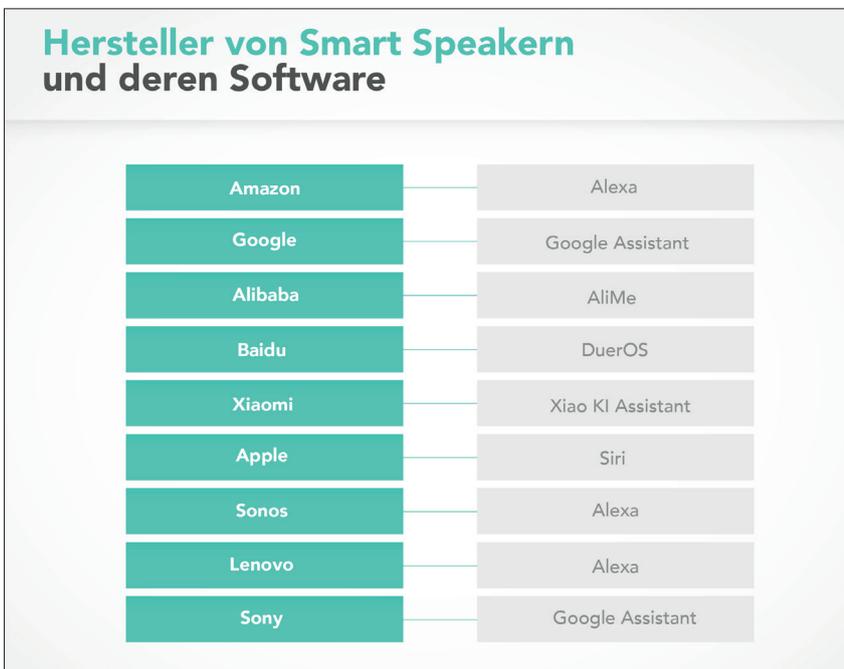


Abb. 2: Hersteller von Smart Speakern und deren Software

	Richtig beantwortet	Verstandene Abfrage
Google Assistant	87,9 %	100 %
Siri	74,6 %	99,6 %
Alexa	72,5 %	99,0 %
Cortana	63,4 %	99,4 %

Abb. 3: Die virtuellen Assistenten im Vergleich (Quelle: GoogleWatchBlog; <http://einfach.st/gwb44>)

wenn man sich die absolute Mengenzahl der abgesetzten Smart Devices ansieht. Hier konnten sowohl Amazon als auch Google ihre Absatzzahlen um ein Vielfaches vergrößern.

Wie in Abbildung 1 zu sehen, ist Amazons Marktanteil trotzdem von über 90 % in Q1 2016 auf gerade einmal 31,6 % in Q3 2018 gefallen. Dennoch bleibt Amazon der größte Player. Google folgt auf Rang zwei mit 22,7 %.

Im weltweiten Vergleich folgen darauf drei Unternehmen, die allesamt aus China stammen: Alibaba, Baidu und Xiaomi. Apple hinkt im Vergleich noch weit hinterher. Weitere Wettbewerber wie der Sonos One, Lenovos Smart Assistant oder Sonys Smart-Lautsprecher LF-S50G konnten sich auf dem hart umkämpften Markt kaum bis gar nicht durchsetzen. Abzuwarten bleibt, ob Samsung im April 2019 erfolgreich mit seinem Smart Speaker Bixby in den Markt einsteigen kann.

Hersteller setzen auf unterschiedliche Suchmaschinen

Entscheidend für den Markterfolg der Speaker ist ihre Performance. Hier liegt auch der große Unterschied zwischen Amazon und Google. Amazons Echo-Reihe greift auf den konzern-eigenen virtuellen Assistenten Alexa zurück. Google setzt hingegen auf seinen Google Assistant. Kleinere Wettbewerber wie Sony, Sonos und Lenovo nutzen die virtuellen Assistenten der beiden großen Wettbewerber, während Apple (Siri) und Microsoft (Cortana) auf eigene Lösungen zurückgreifen. Die asiatischen Wettbewerber nutzen jeweils ihre eigenen Datensets (Abb. 2) und werden im weiteren Verlauf des Artikels außer Betracht gelassen.

Der Google Assistant versteht uns anscheinend am besten

Aufgrund der unterschiedlichen Software der Sprachassistenten (und der damit verknüpften Suchmaschine) variieren die Antworten, welche die User auf ihre Suchanfrage erhalten. In einer Studie des googlewatchblogs, in der jeder Assistent 800 verschiedene (englische) Fragen beantworten sollte, wies Googles Assistant die besten Resultate auf (Abb. 3). Mit einigem Abstand folgen Apples Siri und Amazons Alexa. Im Test zeigt sich deutlich, dass Microsofts Cortana nicht mit den anderen Wettbewerbern mithalten kann. Nicht einmal zwei Drittel der Fragen konnte der Sprachassistent korrekt beantworten.

In der Studie wurden fünf verschiedene Arten von Suchanfragen untersucht: lokale, navigationale, informationale, transaktionale sowie Befehle zur Sprachsteuerung. Auffällig ist, dass der Google Assistant in vier der fünf Kategorien am besten abschnitt. Selbst bei den transaktionalen Suchanfragen, der vermeintlichen Stärke von Alexa, lag Google auf Platz eins. Lediglich in der für SEO-Experten irrelevanten Rubrik der Befehle (z. B. Öffnen einer bestimmten App) konnte Apples HomePod besser als Googles Assistant abschneiden.

Amazons Produkte punkten allerdings mit einem anderen Feature – ihren „Skills“. Dabei handelt es sich um zusätzlich installierbare Apps für den Sprachassistenten, mit denen sich weitere Fähigkeiten hinzufügen lassen. Diese reichen von Kochbüchern bis hin zur auditiven Begleitung bei Meditationsübungen. Somit kann Alexa auch Nischenwissen „erlernen“. Seit September 2018 hat Amazon seine Plattform nach außen hin geöffnet, sodass selbst Privatpersonen ihre eigenen Skills entwickeln können (www.developer.amazon.com/de/alexa-skills-kit). Goo-

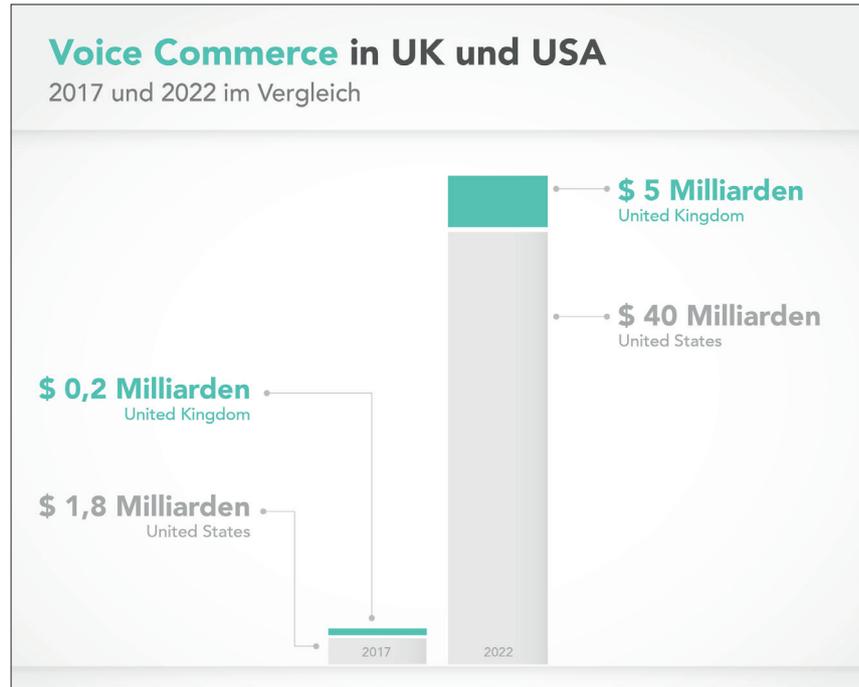


Abb. 4: Umsatz mit Voice Commerce im Vereinigten Königreich und den USA 2017 vs. 2022 (Quelle: Voicebot.ai; <http://einfach.st/voice33>)

gle hat seine mit künstlicher Intelligenz ausgestatteten Sprachassistenten ebenfalls insofern optimiert, als dass diese weitere Fähigkeiten „erlernen“ können. Die sogenannten „Actions“ sind jedoch stark eingeschränkt und stellen zumindest zum heutigen Zeitpunkt keine Konkurrenz zu Alexas Skills dar.

Wer nutzt Alexa und Co. – und wofür?

Smart Speaker hören sich sehr praktisch an, aber wer nutzt die intelligenten Assistenten wirklich? Laut einer Studie von eMarketer sind die durchschnittlichen Käufer eines Smart Speakers wohlhabende ältere Millennials. Aber auch in anderen Altersgruppen werden sie immer populärer. Dieser Trend wird sich in den nächsten Jahren weiter verfestigen. Statista prognostiziert, dass sich der Gesamtumsatz intelligenter Lautsprecher von 2017 bis 2022 mehr als verdreifachen wird. Der Hype scheint durchaus berechtigt, auch wenn der Umsatz noch lange nicht mit dem von Smartphones zu vergleichen ist. Während der Umsatz von Smart Speakern im Jahr 2017 bei \$1,5 Mrd. lag, kamen die Smartphones auf knapp

\$440 Mrd. Der Trend geht dennoch in eine eindeutige Richtung. Bis 2021 rechnet Statista mit mehr als 1,8 Mrd. Nutzern digitaler Assistenten.

Entscheidend ist aber, zwischen den unterschiedlichen Nutzungsmustern zu differenzieren. Handelt es sich um eine Sprachsuche oder Sprachsteuerung? Nur Erstere sind für SEO-Experten relevant. Zwar werden die Suchanfragen immer tiefgründiger und transaktionaler, dennoch wurden die Sprachassistenten laut einer Umfrage von Yext im letzten Jahr vor allem genutzt, um Musik abzuspielen (40 %), Wissensfragen zu beantworten (39 %), das Wetter vorherzusagen (38 %), Erinnerungsfunktionen einzurichten (31 %) und um Personen anzurufen (29 %). Erst danach folgen die für SEO-Experten relevanteren Themen wie Wegbeschreibungen (27 %), Unternehmensinformationen (21 %), Online-Shopping (16 %) und Online-Essensbestellungen (13 %).

Schaut man sich die Entwicklung der Suchanfragen in den letzten Jahren an, wird aber deutlich, dass der Trend von den navigational getriebenen Sprachbefehlen (Apps öffnen) hin

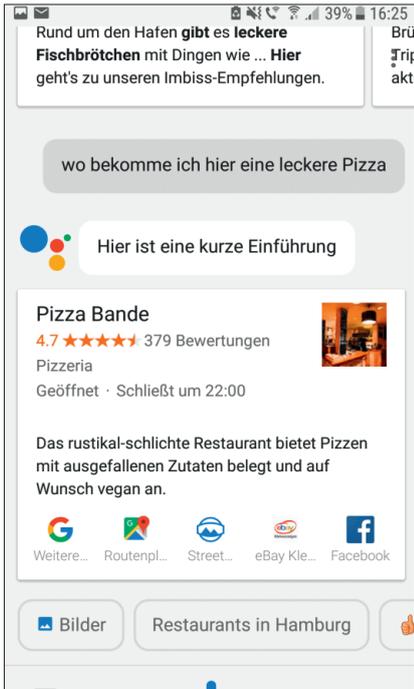


Abb. 5: Beispiel für eine lokale Sprachsuche via Google Assistant mit Display-Anzeige (Quelle: omt.de; <http://einfach.st/omt32>)

zu informativen und transaktionalen Sprachsuchen geht, die für Online-Marketer weitaus spannender sind. Gerade die transaktionalen Sprachsuchen erfordern eine flüssige und reibungs-freie Bedienung und Wiedergabe der Inhalte, um die UX möglichst organisch zu gestalten. Auf Basis der kontinuierlich optimierten Software und des steigenden Absatzes der Sprachassistenten geht OC&C Strategy Consultants davon aus, dass sich der durch transaktionale Sprachsuchen generierte Umsatz in den Vereinigten Staaten und dem UK bis zum Jahr 2022 um mehr als das 20-fache steigern wird (Abb. 4).

Gerade die lokalen Sprachsuchen werden für die Marketer sehr spannend sein, da diese nahe an der Konvertierung liegen. Anfragen wie: „Ok Google, wo ist das nächstgelegene italienische Restaurant?“, sind keine Seltenheit (Abb. 5). Die Unternehmen müssen sich aber, wie bereits oben erwähnt, auf die verschiedenen virtuellen Assistenten einstellen. Während Googles Assistant bei lokalen Sprachsuchen vor allem auf die eigene Plattform Google My Business zurückgreift, nutzt Alexa vermehrt

die Bewertungen auf Portalen wie TripAdvisor oder Yelp. Umso wichtiger ist es daher, die Unternehmensadressen, Öffnungszeiten und Kontaktdaten auf den großen Portalen aktuell zu halten.

Ist der Hype schon da oder nicht?

„Utterances are the new hashtags“ (<http://einfach.st/slidesh4>). Aber ist das wirklich wahr? Sind „Äußerungen die neuen Hashtags“ oder ist Albert Creixell, Business-Development-Manager von Amazon Alexa, da ein wenig über das Ziel hinausgeschossen? Viele Marketer bezeichnen den Trend hin zur Sprachsuche als Hype und tun das Thema als überschätzt ab. Allerdings ist diese Ansicht nicht ganz richtig beziehungsweise verfrüht. Noch kann niemand das gesamte Potenzial der neuen Technologie abschätzen. Mit der Zeit werden aber immer mehr SEO-Experten neue Möglichkeiten ausprobieren und entdecken, um die Sprachsuche zum eigenen Vorteil auszunutzen. Sobald dies der Fall ist, geht der „Hype“ erst richtig los. Wer dann aber erst anfängt, sich mit dem Thema zu beschäftigen, kommt deutlich zu spät, wie uns ein anderes Beispiel zeigt.

AMP und das Voice Search Markup Speakable im Vergleich

Als Google, WordPress und Adobe sich 2015 zusammenschlossen, um das AMP-Projekt (Accelerated Mobile Pages) schnellstmöglich umzusetzen, reagierten die meisten Publisher zunächst eher zögerlich. Schon vor vier Jahren war klar ersichtlich, dass sich Suchanfragen via Smartphone immer größerer Beliebtheit erfreuen und erfreuen werden. Um der steigenden Nachfrage nach mobilen Webseiten nachzukommen, entwickelten sie ein System, das die Ladezeit der Webpages möglichst geringhalten sollte, indem die Content-Formate komprimiert

werden. Zunächst war das Projekt, genau wie nun Googles Voice Search Markup Speakable, in einer Testversion für bestimmte News-Seiten verfügbar. Das gesamte Projekt ist darauf ausgelegt, die Webseiten so schnell wie möglich zu laden. Damit bringt das System, sowohl für User, Publisher als auch Werbetreibende, viele Vorteile mit sich und wurde dementsprechend schnell ausgeweitet. Seit 2017 ist eine AMP-Optimierung Standard für News-Sites und stand schon kurz nach der Testphase mit den Publishern allen Websites offen. Ähnlich könnte es auch mit der Voice Search Optimization laufen.

Derzeit bietet Google mit der BETA-Version des Speakable-Markups englischsprachigen Unternehmen verschiedene Leitlinien für die Voice-Search-Optimierung der eigenen Unternehmenswebseite an. Derzeit handelt es sich noch um eine Testversion. Genauere Angaben, wann das Markup auf weitere Sprachen ausgeweitet wird, hat Google noch nicht getätigt. Hier ist also noch mit einigen Entwicklungen zu rechnen.

Derzeit kooperiert Google vor allem mit mehreren englischsprachigen News-Websites, die das Markup testen sollen. Der britische Guardian hat eigens für die Voice-Search-Optimierung ein Voice-Lab ins Leben gerufen, mit dem sich das altherwürdige Unternehmen als Thought Leader im Bereich Voice Search positionieren will. Unternehmen und Marken scheinen also durchaus die Potenziale der Voice Search zu erkennen. Speakable ist allerdings nicht nur für News-Seiten öffentlich zugänglich, sondern kann bereits von allen englischsprachigen Webpages genutzt werden.

Genau hierin liegt die Chance für die Unternehmen: Schon heute müssen sie sich mit der Optimierung der eigenen Inhalte für die Sprachsuche beschäftigen. Nur so ist es möglich,

Required properties	
cssSelector	<p>Text</p> <p>Addresses content in the annotated pages (such as class attribute). Use either <code>cssSelector</code> or <code>xPath</code>; don't use both. For example:</p> <pre>["headline", "summary"]</pre>
xPath	<p>Text</p> <p>Addresses content using xPaths (assuming an XML view of the content). Use either <code>cssSelector</code> or <code>xPath</code>; don't use both. For example:</p> <pre>/html/head/title</pre>

Abb. 6: Verfügbare Content Locators von Google zur auditiven Wiedergabe von Inhalten (Quelle: Google; <http://einfach.st/gdev43>)

sich einen strategischen Vorteil gegenüber den Wettbewerbern zu verschaffen, sobald die Sprachsuche aus den Kinderschuhen tritt.

Wie optimiert man für Position Zero?

Die Sprachsuche unterscheidet sich in einem Punkt klar von der traditionellen Suchanfrage via Text: Es werden nicht mehr die klassischen zehn Ergebnisse einer Google-Suchanfrage angezeigt, sondern nur noch ein einziges Ergebnis wird wiedergegeben. Sofern die Suche mit einem Display genutzt wird, wird das Ergebnis auch als weiterführender Link auf dem Display angezeigt. Es geht nur noch um die passendste, beste Antwort – die Position Zero.



Abb. 7: Gründe zur Nutzung von Sprachsuchen (Quelle: Mindshareworld.com; <http://einfach.st/mindsh3>)

Weingut Roth 2018
 Silvaner trocken, BIO-Wein,
 Alkohol 12,5 Vol.%,
 Restzucker 3,8 g/L,
 Säure 6,5 g/L

ONLINE SCHOPPEN



Den exklusiven **SEOVANER** und andere Weine gibt es online nur unter onlineschoppen.de

Doch welche Faktoren spielen hierfür eine Rolle?

Selbstverständlich sind die Kriterien ein gut gehütetes Geheimnis. Google gibt allerdings ein paar der entscheidenden Punkte preis. So unterscheidet der Algorithmus zwischen technischen und inhaltlichen Leitlinien, Website-Erstellung und strukturellen Daten. Es wird empfohlen, nur die allerwichtigsten Bullet Points als speakable zu markieren und diese mit dem Markup zu kennzeichnen, sodass Google schnell erkennt, welche Fakten relevant sind. Hierbei sollte jeder Publisher darauf achten, diejenigen Stellen zu markieren, die besonders für die Text-to-Speech-Ausgabe geeignet sind. Noch mehr an Relevanz haben präzise Überschriften gewonnen, die dem Leser eine exakte Beschreibung dessen geben, was sie in dem folgenden Absatz zu erwarten haben. Google weist außerdem darauf hin, dass die Antworten auf die Sprachsuchen möglichst kurz und prägnant zu halten sind. So sollten die Sektionen nicht länger als zwei bis drei Sätze bzw. 20-30 Sekunden lang sein. Zur Kennzeichnung der Inhalte empfiehlt Google die Nutzung einer der beiden Möglichkeiten aus Abbildung 6.

Verschiedene Studien haben darüber hinaus ergeben, dass die Webpages, die als Suchergebnisse wiedergegeben werden, deutlich schneller laden als die durchschnittliche Webseite. Außerdem ist auffällig, dass Featured Snippets aus der organischen Suche in vielen Fällen in der Sprachsuche wiedergegeben werden.

Inzwischen bezieht Google auch PDF-Dokumente und Tweets in die Featured Snippets ein. Somit können selbst Whitepaper die Position Zero einnehmen und den Nutzer auf die eigenen Inhalte aufmerksam machen. Es ist sehr wahrscheinlich, dass sich vereinzelte Parameter noch verändern werden oder sogar ganz neu hinzukommen. Die bekannten Faktoren zeigen dennoch

schon einmal, in welche Richtung es geht und worauf der Google Assistant Wert legt.

Was genau ist das Speakable Markup?

Man kann sich das Speakable Markup ein wenig so vorstellen wie die Meta Description für den Google Assistant. Unterschiedlich ist jedoch die empfohlene Länge. Während bei einer Meta Description höchstens 160 Zeichen (entspricht ungefähr einem längeren Satz) vollständig dargestellt werden, können die Audio-Sektionen zwei bis drei Sätze lang sein. Somit empfiehlt es sich, einen gesonderten Absatz für die Sprachausgabe zu erstellen und diesen mit dem `cssSelector` oder `xPath` zu markieren. Nur so wird die Chance maximiert, als Antwort auf eine Sprachsuche angegeben zu werden. Dies bedeutet zwar zunächst mehr Arbeit für den Marketer, allerdings zahlt sich diese definitiv aus.

Die Fallback-Lösung wäre die kurzfristige Nutzung der Meta Description, um zunächst wenigstens etwas für die Sprachausgabe bereitstellen zu können. Allerdings besteht so immer die Gefahr, dass die Wettbewerber sich vor dem eigenen Unternehmer platzieren und man so gar nicht mehr gefunden wird.

Ist die Position Null immer gut fürs Geschäft?

Das eigene Unternehmen ist auf der Position Zero – was jetzt? Die Featured Snippets haben auch eine Schattenseite. Aufgrund der genauen Antworten zu den Fragen ist die Sprachsuche nach der Wiedergabe des Featured Snippets meist abgeschlossen. Ein großes Problem für Unternehmen, da der Touchpoint mit dem Kunden zumeist abgeschirmt geschieht. Dennoch ist es wichtig, sich mit den nächsten Schritten zu befassen.

Expectation-Management ist hier das Stichwort. Wie kann ich dem potenziellen Kunden eine möglichst

angenehme und reibungsfreie UX garantieren?

Wie in Abbildung 7 zu sehen ist, wird die Sprachsuche vor allem aufgrund ihres Komforts genutzt (Speak Easy Reports 2017). Sie ist bequem, einfach zu nutzen und ermöglicht das gleichzeitige Multitasking. Der Mehrwert liegt dementsprechend in der Audio-Wiedergabe der Antworten. Um die UX bestmöglich zu gestalten, können Unternehmen die Bedienung der gesamten Webseite oder zumindest eines Teils via Sprachsteuerung anbieten. Mögliche Software-Lösungen stehen hierfür auch schon bereit. Unter anderem können Unternehmen mithilfe der Web Speech API Audio-Antworten einbinden. Dass die Nutzung solcher Tools auch für kleine und mittelständische Unternehmen möglich ist, zeigt Hertie. Das Online-Kaufhaus nutzt bereits das `speechRecognition`-Tag in seinem Seitenquelltext. Durch solche Einbindungen können vor allem Online-Shops von der neuen Technologie der Sprachsuche profitieren und sich früh als Thought Leader positionieren. Gleichzeitig steigt die Wahrscheinlichkeit der Konvertierung aufgrund der besseren User Experience.

Fazit

Voice Search bietet vielseitige Nutzungsmöglichkeiten. Aufgrund des hohen Potenzials, der stetig steigenden Akzeptanz der Nutzer, gepaart mit dem technologischen Fortschritt und der immer weiter optimierten Usability der virtuellen Assistenten, ist es nur eine Frage der Zeit, bis die Voice Search Optimization ihren Kinderschuhen entwächst. Wann genau dieser Zeitpunkt kommt, ist nur sehr schwer abzuschätzen. Sicher ist aber, dass schon jetzt die Weichen für die Voice Search Optimization gestellt werden. Nur wer sich heute auf die kommenden SEO-Herausforderungen vorbereitet, wird am Ende ganz oben ranken und den strategischen First Mover Advantage besitzen. ¶