

Digitale Sprachassistenten: Wie Alexa und Co Sprache und Gesellschaft verändern

von **Vera Linss** am **12. April 2018**

2017 war das Jahr von Alexa, Okay Google, Siri oder Cortana. Digitale Sprachassistenten hat der Bundesverband Digitale Wirtschaft sogar als Megatrend des Jahres bezeichnet. Demnach hatten drei Viertel aller Bundesbürger bereits Kontakt mit den sprechenden Helfern oder können sich vorstellen, diese anzuschaffen. Nachdem blmplus schon mehrfach über die Bedeutung von Alexa & Co für den Medienmarkt berichtet hat, soll nun ein Blick auf mögliche gesellschaftliche Auswirkungen geworfen werden.

700 Millionen Menschen nutzen weltweit digitale Sprachassistenten

Weltweit nutzen über 700 Millionen Menschen einen digitalen Sprachassistenten. Die Nutzerzahl soll sich bis 2020 mehr als verdoppeln. Das wirft kritische Fragen auf. Etwa dazu, wie die Technik unsere Sprache und unser Zusammenleben prägt.

Denn was so praktisch daherkommt, hat auch seine Tücken. Das Versprechen jedenfalls, dass die Geräte immer zuhören und uns jeden Wunsch von den Lippen ablesen, löst sich in der Realität selten ein.

Im Gegenteil: Verständigungsprobleme und ein rauer Befehlston bestimmen den Umgang mit den neuen technischen Mitbewohnern. Für Holger Schulze, der an der Universität Kopenhagen das Sound Studies Lab leitet, ist dies ein Zeichen dafür, dass trotz ihrer massenhaften Verbreitung der Umgang mit digitalen Sprachassistenten noch keine Normalität ist.

„Wir befinden uns kulturell immer noch im Prozess der Annäherung“, urteilt der Musikwissenschaftler mit Blick sowohl auf die Technik als auch auf die Nutzer.

Alexa verlangt einen rauen Befehlston und antwortet mit synthetisierter Stimme

Ein wichtiges Merkmal der digitalen Sprachassistenten etwa ist ihre synthetisierte Stimme. Auch wenn viele mit den Geräten interagieren, als handele es sich dabei um einen Menschen: Noch hört man deutlich, dass man in Wirklichkeit mit einer Maschine spricht, die zudem noch nicht mal so ausgereift ist, dass sie Mehrdeutigkeiten oder Nuscheleien verstehen kann.

Dass man die Maschine an ihrem Klang immer noch als solche erkennt, könnte eigentlich von Vorteil sein. Mensch und Technik lassen sich so leichter voneinander abgrenzen. „Historisch betrachtet gibt es aber das Bedürfnis, genau das zu nivellieren“, sagt Holger Schulze. Die erste Sprechmaschine, die Euphonia, wurde bereits 1840 vorgestellt. Seitdem ging es darum, einen Klang zu schaffen, der immer echter wirkt.

Wie sehr dabei gesellschaftliche Normen, etwa das Geschlechterbild, durch Technik manifestiert werden, hat 150 Jahre später die Einführung von Auto-Navigationssystemen gezeigt. In der Testphase setzte man für die Sprachausgabe weibliche Stimmen ein – zum Missfallen der Fahrzeughalter. „Sie können sich gut den Protest der patriarchalen Autofahrer vorstellen: ´Ich lass mir doch nicht von einer Frau neben mir sagen, wie ich fahren soll! ´“, erklärt Musikwissenschaftler Schulze die damalige Ablehnung.

Deshalb habe man daraufhin zunächst auf männliche Stimmen zurückgegriffen. „Durch die Rolle der servilen Computerassistentin in Science-Fiction-Filmen hat es dann eine Art kulturelle Aneignung der weiblichen Computerstimme durch die Medien gegeben. Dadurch haben Sprachassistenten heute diesen Sekretärinnen-Charakter, der auch gerade für Patriarchen annehmbar ist.“

Nuscheln verboten: Digitale Sprachassistenten verstehen nur explizites Sprechen

Nicht nur das ist inzwischen Stoff für gesellschaftliche Debatten. Alexa, Siri und Co. können auch noch auf andere Weise beeinflussen, wie Menschen miteinander kommunizieren. Damit die Maschine alles versteht, spricht man sie in der Regel mit lauten, klaren Sätzen und meist im Imperativ an. Das könnte unerwünschte Folgen haben, befürchtet Holger Schulze von der Universität Kopenhagen.

Der Musikwissenschaftler beobachtet die gesellschaftliche Tendenz des vollständigen und permanenten expliziten Sprechens, da man der Maschine ja alles konkret nennen muss. „Uneigentliches Sprechen ist gar nicht denkbar. Ich muss einen Habitus einnehmen, der mich zu einer Stimmdeformation zwingt. Gefordert ist die permanente Professionalisierung des Sprechens“, sagt Schulze.

Und mit Blick auf den ganz normalen Alltag ergänzt er: „Das familiäre intime Vorsichhinnuscheln, das auch wichtig ist für die Kommunikation, wird dabei verdrängt.“ Zumindest beim derzeitigen Entwicklungsstand der Assistenten. Dass sich dies einmal ändern wird, traut Schulze der Technik durchaus zu.

In eigener Sache:

Der Technik der digitalen Sprachassistenten trauen – mit Blick auf den Markt – viele Experten ein Riesenpotenzial zu. Deshalb wird sich blmplus künftig immer mal wieder mit verschiedenen Themenaspekten der neuen Technik beschäftigen, wie zum Beispiel mit konkreten Auswirkungen von Alexa & Co auf den Radiomarkt.