

Voice-Anwendungen eignen sich vor allem für Call-Center – Entwickler achten zu wenig auf die Usability

# Die Benutzer mögen Sprachapplikationen nicht

Wenn Anwender die Wahl haben, meiden sie Sprachdienste. Das haben Fraunhofer-Forscher bei einer Studie zur Benutzbarkeit von Voice-Applikationen herausgefunden. Einer der Gründe: Viele Anwendungen sind technikgetrieben und nicht benutzerorientiert gestaltet. Softwareprogrammierer haben es meist verstanden, Entwickler von Sprachanwendungen müssen es noch lernen: Die Benutzerfreundlichkeit von Anwendungen entscheidet, ob diese Erfolg haben. Das wird aber momentan kaum beachtet „Usability-Experten spielen bei Erstellung dieser Applikationen keine Rolle, in der Regel werden diese von Linguisten oder Programmierern gebaut,“ konstatiert Matthias Peissner vom Fraunhofer-Institut IAO. Er hat

die erste Studie des IAO erstellt, die die Benutzerfreundlichkeit solcher Anwendungen untersucht hat. Dass Usability-Experten nicht zum Zuge kommen, hat Konsequenzen, die sich in den Ergebnissen der jetzt veröffentlichten Untersuchung niederschlagen. Wenn Benutzer die Wahl haben, entscheiden sie sich eher für Mail oder den persönlichen Kontakt als für eine sprachgestützte Applikation. Und das trotz Fortschritten bei der Technik. Drei zentrale Fehler sieht Peissner momentan beim Design:

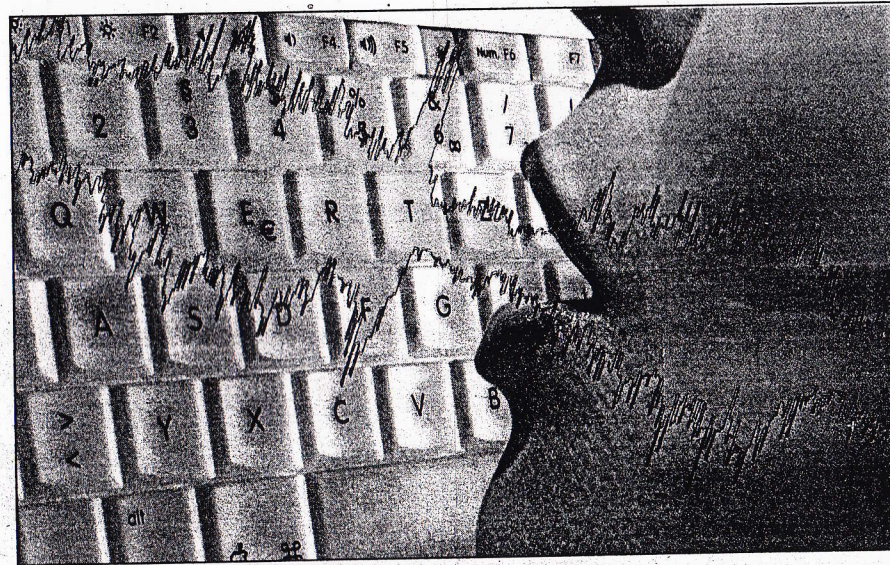
- Die Systemausgaben sind zu lang. Das führt dazu, dass die Benutzer sich die Anweisungen nicht merken können und frustriert werden.
- Die meisten Systeme sind streng menübasiert. Erfahrene

Benutzer verlieren Zeit, weil sie sich durch die gesamte Menüstruktur hangeln müssen und nicht direkt zu einem Endpunkt gelangen.

● Bei der Gestaltung der Schnittstellen werden Erkennungsfehler der Software nicht genügend berücksichtigt. Diese Fehler entstehen, wenn das Programm eine Ansage des Anwenders falsch versteht. Das führt zu Irrwegen und möglicherweise zum Abbruch der Transaktion, ohne dass der Benutzer nachvollziehen kann, warum dies der Fall ist.

## Standards erlauben modularen Aufbau

Der Experte empfiehlt Vorsicht bei der Entwicklung von sprachbasierten Applikationen, auch wenn es bei der Technik deutliche Fortschritte gab: So ist zum einen die Spracherkennung selbst erheblich besser geworden. „Die meisten Leute haben aber in der Vergangenheit schlechte Erfahrungen gemacht und sind deshalb skeptisch“, so seine Erkenntnis. Erleichterung bei der Gestaltung bringt der Standard Voice-XML (siehe Kasten), der einen modularen Aufbau der Anwendungen aus vorgefertigten Bausteinen erlaubt. „Der Trend geht momentan dahin, dass Anbieter Standarddialogelemente wie die Abfrage von Pin-Nummern entwickeln und Content-Provider die Applikationen einfach zusammenklicken können“, so Peissner. Der Aufwand für die Anpassung sei trotzdem nicht zu unterschätzen. Der Forscher sieht aber dennoch einen Vorteil: Für Anwender ist es mittlerweile realis-



Die Benutzer geben der Tastatur bei der Eingabe von Anfragen den Vorzug vor Sprachapplikationen. Nur in Call-Centern finden Voice-Anwendungen notgedrungen Akzeptanz. Doch auch dort spielt die Benutzbarkeit oft nur eine untergeordnete Rolle.  
Foto: Photopool/Zillmann

## Zwei Wege führen zum Ziel

Für sprachbasierte Applikationen existieren derzeit zwei Standards, die auf XML basieren: **Voice-XML** ([www.voicexml.org](http://www.voicexml.org)) und **Salt** (Speech Application Language Tags).

- Voice-XML wurde entwickelt, um das Design von Sprachanwendungen zu vereinfachen. Dies wird dadurch erreicht, dass der Sprachdialog und Anwendungslogik getrennt behandelt werden. Die Architektur solcher Anwendungen umfasst in der Regel zwei Stufen: einen **Dokumentenserver** und einen **Interpreter** (Übersetzer). Der Dokumentenserver verarbeitet die Anfragen, die er über den Interpreter erhält. Diesem liefert er die Voice-XML-Daten, welche der Übersetzer in Sprache umwandelt.
- Salt ist dagegen ein Satz von Befehlen, der in HTML-Dokumente eingebaut werden kann. Der Ansatz, der unter anderem von Microsoft unterstützt wird, ist leichtgewichtiger als Voice-XML und dient primär dazu, Webseiten um Sprachfunktionen zu erweitern. Zudem wurde Salt **multimodal** ausgelegt. Das heißt, dass der Anwender mit dem System über Sprache sowie über andere Eingabemöglichkeiten kommunizieren kann – etwa durch Mausclick oder Berühren eines Touchscreens. Die Kommunikation mit Voice-XML-Systemen verläuft dagegen meist per Telefon ([www.microsoft.com/speech/evaluation/speechtags](http://www.microsoft.com/speech/evaluation/speechtags)).  
fg

tischer geworden, ihre sprachbasierten Applikationen vor der Inbetriebnahme zu testen. Der Grund: Änderungen, die sich nach dem Probelauf ergeben, sind durch die Standardisierung weniger aufwändig geworden.

Sinn mache der Einsatz von Voice-Eingabe aber trotz der Hürden dort, wo Sprache echten Mehrwert bietet. Das sei in Call-Centern der Fall. Das könne aber auch bei Anwendungen im Auto sein, bei schlechten Lichtverhältnissen in industriellem Umfeld oder dann, wenn Menschen ihre übrigen Sinne für andere Interaktionsformen benötigen. Im Produktbereich sei dies für Wartungsaufgaben denkbar, in denen auf einem Display Informationen eingeblendet werden.

Innerhalb dieser Informationen könnte ein Benutzer dann per Sprachbefehl navigieren.

Damit dies wirklich Nutzen bringt, müssen die Entwickler aber noch dazulernen. Hier ergeben sich aus den Ergebnissen der Untersuchung konkrete Hinweise:

- Empfohlen wird eine „progressive Assistenz“. Wer nicht weiterkommt, erhält immer detailliertere zusätzliche Hinweise. Das kann eskalieren bis zu dem Vorschlag, eine Verbindung zu einem menschlichen Berater herzustellen.
- Erwiderungen des Systems müssen freundlich und persönlich sein. Anweisungen wie „Sprechen Sie bitte deutlich“ verunsichern und haben den gegenteiligen Effekt.

● Zusammenfassende Bestätigungen am Ende einer Eingabe sind hilfreich.

● Verwendete Begriffe müssen klar voneinander abgegrenzt und für den Benutzer einleuchtend sein.

● Das System sollte nicht mehr als sechs Optionen bei einer Eingabe bieten.

● Eine Mischung aus system- und benutzerinitiierten Programmen ist sinnvoll. Bei systeminitiiert Logik definiert das Produkt die Navigationspfade, bei der zweiten Variante können erfahrene Anwender bekannte Pfade überspringen. pg

Interessenten können die Studie unter [www.fraunhofer-iao.de](http://www.fraunhofer-iao.de) bestellen. Erste Hinweise, ob eine Anwendung sprachgeeignet ist, bietet ein Test unter [www.clarity-ag.net](http://www.clarity-ag.net).