

Sprachapplikationen: Die Kundenperspektive

Matthias Peissner, Jakob Biesterfeldt

Competence Center Human-Computer Interaction
Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO)
Nobelstraße 12
70569 Stuttgart
matthias.peissner@iao.fhg.de
jakob.biesterfeldt@iao.fhg.de

Abstract: Der vorliegende Beitrag fasst die Ergebnisse einer empirischen Studie über die Akzeptanz und Usability (Gebrauchstauglichkeit) von telefonbasierten Sprachapplikationen in Deutschland zusammen. Eine repräsentative telefonische Nutzerbefragung liefert Daten zur Nutzung und Akzeptanz von Sprachapplikationen. Detaillierte Erkenntnisse zur Gebrauchstauglichkeit aktueller Systeme und zu den relevanten Qualitätsaspekten der sprachlichen Benutzungsschnittstelle stammen aus Usability Tests mit sechs ausgewählten deutschen Sprachapplikationen.

1 Einleitung

Seit einigen Jahren sagen Marktforscher dem kommerziellen Einsatz von Sprachtechnologien ein starkes Wachstum voraus (vgl. [KG02], [Da03], [JL03], [FR00], [Ra02]). Deutschland befindet sich dabei auf einer aussichtsreichen Position, großen Anteil an dieser Entwicklung zu haben (vgl. [JL03]). Dennoch werden telefonbasierte Sprachapplikationen im deutschsprachigen Raum weitaus weniger angeboten und genutzt als es die wirtschaftlichen und technologischen Potenziale erwarten lassen. Warum?

Die hier zusammengefasste Studie [PB04] liefert empirische Daten zur aktuellen Nutzung, Qualität und Akzeptanz von telefonbasierten Sprachapplikationen in Deutschland. Dabei umfasst der Begriff der Sprachapplikation Systeme mit akustischer Eingabe und akustischer Ausgabe. Tastengesteuerte Systeme (DTMF) sind ausgeschlossen. Die Studie nimmt stets die Perspektive des potenziellen Benutzers ein.

2 Die Kundenperspektive – Ergebnisse der telefonischen Befragung

Um den Bekanntheitsgrad von Sprachapplikationen in Deutschland zu erheben und um Nutzungsverhalten und typische Einstellungen potenzieller und tatsächlicher Benutzer zu erfassen, wurden im Oktober 2003 computerunterstützte telefonische Interviews (CATI) mit insgesamt 1005 Personen durchgeführt. Die Stichprobe ist repräsentativ für die deut-

sche Gesamtbevölkerung bezüglich Alter, Geschlecht und Bundesland. Die CATI wurde vom Fraunhofer Institut für Arbeitswirtschaft und Organisation (IAO) initiiert, konzipiert und ausgewertet. Die Durchführung übernahm ein Marktforschungsinstitut.

2.1 Bekanntheit und Nutzung von Sprachapplikationen

Knapp die Hälfte aller Befragten (46%) gibt an, Sprachapplikationen nicht zu kennen. Nur 32% können mindestens eine Applikation namentlich nennen. Diejenigen, die Sprachapplikationen nutzen, tun dies eher sporadisch. Die meisten (40%) geben an, im letzten halben Jahr weniger als fünfmal eine Sprachapplikation genutzt zu haben. Die meisten Befragten, die angeben, bereits Sprachapplikationen genutzt zu haben, sind zwischen 25 und 45 Jahre alt (53,5% der Nutzer). Die Gruppe der Nutzer unterscheidet sich von der bundesdeutschen Gesamtbevölkerung außerdem durch bessere Computerkenntnisse und durch einen höheren Anteil an Handybesitzern.

2.2 Zufriedenheit der Nutzer

Mehr als jeder Vierte (76 bzw. 26%) der 289 Nutzer gibt an, „sehr unzufrieden“ mit der von ihm am häufigsten genutzten Sprachapplikation zu sein. Nur gut jeder Dritte (102 bzw. 35%) ist „sehr zufrieden“ oder „eher zufrieden“.

2.3 Einstellungen und Erwartungen gegenüber Sprachapplikationen

Die detaillierte Betrachtung kritischer Einstellungen und Erwartungen tatsächlicher und potenzieller Nutzer kann als Ausgangspunkt genutzt werden, um mangelnde Nutzungszufriedenheit und Nutzungsbereitschaft besser zu verstehen. Einerseits können erste Hinweise abgeleitet werden, in welchen Bereichen der Sprachapplikationen besonderer Verbesserungsbedarf besteht. Andererseits wird klar, welche Vorurteile in einer positiven Marktkommunikation und Öffentlichkeitsarbeit besonders berücksichtigt werden sollten. Zwei der abgefragten Eigenschaften werden besonders schlecht bewertet: Die Befragten sind der Meinung, heutige Sprachapplikationen könnten eher nicht auf individuelle Bedürfnisse eingehen. Zweitens wird der Spaßfaktor als mangelhaft eingestuft, insbesondere von denjenigen, die bereits erste Nutzungserfahrungen sammeln konnten.

2.4 Einstellungen gegenüber Unternehmen, die Sprachapplikationen verwenden

Einige Unternehmen fürchten einen Imageverlust durch den Einsatz von Sprachapplikationen. Eine Technologie, die nicht perfekt funktioniert, könnte dem Kundenvertrauen schaden und damit nachhaltige Risiken für Kundenbindung und Marktstärke mit sich bringen. Tatsächlich stellt sich dieser Zusammenhang bei weitem nicht derart eindeutig dar. Zwar scheint der Einsatz von Sprachapplikationen auch mit negativen Folgen für das Firmenimage verbunden zu sein (weniger „einladend“ und „persönlich“). Doch zeigt sich insbesondere, dass Unternehmen, die Sprachapplikationen einsetzen, als „innovativ“ und „professionell“ wahrgenommen werden.

3 Aktuelle Sprachapplikationen – Qualität und Erfolgsfaktoren

Um einen repräsentativen und zugleich detaillierten Überblick über die Qualität deutscher Sprachapplikationen zu erlangen, haben wir im Oktober 2003 sechs ausgewählte Applikationen in Usability Tests mit insgesamt 30 potenziellen Benutzern analysiert:

- **Berti:** Informationen zur Ersten Fußball-Bundesliga. 0 91 31/61 00 17
- **EVA:** Informationen wie News, Wetter, Verkehr, etc. und Service-Angebote wie Mailbox, Kundenverwaltung und Adressbuch. 11 44, nur für e-plus Kunden
- **Fahrplanauskunft der Bahn:** Abfrage von Zugverbindungen. 08 00/1 50 70 90
- **FAZ Nachrichten:** Nachrichten zu verschiedenen Rubriken, wie Sport, Politik, Wirtschaft, Kultur, Börse, etc. 01 90/15 14 44
- **Postbank Telefonbanking:** Einfache Transaktionen und allgemeine Information zu Wertpapieren oder Postbank-Produkten angeboten. 0180 / 3 04 07 00
- **Quellefon:** Verschiedene Dienste des Quelle Versandhauses: Bestellung von Katalogartikeln 0180 / 5 33 66 22 , Bestellung von Katalogen 0180 / 3 23 25 24 und Informationen zum persönlichen Kundenkonto 0180 / 3 68 78 88.

3.1 Allgemeines Qualitätsniveau und negative Vorbehalte

Insgesamt scheinen die getesteten Sprachapplikationen besser zu sein als die allgemein verbreiteten Vorurteile und Vorbehalte. Die meisten Probanden haben nach den Tests ein positiveres Bild von Sprachapplikationen als vorher. Die Bewertung wesentlicher Eigenschaften, wie Übersichtlichkeit, Passung auf individuelle Bedürfnisse, bequeme Handhabung, etc. fällt in den meisten Fällen nach der Benutzung positiver aus als vorher.

3.2 Qualitätskriterien

Die wichtigsten Problemfelder und Erfolgsfaktoren, die für die Schaffung von Akzeptanz relevant sind, können wie folgt beschrieben werden:

- **Angebot an Inhalten und Funktionalitäten:** Ein Kosten-Nutzen-Verhältnis, das dem Kunden vorteilhaft erscheint, ist grundlegende Motivation für die Nutzung einer Sprachapplikation. Um gegen alternative Angebote bestehen zu können, ist es entscheidend, Potenziale und inhärente Probleme der Sprachinteraktion zu erkennen (vgl. [PH03]) und geeignete Nutzergruppen und Nutzungssituationen zu identifizieren. Außerdem muss auf Grundlage der Ergebnisse dieser Studie der grundsätzliche Ansatz eines Sprachportals in Frage gestellt werden. Die Benutzer bevorzugen Applikationen mit einem klaren und eng begrenzten Funktionsbereich. Komplexe Applikationen mit heterogenen Inhalten, wie z.B. EVA oder FAZ erschweren die Orientierung und führen schnell zu kognitiver Überbelastung der Benutzer.
- **Automatische Spracherkennung:** Angesichts aktueller Möglichkeiten sind die Spracherkennungsquoten der getesteten Applikationen zum Teil sehr enttäuschend. Nur zwei der sechs Applikationen erreichen insgesamt eine Erkennungsquote von über 80%. Zwar ist eine zuverlässig funktionierende Spracherkennung ein wichtiges Merkmal einer guten Sprachapplikation. Ihre Bedeutung für die Akzeptanz und Be-

nutzungsfreundlichkeit wird jedoch meistens stark überschätzt (vgl. [Pe02]). Der Einfluss der Spracherkennungsqualität auf die subjektiv empfundene Bedienqualität ist bei den verschiedenen Testapplikationen unterschiedlich ausgeprägt. Während bei Bertl der Zusammenhang zwischen subjektiver Anstrengung und Spracherkennungsperformanz recht stark ist (Bestimmtheitsmaß $R^2=0.68$), besteht z.B. bei Qel-telefon und Postbank überhaupt kein Zusammenhang in dem Sinn, dass die Benutzer bei einer besseren Spracherkennung ein geringeres Maß an Anstrengung berichteten.

- **Management von Erkennungsfehlern:** Eine der wichtigsten Aufgaben des Voice User Interface Designs ist es, die Schwächen der automatischen Spracherkennung (ASR) zu kompensieren und die negativen Folgen von Spracherkennungsfehlern zu minimieren. Zunächst muss sichergestellt werden, dass der Benutzer das Auftreten eines Erkennungsfehlers bemerkt. Die Korrektur von Fehlern muss schnell und ohne „Umwege“ möglich sein. Außerdem ist es wichtig, dass Erkennungsfehler keine schwerwiegenden negativen Konsequenzen, wie etwa Datenverlust nach sich ziehen. Schließlich kann gutes Prompting beim Anrufer Akzeptanz für das Auftreten von Erkennungsfehlern schaffen und so ein kooperatives Nutzerverhalten fördern.
- **Unterstützung bei der Formulierung zielführender Benutzeräußerungen:** Bei den meisten Sprachapplikationen ist der Kunde auf eine recht kleine Menge zulässiger Benutzeräußerungen festgelegt. Um den Benutzer dennoch sicher und effizient an sein Ziel zu führen, müssen die Prompts so gestaltet sein, dass dem Anrufer der Funktionsumfang klar wird, dass er weiß, was er sagen kann und wie er sprechen muss. Dabei sind explizite Erläuterungen oft wesentlich langwieriger und weniger effektiv als eine implizite Beeinflussung der Sprachperformanz des Benutzers durch entsprechende Vorbildfunktion der akustischen Systemausgaben. Auf Seiten der Grammar ist es wichtig, dass die hinterlegten Benutzeräußerungen möglichst intuitiv sind, dem alltäglichen Sprachgebrauch der Zielgruppen entsprechen und leicht erlernbar, leicht erinnerbar und leicht sprechbar sind. Zugleich ist darauf zu achten, dass das Grammar – Design die Schwächen und Stärken der verwendeten ASR berücksichtigt und damit eine optimale Erkennung begünstigt.
- **Navigationsmöglichkeiten und Dialogabläufe:** Mixed Initiative Dialoge zeichnen sich dadurch aus, dass geübte Benutzer bekannte Funktionen direkt aufrufen können, ohne lange Menüs durchlaufen zu müssen. Wenn der Benutzer nicht alle benötigten Informationen angibt, die Interaktion ins Stocken gerät (wiederholte Zurückweisung, kein Benutzerinput) oder wenn Hilfe aufgerufen wird, übernimmt das System die Initiative, präsentiert mögliche Benutzeräußerungen und führt so den Benutzer durch den Dialog. Nur ein derart flexibles Dialogdesign kann den Bedürfnissen heterogener Benutzergruppen gerecht werden, und damit auch langfristig Akzeptanz und Kundenbindung schaffen. Außerdem sollte sich der Benutzer stets in der Lage sehen, den Dialog nach seinen Vorstellungen und Bedürfnissen zu kontrollieren und zu steuern, z.B. durch die Möglichkeit, mit sprachlicher Eingabe den aktuellen Prompt zu unterbrechen (Barge In). Schließlich sollten die wichtigsten und häufigsten Nutzungsszenarien besonders schnell und komfortabel bedienbar sein.
- **Akustische Informationsdarstellung:** Der wichtigste Grundsatz für die akustische Ausgabe gesprochener Informationen ist: „Fasse dich kurz!“. Gesprochene Texte müssen deutlich einfacher konstruiert sein als Schrifttexte: Die Sätze müssen kurz sein, die Satzkonstruktionen übersichtlich. Die Bezüge zwischen den Sätzen müssen klar sein. Außerdem müssen die verwendeten Begriffe, die Bezeichnungen spezifi-

scher Funktionen und deren Oberbegriffe, klar und verständlich sein und direkt der Erfahrungswelt des Benutzers entsprechen. Auch nonverbale Gestaltungsaspekte sind von Bedeutung: die akustische Qualität beeinflusst die Verständlichkeit, Earcons und Hintergrundgeräusche können zur besseren Orientierung beitragen.

- **Anmutung und emotionale Wirkung:** Das „Hear&Feel“ einer Sprachapplikation ist nicht nur ein entscheidender emotionaler Faktor, sondern beeinflusst auch das Dialogverhalten des Benutzers und darüber schließlich Erfolg und Effizienz der Interaktion. Wichtigster Grundsatz dabei ist, dass insgesamt ein in sich stimmiges System entsteht: Nutzergruppen, Anwendungsdomäne, Serviceprovider, Persönlichkeit, Stimme, Dialogansatz und Spracherkennungsqualität müssen zusammen passen.

Zusammenfassung und Ausblick

Die dargestellte Studie veranschaulicht den aktuellen Stand von Sprachapplikationen in Deutschland. Bekanntheitsgrad, Nutzungshäufigkeit und Nutzungszufriedenheit sind noch relativ gering. Neben umfassender Öffentlichkeitsarbeit und Marketingkampagnen werden insbesondere gute Beispielapplikationen benötigt. Zwar scheinen die aktuellen Sprachapplikationen etwas besser als ihr Ruf, den hohen Anforderungen einer benutzer-gerechten Dialogführung entsprechen sie jedoch kaum. Die dargestellten Qualitätskriterien können als Ausgangspunkt einer ergonomischen Gestaltung dienen. Sie ersetzen jedoch nicht den gezielten Einsatz von Methoden des Usability Engineering, um in einem iterativen und benutzerzentrierten Entwicklungsprozess zu einer hochwertigen Sprachapplikation zu gelangen, die breite Zustimmung und Akzeptanz finden wird.

Literaturverzeichnis

- [Da03] Datamonitor: Voice Automation: Past Present and Future. In: www.speechtechmag.com/whitepapers/Intervoice.pdf, 2003, zugegriffen am 30.04.2004
- [FR00] Forrester Research: The Web's Speech Impediments, In: <http://www.forrester.com/ER/-Research/Report/Summary/0,1338,9209,FF.html>, zugegriffen am 30.04.2004
- [JL03] Joscelyne, A.; Lockwood, R.: Benchmarking HLT progress in Europe. Kopenhagen 2003
- [KG02] The Kelsey Group: Speech Technology revenues to reach \$2 billion by 2006 according to the Kelsey Group, In: <http://www.kelseygroup.com/press/pr020430.htm>, 2002; zugegriffen am 30.04.2004
- [Pe02] Peissner, M.: What the Relationship between Correct Recognition Rates and Usability Measures Can Tell Us About the Quality of a Speech Application. In: Proceedings of the 6th International Scientific Conference on Work With Display Units, Berlin 2002
- [PB04] Peissner, M.; Biesterfeldt, J.; Heidmann, F.: Akzeptanz und Usability von Sprachapplikationen in Deutschland. Stuttgart 2004
- [PH03] Peissner, M.; Heidmann, F. und Corves, L.: Usability Engineering von Sprach-Applikationen. In: Machate, J. & Burmester, M. (Hrsg.): User Interface Tuning – Benutzungsschnittstellen menschlich gestalten. Frankfurt a.M. 2003
- [Ra02] Raebel, R.: Sprachtechnologie – Mehr als Spracherkennung. In: <http://www.frankfurt-main.ihk.de/presse/ihk-wirtschaftsforum/2002/0201/sprachtechnologie/index.html>, zugegriffen am 04.10.2003