

Erfolgsfaktoren für den Einsatz von Sprachinteraktion



Institutszentrum Stuttgart der Fraunhofer-Gesellschaft



Fraunhofer Institut
Arbeitswirtschaft und
Organisation

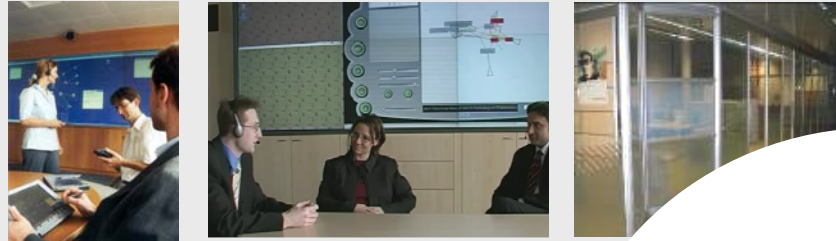
Erfolgsfaktoren für den Einsatz von Sprachinteraktion

Matthias Peissner

Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
Competence Center Human-Computer Interaction



Interaktionslabor / Ambient Intelligence Lab



Innovative Interaktionstechniken
Ambient Intelligence

Usability Labor



Usability Engineering
User Interface Gestaltung

»Mensch-Technik Interaktion« am Fraunhofer IAO



m-Lab



Mobile Software-Lösungen
Mobile ICT-Unterstützung
(<http://www.m-lab.iao.fraunhofer.de>)

Virtual Engineering



Virtual Reality / Augmented Reality
Hardware-Ergonomie und Produktdesign
Vehicle Interaction

Einsatz und Perspektiven der Sprachinteraktion Wofür ist Sprache gut?



»Speech is the bicycle of user interface design:
It is great fun to use and has an important role, but it
can carry only a light load.

Sober advocates know that it will be tough to replace
the automobile: graphic user interfaces.«

Ben Shneiderman, 1998

Shneiderman, B.: Designing the user interface, third edition, Addison-Wesley. 1998.



Potenziale der Sprachinteraktion (1)

Geringe Anforderungen an das Endgerät

- Zur Benutzung eines Sprachservices genügt ein einfaches Telefon (oder Lautsprecher / Mikrofon): klein, preiswert, mobil, nachhaltig.

Immer und Überall

- Mobiler Zugriff auf umfangreiche Inhalte und Funktionalitäten
- Konstant hohe Servicequalität (v.a. Call Center)

Intuitive Bedienung

- Sprache ist die natürlichste Form der menschlichen Kommunikation. Auch mit Maschinen?
- Keine „Übersetzung“ in Mausklicks und Menüs nötig, man kann direkt sagen, was man will.

Praktisch und Effizient

- Direkte Erreichbarkeit aller Informationselemente der Applikation
- Kein mühsames Durchlaufen hierarchischer Menüs und langer Listen



Potenziale der Sprachinteraktion (2)

- Hände frei - Augen frei** – Besonders beim Autofahren, im industriellen Arbeitskontext und bei schlechten Lichtverhältnissen ein wesentlicher Vorteil
- Weniger Aufmerksamkeit** – Visuelle Aufgaben fordern unsere gesamte Aufmerksamkeit. Dialoge führen wir praktisch „nebenbei“.
- Vielfältige Einsatzmöglichkeiten** – Auskunftssysteme, Transaktionssysteme, Sprachsteuerung, Form-Filling, Aufnahme und Verarbeitung von Kommentaren, einfache Klassifikationsaufgaben, Kommissionierung, etc.



Probleme der Sprachinteraktion (1)

**Keine 100%ige
Spracherkennung möglich**

- Natürliche Variabilität der menschlichen Sprache
- Besonders problematisch sind große Vokabulare und die Erkennung komplexer Benutzeräußerungen
- Trotz enormer technologischer Entwicklungssprünge, auch in absehbarer Zukunft keine perfekte Erkennung

**Akustische Darbietung
umfangreicher Informationen**

- ist zeitaufwendig und erfordert volle Aufmerksamkeit.
- Die Reihenfolge der Informationsaufnahme kann vom Benutzer nicht so leicht und komfortabel kontrolliert werden wie bei einem GUI.

**Überblick über
Funktionsangebot**

- Ein grafisches Interface ermöglicht ein vergleichsweise schnelles Erfassen aller angebotenen Inhalte und Funktionen

**Was kann ich sagen?
Wie muss ich sprechen?**

- Die Beschränkungen der automatischen Spracherkennung (ASR) müssen bekannt sein, um die Applikation effektiv nutzen zu können.



Probleme der Sprachinteraktion (2)

Individuelle Sprechgewohnheiten und Dialekte

- Erschweren die zuverlässige Spracherkennung, insbesondere bei Systemen, die von sehr unterschiedlichen Gruppen genutzt werden.

Negative Vorerfahrungen und Vorbehalte

- Häufig negative Erfahrungen mit Diktiersystemen.
- Negative Vorurteile stellen ein wesentliches Akzeptanzhindernis dar und können die Interaktion negativ beeinflussen.

Unrealistische Erwartungen

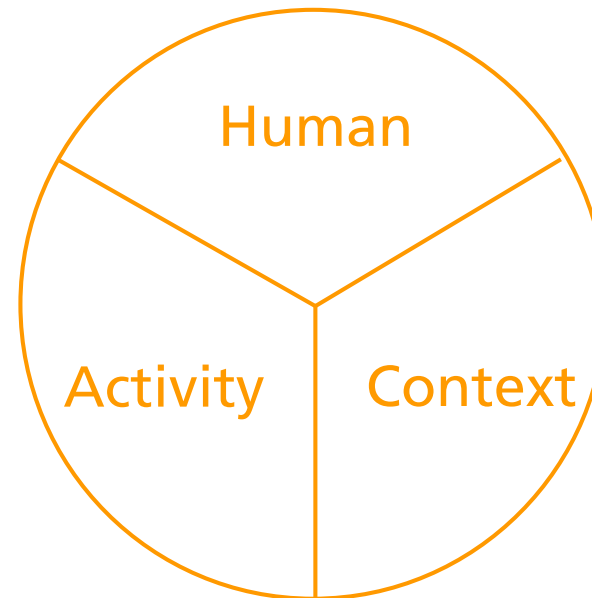
- Viele Benutzer möchten sprechen, wie mit einem menschlichen Partner.
- Systeme, die Sprache „verstehen“ sind noch nicht realistisch. Auch schon mit Sprach-„Erkennung“ ist eine intuitive und effiziente Bedienung in vielen Bereichen möglich.



Kriterien für den industriellen Einsatz von Sprachinteraktion

		sehr gut geeignet	geeignet	möglicherweise geeignet	nicht geeignet
Mitarbeiter	Mitarbeiter mit wenig Computerefahrung		*		
	Mitartermotivation und Offenheit der Mitarbeiter gegenüber technischen Neuerungen			*	
	Bewegungsfreiheit		*		
	Körperliche Beeinträchtigungen, die manuelle Systemeingaben erschweren oder unmöglich machen		*		
	Einschränkungen des Hörvermögens			*	
	Einschränkungen des Sprachvermögens			*	
	Einschränkungen des Sehvermögens		*		
	Schreib-/Leseschwäche		*		
	Fremdsprachige Mitarbeiter		*		
	Mitarbeiter mit verschiedenen Dialekten oder Akzenten		*		
Arbeitsprozess	Freies Sprechen		*		
	Eingabe großer Mengen an Informationen nötig	*			
	Größe des zu erkennenden Vokabulars pro Ebene			*	
	Die benötigten Informationen lassen sich leicht in Worte fassen	*			
	Ausgabe großer Mengen an Informationen nötig				*
	Parallele Informationsausgabe erforderlich			*	
	Parallele Kommunikation mit Mitmenschen und System erforderlich			*	
	Sofortige Systemreaktion oder 100% Verlässlichkeit erforderlich				*
	Benutzererkennung erforderlich		*		
	Eingabe vertraulicher Informationen in Gegenwart anderer nötig				*
	Ist die Ausgabe sicherheitsrelevanter Informationen an einen oder mehrere Mitarbeiter erforderlich	*			
	Konzentration der Augen des Mitarbeiters auf Hauptaufgabe nötig	*			
	Mitarbeiter benötigen Hände für Hauptaufgabe	*			
	Saubere Hände des Mitarbeiters erforderlich	*			
	Arbeitsprozess erfordert Keimfreiheit oder Staubfreiheit	*			
	Mitarbeiter tragen Mundschutz			*	
	Arbeitsprozess erfordert Stille				*
Arbeitsprozess erfordert Mobilität	*				
Arbeitsablauf kommt häufig am Tag vor		*			
Arbeitsprozess wird von eingeschränktem Nutzerkreis ausgeführt		*			
Arbeitsumfeld	Hintergrundgeräusche mit hoher Varianz in der Tonlage vorhanden			*	
	Unvorhersehbare/unbekannte Geräusche vorhanden			*	
	Arbeitsumfeld erschwert visuelle Wahrnehmung	*			
	Temperaturen im Arbeitsumfeld		*		
	Eingeschränktes Platzangebot		*		
Keine Tastatur/Bildschirm verfügbar	*				

Tabelle 3: Relevanz der Kriterien



Human Performance Model nach Bailey, 1996

Mitarbeiter

- Wenig Computerefahrung
- Schreib-/Leseschwäche
- Fremdsprachen/Dialekte

Aktivitäten

- Hände und Augen frei
- Input leicht in Worte zu fassen
- Mobilität
- Freiheit von Keimen/Staub
- Häufig wiederholte Aufgaben

Arbeitsumfeld

- Vis. Wahrnehmung erschwert
- Platzmangel → kein Bildschirm/Tastatur



Sprache im industriellen Einsatz: Feasibility Studien und Guidelines

Multiple User Interfaces für die Halbleiterproduktion



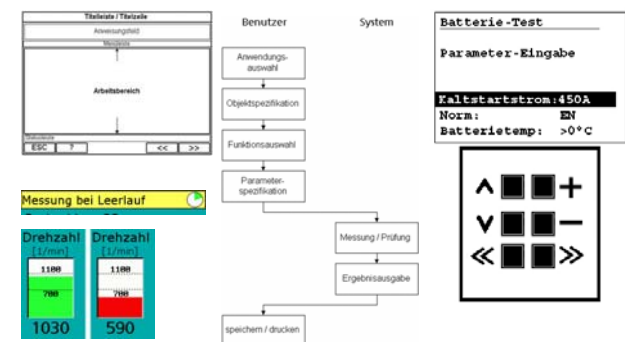
- Für Operator UI:
- Eingabe von Kommentaren bei Ausschuss, Störungen/Fehlern
- Ansonsten:
- Wartung, Service
 - Inbetriebnahme



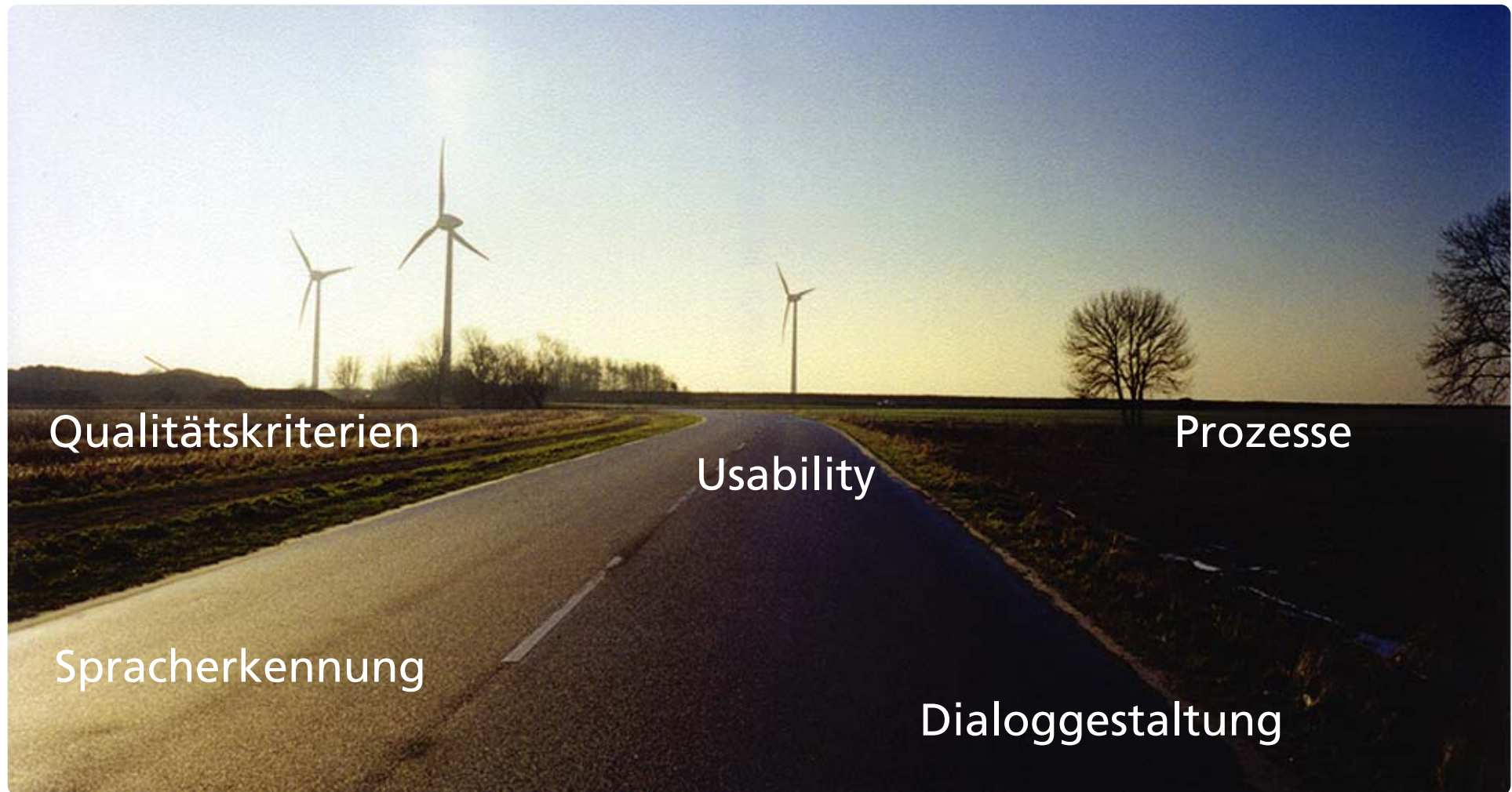
HMI Design für KfZ Prüfgeräte



- Einsatz von Sprachinteraktion in Diagnose- und Prüfaufgaben
- Arbeitsaufgaben und Abläufe
- Datenformate (z.B. Browsing, Form-Filling, Repräsentation, Objekt- und Funktionsauswahl)
- Fehlermanagement
- Etc.



Qualität von Sprachapplikationen



Qualitätsoffensive 2007/08

- Interdisziplinäres Team von mehr als 20 Fachexperten
- Position Papers und relevante Publikationen
- 2-Tagesworkshops
- Arbeitsgruppen zu
 - »Strategie & Business Logik«
 - »Sprachtechnologie«
 - »Dialogplattformen«
 - »VUI Design & Usability«
- Kontinuierliche Abstimmung in Telefonkonferenzen mit Sponsoren
- Vorbereitung, Integration und Dokumentation der Arbeiten durch Fraunhofer IAO



Qualitätsoffensive 2007/08

Leitfaden für gute Sprachapplikationen



- In einem Verbundprojekts unter Leitung von Fraunhofer IAO ist ein **offener Qualitätsstandard für Sprachapplikationen** erarbeitet worden.
- Die entwickelten **Qualitätskriterien** fördern ein ganzheitliches Verständnis der Qualität von Sprachapplikationen.
- Die Übersicht der wichtigsten qualitätsrelevanten **Maße** dient als Grundlage für eine standardisierte Bewertung von Sprachapplikationen auf der Basis konkreter Kenngrößen.
- Die Beschreibung und Bewertung von **Verfahren** zur kontinuierlichen Evaluation und Optimierung unterstützt den Transfer in konkrete Entwicklungsprojekte.
- Die Verständigung auf eine **gemeinsame Terminologie** für zentrale Konzepte (insbesondere **Applikationskomponenten**) erhöht die Transparenz und die Vergleichbarkeit unterschiedlicher Lösungen und Angebote.

Peissner, M.; Hipp, C. & Steimel, B. (2007). Qualitätskriterien, Maße und Verfahren für Sprachapplikationen - Leitfaden

Qualitätskriterien für Sprachapplikationen (1-5)

- | | |
|--|---|
| 1. Angemessener Funktionsumfang und Inhaltsangebot | »Eine Sprachapplikation ist gut, wenn die Sprachapplikation durch ein attraktives und vollständiges Funktionsangebot einen Mehrwert für die Kunden schafft.« |
| 2. Einwandfreie Funktionsfähigkeit und Leistungsfähigkeit | »Eine Sprachapplikation ist gut, wenn ein sicheres, performantes und fehlerfreies Funktionieren des Systems gewährleistet ist – auch bei zu erwartenden Belastungsspitzen.« |
| 3. Administrierbarkeit und effizienter Betrieb | »Eine Sprachapplikation ist gut, wenn technische Aufwände nach Inbetriebnahme des Systems minimal gehalten werden können.« |
| 4. Erweiterbarkeit und Skalierbarkeit | »Eine Sprachapplikation ist gut, wenn die Systemarchitektur zukünftige Erweiterungen und Veränderungen leicht ermöglicht.« |
| 5. Wirtschaftlichkeit | »Eine Sprachapplikation ist gut, wenn der Betrieb der Sprachapplikation wirtschaftlich rentabel ist.« |
-



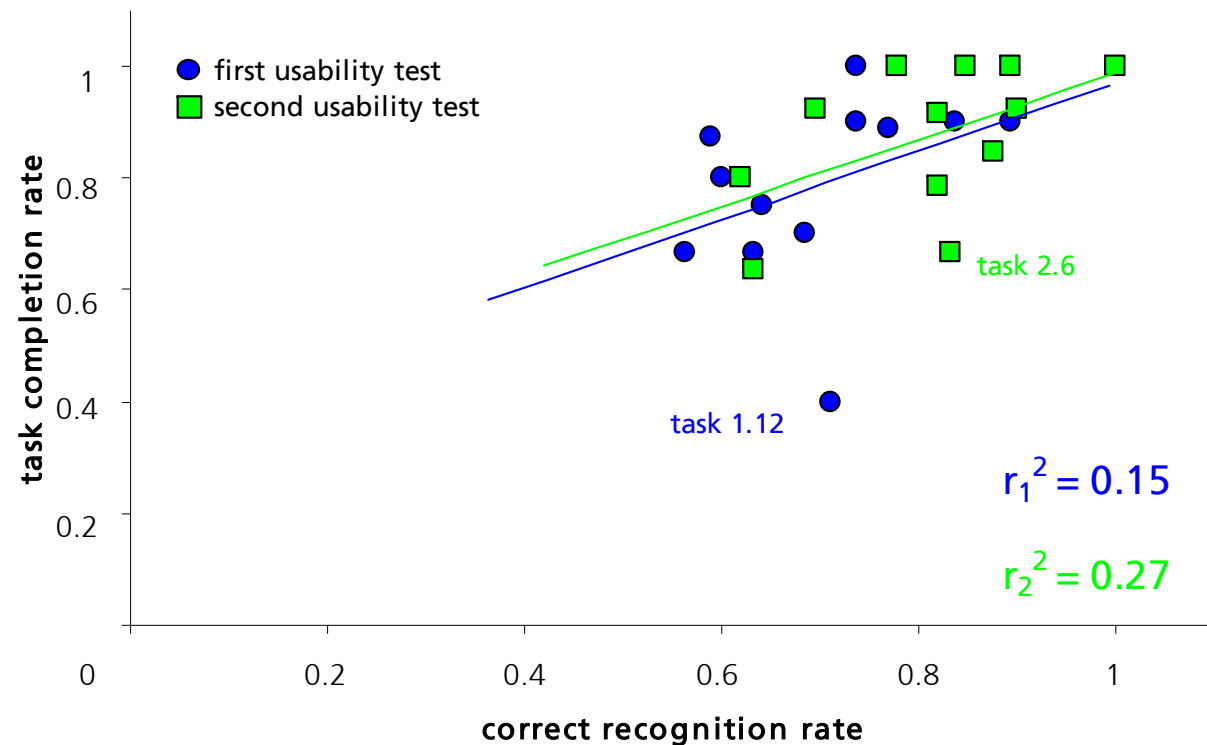
Qualitätskriterien für Sprachapplikationen (6-10)

- | | | |
|--|--|---|
| 6. Zuverlässige Erkennung der Benutzeräußerungen | »Eine Sprachapplikation ist gut, wenn die Spracherkennung einen angemessenen Umfang zu erwartender Benutzeräußerungen zuverlässig erkennt.« | → |
| 7. Effektives Fehlermanagement | »Eine Sprachapplikation ist gut, wenn Erkennungsfehler und Bedienfehler keinen großen Schaden anrichten.« | |
| 8. Effektive und flexible Dialogabläufe | »Eine Sprachapplikation ist gut, wenn die Navigationsstruktur die Benutzer unterstützt, schnell und sicher ihr Ziel zu erreichen.« | → |
| 9. Verständliche und zielführende Systemausgaben | »Eine Sprachapplikation ist gut, wenn die akustischen Systemausgaben den Benutzer bei der Orientierung und der Formulierung zielführender Äußerungen unterstützen.« | |
| 10. Anmutung und emotionale Adressierung | »Eine Sprachapplikation ist gut, wenn eine positive und angemessene Einstellung des Benutzers gegenüber der Sprachapplikation, ihrer Nutzung und ihres Betreibers erzielt wird.« | → |
-



Spracherkennung und Usability

- Der Zusammenhang von Spracherkennungsqualität und subjektiv empfundenem Bedienkomfort variiert stark zwischen verschiedenen Applikationen
- Die Bedienqualität von Sprachdialogsystemen hängt im Durchschnitt nur zu etwa 15 – 30 % von der ASR Performanz ab
- Die Gestaltung der Benutzungsschnittstelle (VUI) ist ein entscheidender Qualitätsfaktor

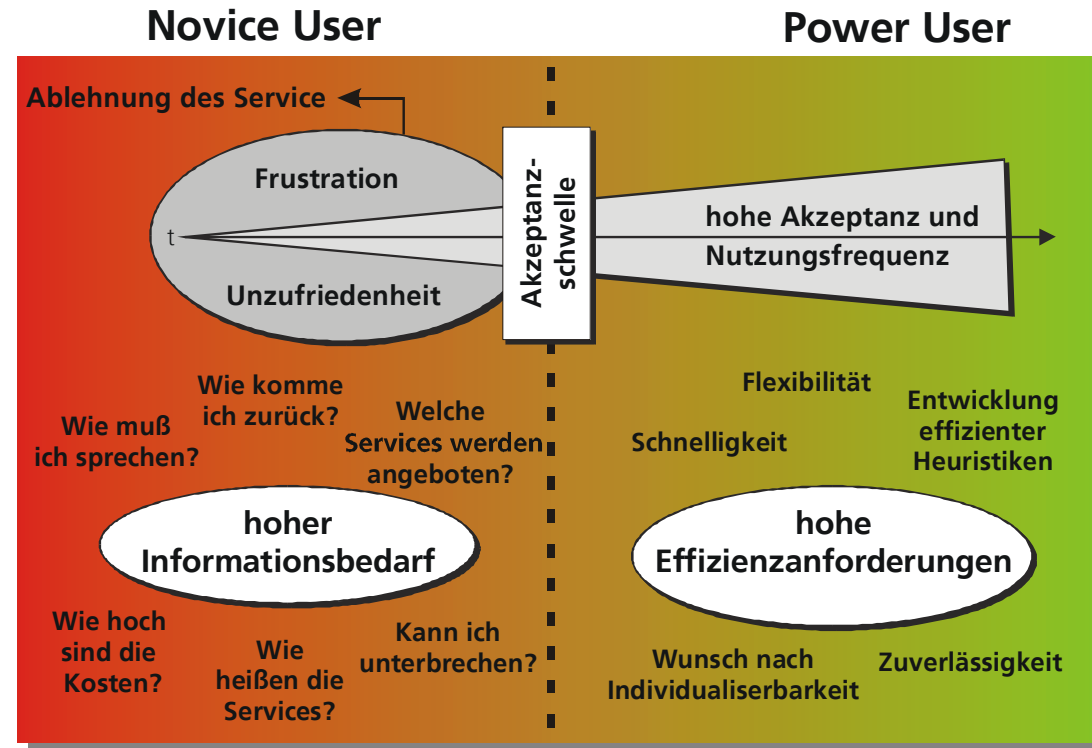


Peissner, M. (2002) "What the Relationship between Correct Recognition Rates and Usability Measures Can Tell Us About the Quality of a Speech Application".



Mixed Initiative: Benutzerführung und Effizienz

- Neue Benutzer bevorzugen eine enge Führung (Menüs).
- Schon nach kurzer Zeit werden „Abkürzungen“ durch das System erwartet. Der Benutzer möchte selbst die Initiative ergreifen und schnell sein Ziel erreichen.
- Wird das nicht unterstützt, resultiert Frustration. Regelmäßige Nutzung und langfristige Akzeptanz bleiben aus.



Peissner, M., Heidmann, F. & Corves, L. (2003) Usability Engineering von Sprachapplikationen. In: Machate, J. & Burmester, M. (Hrsg.): User Interface Tuning – Benutzungsschnittstellen menschlich gestalten. Frankfurt a.M., pp. 299-317.

Prompting: kleine Variationen, große Wirkung

Attraktivität	Stimme und Kommunikationsstil haben großen Einfluss auf die wahrgenommene Service-Attraktivität.
Komplexität und Benutzungsfreundlichkeit	Verschiedenes Prompting führt bei ansonsten identischem Systemdesign zu unterschiedlicher Wahrnehmung bzgl. Komplexität und Benutzungsfreundlichkeit.
Qualität der ASR	Ähnliche Effekte können für die Wahrnehmung der Spracherkennungsqualität beobachtet werden.
Benutzerverhalten	Unterschiedliches Prompting beeinflusst auch wesentliche Aspekte des Benutzerverhaltens: Länge der Benutzeräußerungen, Folgen von Erkennungsfehlern, etc.



Qualität messen und im Prozess verankern

Beispiel: Self-Service Portal

	Projektvorbereitung & Analyse	Konzept & Design	Implementierung	Integration & Inbetriebnahme	Betrieb inkl. Wartung
**	Potenzialanalyse Anforderungsanalyse Business Case Planung	Beispielsammlung (NLU)	Traversaltest Funktionstest	Funktionstest Lasttest Tuning Usability Test	Manuelle Transkription und Annotierung Logfile-Analyse und Reporting Tuning
*	Balanced Scorecard (BSC)	Expertenevaluation oder WOZ-Test	Friendly-User Test		Balanced Scorecard (BSC)

** Diese Verfahren sollten unbedingt eingesetzt werden
* Empfohlen bei besonderen Qualitätsansprüchen

Hipp, C., Paulke, S., Peissner, M. & Steimel, B. (2008). Qualitätsleitfaden - Kochbuch für gute Sprachapplikationen

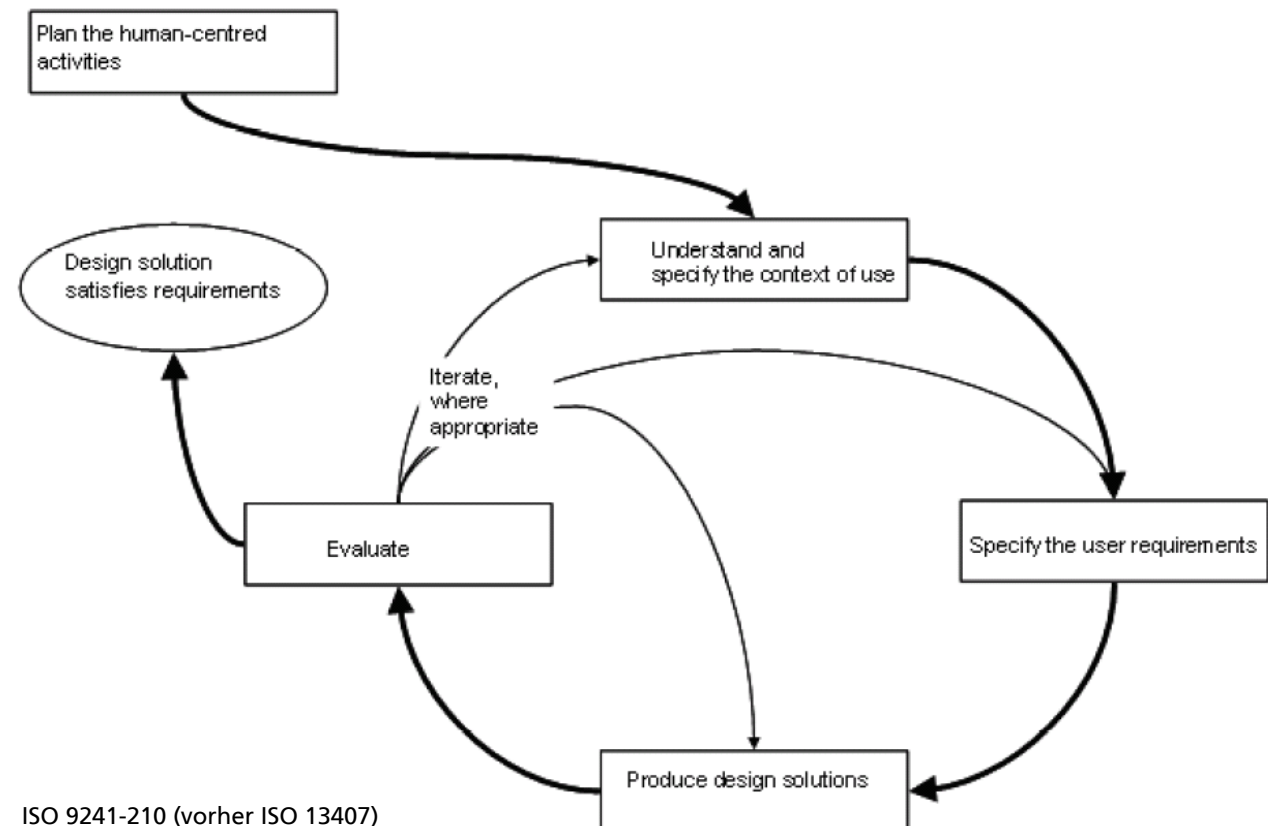


Human-Centred Design

Quality doesn't happen by accident

Principles of HCD

- Understanding users, tasks and environments
- User involvement
- Iteration
- Designing for the user experience
- User centred evaluation
- Multi-disciplinary perspectives



ISO 9241-210 (vorher ISO 13407)



Fazit

- Der Einsatz von Sprachinteraktion erfordert eine detaillierte Analyse der unterstützten Prozesse und Nutzungskontexte.
- Vor allem in Kombination mit grafischen Oberflächen bietet die Sprachinteraktion große Vorteile.
- Die Spracherkennungsqualität ist nur ein Aspekt einer erfolgreichen Applikation - häufig nicht der wichtigste.
- Grundvoraussetzung für die effiziente Nutzung von Sprachsystemen ist ein ergonomisches Voice User Interface.



Danke schön

Erfolgsfaktoren für den Einsatz
von Sprachinteraktion

Matthias Peissner

Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO
Competence Center Human-Computer Interaction

Nobelstraße 12 | D-70569 Stuttgart
matthias.peissner@iao.fraunhofer.de
Telefon +49 (0)711 970 2311

<http://www.hci.iao.fraunhofer.de>