

Ratzka ·

Patternbasiertes User Interface Design für multimodale Interaktion

Schriften zur Informationswissenschaft Band 55

Herausgegeben vom Hochschulverband
für Informationswissenschaft (HI) e.V. Konstanz

Mitglieder des wissenschaftlichen Beirates für die Schriftenreihe:

Nicolas Belkin	Rutgers University
Hans Peter Frei	ETH Zürich
Rainer Hammwöhner	Universität Regensburg
Ilse M. Harms	Universität des Saarlandes
Norbert Henrichs	Universität Düsseldorf
Josef Herget	Donau-Universität Krems
Gerhard Knorz	Hochschule Darmstadt
Jürgen Krause	Universität Koblenz-Landau IZ Sozialwissenschaften
Rainer Kuhlen	Universität Konstanz
Klaus-Dieter Lehmann	Präsident der Stiftung Preußischer Kulturbesitz
Hans-Jürgen Manecke	TU Ilmenau
Achim Oßwald	Fachhochschule Köln
Wolf Rauch	Universität Graz
Harald Reiterer	Universität Konstanz
Marc Rittberger	Hochschule Darmstadt/DIPF
Christian Schlögl	Universität Graz
Dagobert Soergel	University of Maryland
Wolfgang G. Stock	Universität Düsseldorf
Christian Wolff	Universität Regensburg
Christa Womser-Hacker	Universität Hildesheim
Harald Zimmermann	Universität des Saarlandes

Wissenschaftliche Redaktion Wolfgang Semar, HTW Chur

Andreas Ratzka

Patternbasiertes User Interface Design für multimodale Interaktion

**Identifikation und Validierung von
Patterns auf Basis einer Analyse der
Forschungsliteratur und explorativer
Benutzertests an Systemprototypen**

vwh

Verlag Werner Hülsbusch
Fachverlag für Medientechnik und -wirtschaft

A. Ratzka: Patternbasiertes User Interface Design für multimodale Interaktion

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://d-nb.de> abrufbar.

© Verlag Werner Hülsbusch, Boizenburg, 2010

vwh Verlag Werner Hülsbusch
Fachverlag für Medientechnik und -wirtschaft

www.vwh-verlag.de

Einfache Nutzungsrechte liegen beim Verlag Werner Hülsbusch, Boizenburg.
Eine weitere Verwertung im Sinne des Urheberrechtsgesetzes ist nur mit Zustimmung des Autors möglich.

Markenerklärung: Die in diesem Werk wiedergegebenen Gebrauchsnamen, Handelsnamen, Warenzeichen usw. können auch ohne besondere Kennzeichnung geschützte Marken sein und als solche den gesetzlichen Bestimmungen unterliegen.

Druck und Bindung: Kunsthaus Schwanheide

Printed in Germany

Zugleich: Diss., Univ. Regensburg, 2009

– Als Typoskript gedruckt –

ISSN: 0938-8710

ISBN: 978-3-940317-62-9

Andreas Ratzka

Patternbasiertes User Interface Design für multimodale Interaktion

Identifikation und Validierung von Patterns auf Basis einer Analyse der
Forschungsliteratur und explorativer Benutzertests an Systemprototypen



Die Arbeit wurde im Jahr 2009 von der Philosophischen Fakultät III
– Sprach- und Literaturwissenschaften –
der Universität Regensburg als Dissertation angenommen.

D 355

Danksagung

Meinen besonderen Dank möchte ich meinen Betreuern Herrn Professor Rainer Hammwöhner und Herrn Professor Christian Wolff sowie Herrn Dr. Ludwig Hitzenberger aussprechen, welche mir diese Arbeit am Lehrstuhl für Informationswissenschaft ermöglicht haben und stets offen für fachliche Diskussionen waren. Mein Dank gilt ebenfalls den Kollegen im Doktoranden-seminar für ihre Anregungen.

Weiterhin möchte ich den Forscherkollegen danken, die mir auf Konferenzen für mein Forschungsvorhaben wertvolle Hinweise gegeben haben. Insbesondere möchte ich Allan Kelly und Michael VanHilst hervorheben, welche mir im Rahmen der PLoP-Konferenzen (VikingPLoP 2007 bzw. EuroPLoP 2008) umfangreiches Feedback als Konferenz-Mentoren (*shepherds*) gegeben haben.

Für die Durchführung der Benutzertests und die vorbereitenden Handgriffe zur Auswertung der Testdaten möchte ich der studentischen Hilfskraft Sarah Will meinen herzlichen Dank aussprechen. Weiterhin danke ich Bettina Williger, Stefan Engl und Michael Giesz für ihre Unterstützung bei Fragen zur Testmethodik und zur Bedienung des *eye-tracking*-Labors.

Auch möchte ich all denen danken, die die Arbeit im Vorfeld der Einreichung Korrektur gelesen haben: Meinem Vater, meinen Promotionskollegen Rüdiger Heimgärtner und Susanne Mühlbacher, meinem Bruder Wolfgang.

Ferner danke ich meinen Eltern, Kollegen, Freunden und Geschwistern, die mir immer wieder Mut gemacht haben, wenn ich am Gelingen dieser Arbeit zweifelte.

Heidelberg, den 22. Januar 2010

Abstract

Since the 90ies several projects have been investigating multimodal interaction techniques and diverse modality combinations comprising among others speech interaction, gestures, facial expressions, and eye gaze input. The goal of multimodal interaction is to offer more natural, robust, flexible and performant user interfaces.

Traditional approaches for multimodal interaction provide tool support for simplifying implementation and prototyping as well as formalisms for detailed user interface specification. Other approaches are focussing on guideline-based decision support during the initial phases of requirements analysis.

This work aims to complement these traditional approaches. The application of *user interface patterns*, that is of patterns of successful system realisations, has the potential to provide concrete decision support not only to early design phases, but also to later stages of detailed design.

In this work, patterns were identified within a meta analysis of research literature. For this, general principles derived from psychological and HCI-related research results, descriptions of multimodal prototypes that were developed during research projects, as well as empirical research results based on these prototypes were analysed. Altogether, 16 patterns were identified and described.

Within the practical part of this work, both a Wizard of Oz and a stand-alone prototype for e-mail communication were developed for both a PDA and a desktop platform. These prototypes served for exemplary validation of some of the patterns identified in the theoretical part. Subsequently, user tests revealed the plausibility and acceptance of these patterns, especially of *voice-based interaction shortcut*, *multiple ways of input* and *speech-enabled form*.

Analysing the users' eye movements showed, that gaze tracking data can give hints about user intentions. Concrete eye gaze patterns, however, depend strongly on the specific user task. In this field, further research is necessary in order to develop new, more natural and easy to use multimodal operating concepts.

Zusammenfassung

Seit den 90er Jahren wird intensiv in verschiedensten Projekten an multimodalen Interaktionstechniken und verschiedenartigsten Modalitätskombinationen aus Sprache, Gesten, Mimik, Blickbewegungen etc. geforscht. Man hofft, im Rahmen von multimodalen Interaktionsparadigmen natürlichere, robustere, flexiblere und leistungsfähigere Benutzerschnittstellen gestalten zu können.

Traditionelle Entwicklungsansätze für multimodale Interaktion bieten Implementierungs- und Prototyping-Hilfsmittel und Formalismen zur Spezifikation der Benutzerschnittstelle. Andere Ansätze legen ihren Fokus auf guidelinebasierte Entscheidungsunterstützung für die Anfangsphasen der Anforderungsanalyse.

Ziel dieser Arbeit ist es, diese traditionellen Ansätze zu ergänzen. Mit Hilfe von *User Interface Patterns* – von Mustern erfolgreicher Systemrealisierungen – sollen nicht nur frühe Designphasen, sondern auch spätere Stufen des detaillierten Designs durch konkrete Entscheidungshilfen unterstützt werden.

In dieser Arbeit wurde im Rahmen einer Metastudie über die Forschungsliteratur ein Katalog von User Interface Patterns identifiziert. Hierzu wurden allgemeine aus Forschungsergebnissen der Psychologie und Mensch-Maschine-Interaktion abgeleitete Prinzipien, Beschreibungen der im Rahmen von Forschungsprojekten entstandenen multimodalen Prototypen und auf diese basierende empirische Forschungsergebnisse analysiert.

Im Rahmen des praktischen Teils der Arbeit wurden für zwei Szenarien (PDA und Desktop) jeweils ein Wizard-of-Oz-Prototyp und ein eigenständiger „Stand-alone“-Prototyp zur E-Mail-Kommunikation umgesetzt. Diese Prototypen dienten zunächst zur exemplarischen Verifikation einiger der Patterns. Benutzertests zeigten die Plausibilität und Akzeptanz der eingesetzten Muster, insbesondere der Patterns *Sprachkürzel*, *alternative Eingabemöglichkeiten* und *sprachgestütztes Formular*, auf.

Die Untersuchung der Augenbewegungen zeigte, dass Blickbewegungsdaten Hinweise auf Benutzerabsichten bieten können, dies allerdings von der aktuellen Aufgabe abhängt. Hier ist weitere Forschung notwendig, um neue, natürlichere und einfacher zu verwendende multimodale Bedienkonzepte zu entwickeln.

Inhaltsverzeichnis

1	Einleitung	15
1.1	Ziel dieser Arbeit	15
1.2	Aufbau der vorliegenden Arbeit	17
2	Multimodale interaktive Systeme	21
2.1	Modalität	21
2.1.1	Prozessmodell der Mensch-Maschine-Interaktion	23
2.1.2	Bewusst kontrollierte und unbewusst kontrollierte Modalitäten	25
2.1.3	Abgrenzung Modalität, Modus, Medium, Kanal	25
2.2	Multimodalität	26
2.3	Interaktion: zwischen Kommunikation und Aktion	28
2.4	Modalitäten nach dem Vorbild der menschlichen Kommunikation	29
2.4.1	Kanäle der zwischenmenschlichen Kommunikation	29
2.4.2	Sprachliche Interaktion	30
2.4.3	Non-Verbale Interaktion mit <i>perceptual user interfaces</i>	33
2.4.4	Gesten in der MMI	34
2.4.5	Emotionserkennung: Mimik, Gestik und Prosodik	36
2.4.6	Blickinteraktion	37
2.5	Interaktion als werkzeugbasierte Handlung	38
2.5.1	Direkte Manipulation	38
2.5.2	Virtuelle und erweiterte Realität	40
2.5.3	Zweihändige Interaktion	41
2.5.4	Taxonomie von Interaktionswerkzeugen	43
2.5.5	Elementare Interaktionshandlungen	44
2.6	Modality Theory: Ein Beschreibungsmodell für Modalitäten	46
2.6.1	Ausgabemodalitäten	46
2.6.2	Eingabemodalitäten	49
2.6.3	Modality Properties	55
2.7	Multimodale Verarbeitung	57
2.7.1	Multimodale Wahrnehmung	57
2.7.2	Technische Verarbeitung: Fusion und Fission	60
2.7.3	Formen multimodaler Kombination	63
2.8	Das Potential multimodaler Interaktion	65
2.8.1	Adäquate Modalitätsauswahl	67
2.8.2	Kombination mehrerer Interaktionsmodalitäten	72
2.9	Architektur der multimodalen Schnittstelle	78
2.9.1	Das Seeheim-Modell	79
2.9.2	Das Arch-Modell	79
2.9.3	Architekturmuster: <i>Model view controller</i> und <i>presentation abstraction control</i>	81
2.9.4	Die Interaktor-Abstraktion	83
2.9.5	PAC-AMODEUS	86
2.9.6	Multimodale Architekturen auf Basis intelligenter Agenten	86
2.9.7	Bezug zum Software-Engineering	88
2.10	Beispiele multimodaler Systemprototypen	88
2.10.1	Frühe Systeme	88
2.10.2	Einschlägige Forschungsprojekte	89
2.10.3	Einsatzszenarien	95
2.10.4	Zusammenfassung	108

2.11	Standardisierungsansätze durch das W3C	112
2.12	Positionierung dieser Arbeit	115
3	User-Interface-Engineering	117
3.1	Vorgehensmodelle des Software-Engineerings	117
3.2	Entwicklung interaktiver Systeme	120
3.3	Bestandsaufnahme von Entwurfsansätzen für multimodale Interaktion	124
3.3.1	Guideline-basierte Entscheidungsunterstützung	124
3.3.2	Prototyping-Ansätze	128
3.3.3	Abdeckung des <i>usability engineering lifecycles</i>	135
3.4	Modellbasierte Ansätze	135
3.4.1	Bezug zu modellgetriebenen Ansätzen des Software-Engineerings	138
3.4.2	Plattformunabhängige und adaptive Systeme	141
3.4.3	Beispielansatz für adaptive multimodale Interaktion: FAME	143
3.4.4	Probleme und Perspektiven des modellbasierten Ansatzes	144
3.5	Patternbasierte Ansätze	146
3.5.1	Qualitätskriterien für Designpatterns	150
3.5.2	User Interface Patterns	151
3.5.3	Patterns für Softwarearchitekturen interaktiver Systeme	156
3.5.4	Patterns in modellbasierten Ansätzen	157
3.5.5	Patterns für multimodale Interaktion	157
3.6	Fazit	160
4	Identifikation von Patterns für multimodale Interaktion	161
4.1	Ermittlung sinnvoller Interaktionsmodalitäten	162
4.1.1	Modalitätsauswahl auf Basis von Aufgabeneigenschaften	162
4.1.2	Interaction Constraints	163
4.1.3	Ableitung von Patterns?	164
4.2	Potential von Modalitätskombinationen	165
4.3	Potential durch Wechseln der Interaktionsmodalitäten	170
4.4	Identifizieren von Mustern in einschlägigen Anwendungsszenarien	170
4.4.1	Unterstützung von Doppelaufgabenszenarien	172
4.4.2	Kartenbasierte Anwendungen	172
4.4.3	Graphische Design-Anwendungen	173
4.4.4	Mobile Systeme	175
4.4.5	Weitere Aspekte	176
4.5	Fazit	178
5	Eine Patternsammlung für multimodale Interaktion	179
5.1	Aufbau der Patternsammlung	179
5.2	Flexible Interaktion	182
5.2.1	Alternative Eingabemöglichkeiten (<i>multiple ways of input</i>)	184
5.2.2	Globale Kanalkonfiguration (<i>global channel configuration</i>)	187
5.2.3	Kontextadaption (<i>context adaptation</i>)	190
5.3	Effiziente Informationsaufnahme	194
5.3.1	Audiovisuelle Präsentation (<i>audio-visual presentation</i>)	195
5.3.2	Audiovisuelle Arbeitsumgebung (<i>audio-visual workspace</i>)	198
5.4	Robuste Interaktion	201
5.4.1	Redundante Ausgabe (<i>redundant output</i>)	202
5.4.2	Redundante Eingabe (<i>redundant input</i>)	205
5.4.3	Multimodale N-Best-Auswahl (<i>multi-modal n-best selection</i>)	208

5.4.4	Hypothesenreduktion durch Anbuchstabieren (<i>spelling-based hypothesis reduction</i>)	211
5.5	Schnelle Dateneingabe	214
5.5.1	Sprachkürzel (<i>voice-based interaction shortcut</i>)	216
5.5.2	Sprachgestütztes Formular (<i>speech-enabled form</i>)	219
5.5.3	Natürliche Sprache mit Gesten (<i>gesture-enhanced natural speech</i>)	222
5.5.4	Positionsspezifische Geste (<i>location-sensitive gesture</i>)	227
5.6	Graphikbasierte Anwendungen	230
5.6.1	Sprachgestützte Werkzeugleiste (<i>speech-enabled palette</i>)	231
5.6.2	Sprachgestütztes Skizzieren (<i>speech-enhanced sketching</i>)	233
5.6.3	Multimodale Ortssuche (<i>multi-modal spatial search</i>)	236
5.7	Fazit	239
6	Umsetzung multimodaler E-Mail-Prototypen für PDA und Desktop-PC	241
6.1	Interaktives Szenario	241
6.2	Technische Basis	243
6.3	Architekturprinzipien	244
6.3.1	Observer-basierte Architektur	245
6.3.2	Verteilte Architektur	246
6.3.3	Das interaktive System als Zustandsmaschine	249
6.4	Systembeschreibung	250
6.4.1	Wizard-of-Oz-Setting: Desktop	250
6.4.2	Wizard-of-Oz-Setting: PDA	254
6.4.3	Zu simulierende Sprachinteraktion	259
6.4.4	Allgemeine Funktionalität der Versuchsleiterschnittstelle	260
6.4.5	„Stand-alone“-PDA-Setting	266
6.4.6	„Stand-alone“-Desktop-Setting	272
6.5	Fazit	273
7	Explorative Untersuchungen zur multimodalen Interaktion	275
7.1	Empirischer Forschungsstand zur multimodalen Interaktion	275
7.1.1	Studien zur Modalitätspräferenz und Modalitätsverteilung	276
7.1.2	Auswirkung von Multimodalität auf die Sprachsyntax	277
7.1.3	Interindividuelle Unterschiede	279
7.1.4	Fehlerbehandlung und -vermeidung	280
7.1.5	Schnellere Interaktion	280
7.1.6	Auswirkung der sozialen Umgebung auf die Akzeptanz	281
7.1.7	Natürlichkeit	282
7.2	Das Ziel dieser explorativen Forschung	282
7.3	Untersuchungen multimodaler Interaktion am PDA und Desktop-PC	283
7.3.1	Versuchaufbau	284
7.3.2	Versuchsdesign	287
7.3.3	Datenerhebung und Metriken	290
7.3.4	Annahmen und Fragestellungen	292
7.3.5	Testpersonen	296
7.3.6	Subjektive Systembewertungen	298
7.3.7	Interaktionseffizienz (Wizard-of-Oz-Setting)	300
7.3.8	Erkennungsrakurtheit („Stand-alone“-Setting)	306
7.3.9	Modalitätsverteilung	306
7.3.10	Weitere Beobachtungen	315
7.3.11	Folgerungen für die Patterns	317

7.4	Untersuchungen der Augenbewegungen bei multimodaler Interaktion	319
7.4.1	Augenbewegungen als unbewusst kontrollierte Eingabemodalität – Stand der Forschung	319
7.4.2	Fragestellungen	323
7.4.3	Versuchaufbau	323
7.4.4	Datenerhebung und Metriken	324
7.4.5	Beobachtungen	324
7.4.6	Folgerungen	326
7.5	Fazit	326
8	Schlussbemerkungen	329
8.1	Zusammenfassung	329
8.2	Ausblick	330
	Literaturverzeichnis	333
	Abkürzungsverzeichnis	375
	Abbildungsverzeichnis	379
	Tabellenverzeichnis	381
A	Versuchssetting	383
A.1	Checkliste für den Versuchsleiter (Wizard-Setting)	383
A.2	Erfassungsbogen	386
A.3	Systemdemonstration für die Testpersonen (Stand-alone-Setting)	387
A.4	Interviewfragebogen für das Wizard-of-Oz-Setting	388
A.5	Interviewfragebogen für das Stand-alone-Setting	391

Ein weiterer Anhang mit einer detaillierten Beschreibung des Versuchssettings und Daten aus den Benutzertests ist online verfügbar unter:

<http://www.vwh-verlag.de/vwh/?p=458>

1 Einleitung

In den 80er Jahren begann mit dem Erscheinen der Xerox-Star-Workstation (1980) und des Apple-Macintosh (1984) die Ära der graphischen Benutzerschnittstellen, welche sich in den 1990er Jahren mit dem Erscheinen und der Verbreitung des Windows-Betriebssystems etablierte (vgl. Dahm 2006, S. 199). Diese Benutzerschnittstellen basieren auf dem WIMP-Paradigma (Windows, Icons, Menus, Pointing device), das, zumindest im Desktop-Bereich, andere Paradigmen wie Kommandosprachen oder getippte natürliche Sprache allmählich zu verdrängen begann.

Ebenfalls in den 80er Jahren entstanden erste automatische telefonbasierte Informationsportale und Hotlines, die dem Benutzer den Zugriff auf eine Reihe von Anwendungen wie Wetterdaten, Börsendaten, persönliche Nachrichten etc. ermöglichten. Die ersten dieser Systeme wurden mit den Telefontasten (Mehrfrequenzwahl) gesteuert (Aust 1998, S. 21 f.). Um die daraus resultierenden tiefen Menühierarchien zu vermeiden und auch den Zugang mit Hilfe von älteren (Wählscheiben-)Telefonen zu ermöglichen, wurde alsbald an der sprachbasierten Steuerung solcher Anwendungen geforscht, welche seit den letzten Jahren eine immer größere Verbreitung finden – man denke an die automatische Bahnauskunft oder die Kundenhotlines diverser Telekommunikationsanbieter (vgl. auch Aust 1998, S. 22 ff.).

Das „Put-That-There“-System (Bolt 1980) ging einen Schritt weiter, indem es graphische Ausgabe und Zeigegesten auf der einen Seite und Sprachinteraktion auf der anderen Seite zu einer neuen multimodalen Interaktionsform verschmolz. Seit den 90er Jahren wird intensiv in verschiedensten Projekten (vgl. Kapitel 2.10, S. 88) an multimodalen Interaktionstechniken und verschiedenartigsten Modalitätskombinationen aus Sprache, Gesten, Mimik, Blickbewegungen etc. geforscht. Man hofft, im Rahmen von multimodalen Interaktionsparadigmen natürlichere, robustere, flexiblere und leistungsfähigere Benutzerschnittstellen gestalten zu können (Oviatt (1996b), Hedicke (2002, S. 206 f.), vgl. auch Kap. 2.8, S. 65).

1.1 Ziel dieser Arbeit

Um in dem durch die Einbeziehung unterschiedlichster Interaktionstechniken erweiterten Gestaltungsspielraum das Potential multimodaler Interaktion auszuschöpfen, benötigen Entwickler Methoden und Werkzeuge, die sie beim Erstellen multimodaler Systeme unterstützen. Traditionelle Ansätze für multimodale Interaktion bieten Implementierungs- und Prototyping-Hilfsmittel (Niedermaier 2003, vgl. auch Kap. 3.3.2, S. 128) und Formalismen zur detaillierten Spezifikation der Benutzerschnittstelle (Dragičević 2004). Andere

Ansätze legen ihren Fokus auf guidelinebasierte Entscheidungsunterstützung für die Anfangsphasen der Anforderungsanalyse (Bernsen 2001, Bürgy 2002, vgl. auch Kap. 3.3.1, S. 124).

Ziel dieser Arbeit ist es, diese traditionellen Ansätze zu ergänzen. Mit Hilfe von *User Interface Patterns* – von Mustern erfolgreicher Systemrealisierungen – sollen nicht nur frühe Designphasen sondern auch spätere Stufen des detaillierten Designs durch konkrete Entscheidungshilfen unterstützt werden (vgl. Kapitel 3.5, S. 146).

Ein *Muster* oder *Pattern* ist dabei definiert als Regel, welche ein gängiges Entwurfsproblem mit einer bewährten Lösung und einer Beschreibung des Kontextes, in dem diese Lösung anwendbar ist, in Verbindung bringt (Brown et al. 1998, Gabriel 2007). Patterns sind gewissermaßen eine hypertextuell verlinkte Weiterentwicklung von guidelinebasierten Ansätzen. Im Gegensatz zu Guidelines bieten Patterns eine strukturierte Präsentation des Inhalts, klären über Hintergründe und Zielkonflikte auf, verweisen auf alternative oder ergänzende Patterns und bauen auf bereits mehrfach bewährte Lösungen.

Die Muster-Idee stammt ursprünglich aus der Architektur und Stadtplanung (Alexander et al. 1977, Alexander 1979), wurde aber hauptsächlich in verschiedenen Gebieten der Softwareentwicklung wie der objektorientierten Programmierung (Gamma et al. 1995), der Softwarearchitektur (Buschmann et al. 1998) und der Gestaltung von Benutzerschnittstellen (Tidwell 1999, 2005, van Welie u. Trätteberg 2000, Borchers 2000, 2001a, b) rezipiert.

Die Voraussetzung, dass Patterns auf bewährten Lösungen basieren müssen, lässt die Anwendung dieses Ansatzes auf multimodale Interaktion als schwierig erscheinen. Aber auch wenn multimodale Interaktion noch keine Marktdurchdringung erreicht hat, kann die Forschung in diesem Feld auf eine fast dreißigjährige Geschichte zurück blicken, in deren Verlauf zahlreiche Demonstratoren und Forschungssysteme entstanden sind. Für wiederkehrende Interaktionsprobleme wurden Lösungen ermittelt, welche in Nachfolgeprojekten erfolgreich wiederverwendet wurden, sich bewährt haben und somit als Grundlage für User Interface Patterns dienen (vgl. Ratzka u. Wolff 2006).

In dieser Arbeit wird im Rahmen einer Metastudie aus der Forschungsliteratur ein Katalog von User Interface Patterns identifiziert (vgl. Kapitel 4 und 5, S. 161 und 179). Hierzu werden allgemeine aus Forschungsergebnissen der Psychologie und Mensch-Maschine-Interaktion abgeleitete Prinzipien, Beschreibungen der im Rahmen von Forschungsprojekten entstandenen multimodalen Prototypen und auf diesen Prototypen basierende empirische Forschungsergebnisse analysiert. Weiterhin werden Beziehungen innerhalb der Muster für multimodale Interaktion sowie deren Beziehungen zu GUI-Patterns

(Patterns für graphische Benutzerschnittstellen) (vgl. Tidwell 1999, 2005, van Welie u. Trätteberg 2000) dargelegt.

Ein Problempunkt ist die Validierung von User Interface Patterns. Psychologisch orientierte Grundlagenforschung zum Benutzerverhalten ist weder zur Verifikation noch zur Entdeckung von User Interface Patterns uneingeschränkt geeignet, da ein Pattern zu vielschichtig ist, um sich auf messbare Zahlen reduzieren zu lassen. Hierzu wäre eine Vielzahl von Untersuchungen notwendig, um alle Aspekte, Implementierungsoptionen und Einsatzkontexte ausreichend zu würdigen und zu einem generalisierbaren Ergebnis zu kommen. Andererseits beruhen Muster auf bewährten Lösungen und erforschten Prinzipien und besitzen bereits ein gewisses Maß an Validität. Allerdings kann nicht ausgeschlossen werden, dass bei der Formulierung von Patterns und der dabei stattfindenden Abstraktion einzelne Aspekte überbetont und andere wichtige Aspekte übersehen werden.

Daher wurden exemplarische Prototypen einer multimodalen E-Mail-Anwendung für mobile Systeme und Desktop-Plattformen umgesetzt (vgl. Kapitel 6, S. 241). Diese Prototypen dienen zunächst zur exemplarischen Verifikation einiger der identifizierten Muster. Daraufhin wurde mit Hilfe dieser Prototypen in Benutzertests die Plausibilität und Akzeptanz der eingesetzten Muster explorativ ermittelt (vgl. Kapitel 7, S. 275).

1.2 Aufbau der vorliegenden Arbeit

Kapitel 2 (S. 21) beschreibt die theoretischen Hintergründe multimodaler interaktiver Systeme. Dabei werden zunächst grundlegende Begriffe wie Modalität, Modus, Medium, Kanal, Multimodalität geklärt. Im Anschluss wird die Mensch-Maschine-Interaktion aus dem Blickwinkel der menschlichen Kommunikation und der werkzeugartigen Interaktion beleuchtet. Es folgen Betrachtungen zur Modalitätstaxonomie, zu Theorien der multimodalen Verarbeitung, zum softwareergonomischen Potential multimodaler Interaktion und zur Architektur. Daraufhin werden einschlägige multimodale Systeme vorgestellt, welche die Grundlage für spätere Diskussionen zur Identifikation von User Interface Patterns bilden (siehe auch Kap. 4.4, S. 170). Schließlich werden Standardisierungsansätze des W3C im Bereich multimodaler Interaktion vorgestellt.

Nach der Darstellung der theoretischen Grundlagen erfolgt in Kapitel 3 (S. 117) eine Diskussion methodischer Aspekte. Dieses Kapitel stellt zunächst Vorgehensmodelle aus dem Bereich Software-Engineering und allgemeine Ansätze zur Entwicklung interaktiver Systeme vor. Darauf folgt eine Bestandsaufnahme von Entwicklungsansätzen für multimodale interak-

tive Systeme. Schließlich werden mit modellbasierten und patternbasierten Ansätzen zwei gegensätzliche, sich aber gut ergänzende Herangehensweisen erörtert. Während auf dem Gebiet der Entwicklung multimodaler Benutzerschnittstellen bereits ein stetig wachsendes Inventar an Forschungen zu modellbasierten Ansätzen besteht, trifft dies für patternbasierte Herangehensweisen nicht zu, was die Motivation dieser Arbeit begründet.

Anknüpfend an die methodischen Betrachtungen des vorhergehenden Kapitels beschreibt Kapitel 4 (S. 161) den dieser Arbeit zu Grunde liegenden Identifikationsprozess für Patterns multimodaler Schnittstellen, der sowohl Top-down- als auch Bottom-up-Elemente enthält. Zunächst werden Faktoren zur Ermittlung aufgaben- und kontextspezifisch sinnvoller Modalitäten dargestellt. Im Anschluss wird der Mehrwert, der durch die Kombination und den Wechsel von Modalitäten entsteht, untersucht (Top-down-Ansatz). Abschließend wird (Bottom-up-Ansatz) an Hand von einschlägigen Anwendungsszenarien (Doppelaufgabenszenarien, kartenbasierte Anwendungen, graphische Design-Anwendungen, mobile Systeme) die Existenz der zuvor ermittelten Muster nachgewiesen und durch neu identifizierte Patterns ergänzt. Zudem werden die Anwendbarkeit von GUI-Patterns (Patterns für graphische Benutzerschnittstellen) in diesen Anwendungsszenarien und bestehende Beziehungen zwischen den Mustern untersucht.

Die im vorhergehenden Kapitel identifizierten Patterns und Patternbeziehungen werden in Kapitel 5 (S. 179) detailliert beschrieben. Nach einleitenden Bemerkungen zum Aufbau der Patternsammlung werden Muster für flexible Interaktion, Patterns für die effiziente Informationsaufnahme, Muster für robuste Interaktion, Patterns zur schnellen Dateneingabe und Muster für grafikbasierte Anwendungen vorgestellt.

Der in den beiden vorherigen Kapiteln beschriebene methodische Beitrag dieser Arbeit wird in Kapitel 6 (S. 241) durch die praktische Umsetzung von Prototypen im Rahmen eines multimodalen Desktop-basierten und mobilen E-Mail-Szenarios ergänzt. Nach der Beschreibung des interaktiven E-Mail-Szenarios, der technischen Basis und grundlegender Architekturprinzipien werden die einzelnen Prototypen beschrieben. Dabei wurden für beide Szenarien (PDA und Desktop) jeweils sowohl ein Wizard-of-Oz-Prototyp als auch ein eigenständiger „Stand-alone“-Prototyp umgesetzt. Neben GUI-Patterns wie *autocompletion*, *two-panel selector* (Tidwell 2005), *continuous filter*, *shield* (van Welie u. Trætteberg 2000) etc.¹ kamen von den im vorangehenden Kapitel beschriebenen Mustern vor allem die Muster *Sprachkürzel* (S. 216), *Kontext-*

¹ Für eine Kurzbeschreibung hier relevanter Patterns anderer Autoren siehe auch Tabelle 3.3, Seite 151 ff.

adaption (S. 190), *alternative Eingabemöglichkeiten* (S. 184) und *sprachgestütztes Formular* (S. 219) zum Einsatz. Im „Stand-alone“-Setting wurden außerdem die Muster *globale Kanalkonfiguration* (S. 187), *multimodale N-Best-Auswahl* (S. 208) und *Hypothesenreduktion durch Anbuchstabieren* (S. 211) exemplarisch umgesetzt.

Anknüpfend an den praktischen Teil wird in Kapitel 7 (S. 275) der empirische Beitrag dieser Arbeit beschrieben. Nach einer Zusammenfassung des empirischen Forschungsstandes werden die im Rahmen dieser Arbeit durchgeführten explorativen Untersuchungen an den Desktop- und PDA-basierten Wizard-of-Oz- und „Stand-alone“-Prototypen diskutiert. Neben subjektiven Systembewertungen standen insbesondere Fragen der aufgaben- und kontextbezogenen Vorlieben der Benutzer für bestimmte Modalitäten sowie Folgerungen für die eingesetzten Patterns im Vordergrund. Weiteres Ziel war es, einen Eindruck vom tatsächlichen Nutzerverhalten bei multimodaler Interaktion zu gewinnen, um davon ausgehend für die Patterns relevante Aspekte besser einschätzen zu können. Abschließend wird in diesem Kapitel noch eine kleinere Untersuchung der Augenbewegungen bei multimodaler Interaktion mit dem „Stand-alone“-Desktop-Szenario beschrieben (Kap. 7.4, S. 319).