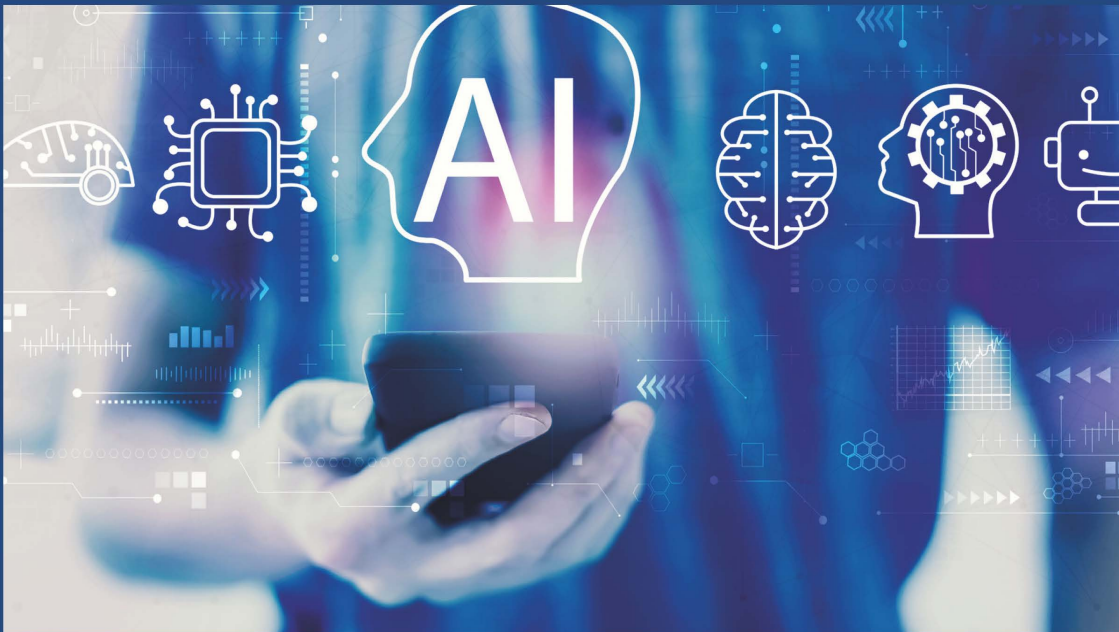


Marc Doerfer

Ergonomiefaktoren auditiver Benutzerschnittstellen

Eine empirische Analyse am Beispiel von Sprachportalen



Cuvillier Verlag Göttingen
Internationaler wissenschaftlicher Fachverlag



Ergonomiefaktoren auditiver Benutzerschnittstellen





Ergonomiefaktoren auditiver Benutzerschnittstellen

Eine empirische Analyse am Beispiel von Sprachportalen

An der Fakultät für Wirtschaftswissenschaft der Bergischen
Universität Wuppertal, Schumpeter School of Business and
Economics, angenommene Inauguraldissertation

Vorgelegt von Marc Doerfer

Mit einem Geleitwort von
Prof. Dr. Gerold Behrens



Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliographische Daten sind im Internet

über <http://dnb.d-nb.de> abrufbar.

1. Aufl. - Göttingen: Cuvillier, 2021

Zugl.: Wuppertal, Univ., Diss., 2020

© CUVILLIER VERLAG, Göttingen 2021

Nonnenstieg 8, 37075 Göttingen

Telefon: 0551-54724-0

Telefax: 0551-54724-21

www.cuvillier.de

Alle Rechte vorbehalten. Ohne ausdrückliche Genehmigung des Verlages ist es nicht gestattet, das Buch oder Teile daraus auf fotomechanischem Weg (Fotokopie, Mikrokopie) zu vervielfältigen.

1. Auflage, 2021

Gedruckt auf umweltfreundlichem, säurefreiem Papier aus nachhaltiger Forstwirtschaft.

ISBN 978-3-7369-7250-6

eISBN 978-3-7369-6250-7



Geleitwort

Was vor 50 Jahren noch kurios war und futuristisch klang, ist heute gebräuchlich und verbreitet: die Kommunikation mit Computern. Interaktive Computer stehen beispielsweise in vielen Haushalten als intelligente persönliche Assistenten wie „Alexa“ (Amazon Echo), „Cortana“, „Siri“ und „Google Assistant“. Das sind Apparate, mit denen wir sprechen können. Sie sind in der Lage, zahlreiche Befehle auszuführen und Informationen zu liefern, können Musik abspielen, Nachrichten senden, Hausgeräte an- und ausstellen sowie steuern, Termine verwalten und u.v.a.m. Für „Alexa“ stehen beispielsweise mehrere tausend Anwendungsbeispiele zur Verfügung. Auch in neuen Autos sprechen wir mit elektronischen Geräten. Es können über Sprachbefehle Radio und Telefon bedient werden; man kann die Klimaanlage regeln und die Navigation steuern.

Im privaten Bereich und im Beruf werden am Telefon Dialoge mit Computern geführt, wenn wir Firmen anrufen. Das können sehr einfache Dialoge sein. So werden wir beim Anruf einer Arztpraxis beispielsweise aufgefordert, eine bestimmte Taste zu drücken, wenn lediglich ein Attest ausgestellt werden soll. Für andere Leistungen müssen andere Tasten gedrückt werden. Nach dem Tastendruck werden die Aufgaben ausgeführt. Dabei werden die notwendigen Informationen durch einfache Dialoge vom Computer abgefragt. Das sind sehr einfache Systeme, die es seit den 1970er Jahren gibt und die seit den 1980er Jahren bis heute in der Geschäftswelt verbreitet sind. Inzwischen wurden aber große Fortschritte bei der Sprachwahrnehmung, dem Verstehen von Sprache und der Sprachausgabe erzielt.

Es können Dialoge in natürlicher Sprache geführt werden. Bei standardisierbaren Aufgaben klappt das schon ganz gut. Daher sind Sprachportale verbreitet und ihre Bedeutung wird zunehmen, denn ihr Einsatz bietet Unternehmen große Vorteile. An erster Stelle stehen die Kostenvorteile. Das Gespräch mit Servicemitarbeitern ist für die Firma teuer. Sie werden durch Sprachportale entlastet und können für anspruchsvollere Aufgaben eingesetzt werden. Sprachportale haben aber auch für Kunden Vorteile.



Sie können die Erreichbarkeit erhöhen und die Wartezeiten verringern. Trotzdem lösen sie bei Nutzern häufig Unbehagen aus. Das ist problematisch, denn diese Kundenkontakte sind für die Kundenzufriedenheit wichtig.

Hier knüpft die Veröffentlichung von Herrn Doerfer an. Übergeordnetes Ziel seiner Arbeit ist die Verbesserung der Nutzerakzeptanz von Sprachportalen. Im theoretischen Teil wird dieses Problem aus verschiedenen Perspektiven beleuchtet. Um die Möglichkeiten und Grenzen eines Dialoges auszuloten, untersucht er Sprachaufnahme, Sprachausgabe und das Verstehen von Texten. Dabei werden sowohl Menschen als auch Computer betrachtet. Das Problem, das die Dialoge am stärksten begrenzt, ist das Verstehen der Inhalte. Früher haben die Computer nur auf bestimmte Zeichen in einer bestimmten Weise reagiert, heute können sie angemessen auf verschiedene Inhalte antworten. Außenstehende gewinnen den Eindruck, dass Computer einfache Texte verstehen. Das eröffnet neue Möglichkeiten für den Einsatz von Sprachportalen und erhöht deren Akzeptanz.

Die Akzeptanz von Sprachportalen wird außerdem durch eine bequeme Handhabung erhöht. Dieses ergonomische Problem wird in der Arbeit behandelt. Es werden zahlreiche Kriterien vorgestellt, durch die ergonomische Aspekte bewertet und somit auch kontrolliert verbessert werden können. Im Mittelpunkt der Veröffentlichung von Herrn Doerfer steht aber die Untersuchung der Nutzerakzeptanz aus psychologischer Sicht. Diese Betrachtungsweise ist bisher bei Sprachportalen vernachlässigt worden. Anregungen für die Bearbeitung dieser Fragestellung findet man in der Akzeptanz- und der Konsumenten-Verhaltensforschung. Auf der Basis der hier gewonnenen Erkenntnisse wird ein Design für die empirische Untersuchung der Nutzerakzeptanz von Sprachportalen entwickelt. In einem Experiment werden die charakteristischen Merkmale von 12 Sprachportalen systematisch variiert. Dadurch kann kontrolliert untersucht werden, wie Systemeigenschaften, die die Nutzerakzeptanz beeinflussen, auf Nutzer wirken.

Auf diese Weise wird der Einfluss von Erfahrungen mit Sprachportalen auf die Bereitschaft untersucht, sie zu nutzen. Darüber hinaus wird gemessen, wie die Integration von Musik in den Dialogen aufgenommen wird. Ein weiterer wichtiger Einflussfaktor auf die Nutzerakzeptanz ist die kognitive Belastung, also das Behalten von Informationen, die für eine erfolgreiche Benutzung des Sprachportals notwendig sind. Darauf wird auch eingegangen. Außerdem wird berücksichtigt, dass die Gestaltung der Inhalte das Aufnehmen und Behalten von Informationen beeinflusst.

Die Ergebnisse zeigen, dass der Einsatz von Musik mit vielen Problemen verbunden ist. Sie wird häufig als störend empfunden, aber nicht immer. Das hängt auch mit den Musikvorlieben der Nutzer zusammen und der Situation, in der sie sich gerade befinden. Diese Informationen liegen aber nicht vor. Ein weiteres Anwendungsproblem: Sehr wichtig ist es, die kognitive Belastung der Nutzer gering zu halten. Sie behalten nur wenige Informationen und werden daher schnell überlastet. Wer beispielsweise zu Anfang eines Dialoges einen großen Überblick über seine Serviceleistungen gibt, kann nicht davon ausgehen, dass der Dialogpartner, der Kunde, sich zu einem späteren Zeitpunkt des Dialogs an Einzelheiten des Angebots erinnert. Häufig werden im Laufe solcher Gespräche aber Informationen vorausgesetzt, die in der einleitenden Anfangsphase gegeben wurden, z.B. Überblicke über Serviceleistungen. In der Arbeit wird gezeigt, dass diese Probleme der kognitiven Überlastung durch eine geschickte Textgestaltung und einen klugen Dialogaufbau reduziert werden können.

Abschließend ein zusammenfassender Überblick: In dieser Untersuchung werden aus theoretischen Überlegungen Hypothesen gewonnen und empirisch überprüft, die für praktische Problemlösungen relevant sind. Somit werden durch eine experimentelle Vorgehensweise theoretische Überlegungen mit praktischen Problemlösungen verbunden. Die theoretischen Überlegungen liefern zahlreiche Anknüpfungspunkte, mit denen weitere Aspekte der Fragestellung vertieft diskutiert werden können.



Die experimentelle Vorgehensweise erzeugt –wie alle wissenschaftlichen Vorgehensweisen– keine Gewissheiten, aber die gewonnenen Aussagen sind belastbarer als Vermutungen, Spekulationen und gewöhnliches Alltagswissen. Die praktischen Hinweise liefern Anregungen für Verbesserungen und sensibilisieren für Probleme. Alles in allem ist eine Veröffentlichung entstanden, die zugleich anregend und informativ ist.

Prof. Dr. Gerold Behrens

Emeritus

Lehrstuhl für Betriebswirtschaft, insb. Marketing

Schumpeter School of Business and Economics

Bergische Universität Wuppertal

Danksagung

Die vorliegende Arbeit ist am Lehrstuhl für Marketing der Schumpeter School of Business and Economics der Bergischen Universität Wuppertal entstanden und wurde im Mai 2020 als Dissertation angenommen. Während der nebenberuflichen Anfertigung der Dissertationsschrift als externer Doktorand habe ich von verschiedenen Seiten viel Motivation, Unterstützung und Inspiration erfahren. Ferner habe ich auf vielen Ebenen dazugelernt. Dafür bin ich äußerst dankbar.

Mein besonderer Dank gilt Herrn Prof. Dr. Gerold Behrens, der mir die Möglichkeit der Promotion an der Bergischen Universität eröffnete und dieses Projekt trotz seiner zwischenzeitlichen Emeritierung kontinuierlich unterstützt und gefördert hat. Danken möchte ich in diesem Zusammenhang auch gern Herrn Dr. Thorsten Böth, der sich seiner Zeit erfolgreich für die Herstellung des Erstkontakts mit dem Lehrstuhl eingesetzt hat.

Sehr herzlich bedanke ich mich darüber hinaus für die Betreuung meines Promotionsprojekts bei Herrn Prof. Dr. Tobias Langner. Herr Prof. Langner hat mich als Lehrstuhlinhaber mit enormer Fachkompetenz und stets hohem wissenschaftlichen Anspruch erfolgreich durch den Promotionsprozess begleitet.

Für das Mitwirken an meinem Prüfungsverfahren möchte ich Frau Prof. Dr. Ina Garnefeld, Herrn Prof. Dr. Michael Fallgatter sowie Herrn Prof. Dr. Stefan Zielke ebenso meinen herzlichen Dank aussprechen, wie Herrn Prof. Dr. Dirk Temme und Prof. Dr. Alexander Fischer für entscheidende fachliche Gespräche. Für die wichtige organisatorische Unterstützung beim empirischen Teils der Arbeit gilt mein besonderer Dank Frau Laura Erner sowie Herrn Prof. Dr. Michael Gehle (+), dem ich die vorliegende Arbeit gerne widmen möchte.

Nicht zuletzt geht ein großes Dankeschön an meine Kollegen, Freunde und an meine Familie für die immens wichtige Motivation und Unterstützung, allen voran meiner lieben Mutter, Christa Doerfer, für ihr unermüdliches Korrekturlesen.

Marc Doerfer

Zusammenfassung

Die Bedienung von Systemen per Sprache erhält vermehrt Einzug in unseren Alltag. Im Bereich der Künstlichen Intelligenz werden Sprachdialoge zwischen Mensch und Maschine entwickelt, die von einem Gespräch zwischen Menschen kaum zu unterscheiden sind. Ein Anwendungsbereich für Sprachapplikationen sind Sprachportale, bzw. sogenannte Interactive Voice Responses (IVR). Diese sind Call Centern als wichtigem Touchpoint zwischen Kunden und Unternehmen vorgeschaltet. Kunden klagen in der Praxis über zu komplex zu bedienende Portale. Studien belegen dies. Im theor. Teil der Arbeit wird auf Basis eines Abgleichs sprachlicher Aspekte im Dialog zwischen Mensch und Maschine analysiert, ob Ergonomiekriterien bereits im Sinne eines kundenzentrierten Designs von Sprachportalen Anwendung finden. Zur Erstellung und Evaluierung solcher Portale werden häufig DIN-Normen zur Mensch-Maschine-Interaktion und zur Software-Ergonomie herangezogen, die jedoch originär für haptische oder grafische Schnittstellen konzipiert wurden. Es wurden daher Kriterien aus der Psychologie abgeleitet, die einen engeren Bezug zu rein auditiven Schnittstellen aufweisen. So wurden 12 Portal-Hauptmenü-Modelle spezifiziert, deren Anzahl an Hauptmenüpunkten variiert, ohne, bzw. mit beruhigender oder aktivierender Musikuntermalung sowie mit oder ohne Chunking up konfiguriert wurden. Diese Modelle wurden von $n = 240$ Probanden im experimentellen Teil der Arbeit bzgl. ihrer Wirkungen auf abhängige Variablen bewertet. Gemessen wurden u.a. Behaltensleistungen der Hauptmenü-Begriffe, die Nutzerzufriedenheit, die empfundene Nützlichkeit und Nutzerfreundlichkeit sowie die potenziellen Verhaltensabsichten im Sinne der Einstellungsakzeptanz. Die gewonnenen Erkenntnisse zeigen, dass unter optimalen Versuchsbedingungen max. 2-3 Items behalten werden können. Musikuntermalungen wirken kontraproduktiv auf die AV, während Chunking up einen eher förderlichen Effekt hat. Musik wird als störend empfunden, wenn diese nicht gemocht wird. Neben diesen Faktoren wirkt auch die bisherige Erfahrung des Nutzers auf die Evaluierung der Portal-Modelle. Deutlich wird ebenfalls, dass ein Sprachportal durchaus Auswirkung auf die empfundene Kundenorientierung und die pot. Kaufbereitschaft des Konsumenten haben kann.

Abstract

To operate systems by language is increasingly entering our everyday lives. In the field of Artificial Intelligence, speech dialogues between human and machine are being developed that are barely distinguishable from a conversation between people. One area of speech applications are voice portals, or so-called Interactive Voice Responses (IVR). They must be used by a caller before a personal dialogue with a call center agent -as an important touchpoint between customers and companies- will be enabled. In practice, customers complain about systems that are too complex to use. Studies prove this. In the theoretical part of the thesis, a research of linguistic aspects in the dialogue between human and machine will be the basis to analyze whether ergonomics criteria are already applied in the sense of a customer-centered design of voice portals. For the creation and evaluation of such portals, DIN standards for human-machine interaction and software ergonomics are frequently used, but these were originally designed for haptic or graphical interfaces. Therefore, criteria from the field of psychology were derived, which have a closer relevance to pure auditory interfaces. Thus, 12 portal menu models have been specified whose number of main menu items have been varied - and presented without, or with calming or activating music background, and with or without chunking up. These portal menu models have been evaluated by 240 raters in the experimental part of the study for their effects on dependent variables. Measurements were made of the recall of menu terms, the user satisfaction, the perceived usefulness and -ease of use as well as the potential behavioral intentions in the sense of an acceptance attitude. The results show that max. 2-3 items can be kept - under ideal experimental conditions. Background music can be counterproductive to the dependent variable, while Chunking up has a more beneficial effect. Music is distracting if it is not liked. Apart from these factors, the previous experience of the user also affects the evaluation of the portal models. It also becomes clear that a voice portal can have effects on the perceived customer orientation and the potential consumer willingness to buy.





I. Inhaltsverzeichnis

I. Inhaltsverzeichnis.....	I
II. Abbildungsverzeichnis.....	V
III. Tabellenverzeichnis.....	VII
IV. Abkürzungsverzeichnis	X
1. Einleitender Teil.....	1
1.1 Einführende Beschreibung	1
1.2 Motivation und Zielstellungen der Arbeit.....	8
1.2.1 Motivation	8
1.2.2 Zielstellungen.....	14
1.3 Methodischer Ansatz und Aufbau der Arbeit	21
1.3.1 Methodischer Ansatz.....	21
1.3.2 Struktureller Aufbau	25
1.3.2.1 Einleitende Beschreibungen	26
1.3.2.2 Inhalte des theoretischen Teils	28
1.3.2.3 Aufbau der empirischen Analysen	29
1.3.2.4 Gegenstand der Diskussion	29
1.4 Abgrenzung des Untersuchungsobjekts	30
1.4.1 Definitionen und Stand der Forschung	31
1.4.2 Einsatzgebiete der Sprachtechnologie	39
1.4.2.1 Call Center als Einsatzgebiet von Sprachdialogsystemen.....	44
1.4.2.2 Aufbau von Sprachdialogsystemen in Call Centern.....	52
1.4.2.3 Ausprägungen von Sprachdialogsystemen.....	52
1.4.2.4 Systemtechnische Spezifika von Sprachdialogsystemen	59
1.4.2.5 Technisch-prozessuale Rahmenbedingungen.....	61
1.4.3 Abgrenzung relevanter Stimuli aus technischer Sicht	86
2. Theoretischer Teil	90
2.1 Menschliche Aspekte	92
2.1.1 Aspekte der Sprachproduktion.....	96



2.1.2	Die Rezeption von Sprache.....	103
2.1.2.1	Kognitive Sprachverarbeitung.....	109
2.1.2.2	Das Sprachverstehen.....	112
2.1.3	Sprache und menschliches Verhalten	117
2.2	Maschinelle Aspekte (Künstliche Intelligenz)	125
2.2.1	Maschinelle Sprachausgabe	126
2.2.2	Maschinelle Spracheingabe	137
2.2.3	Maschinelles Verhalten.....	145
2.3	Applikationsbezogene Einflüsse	154
2.3.1	Ergonomie im Allgemeinen	155
2.3.2	Ergonomische Spezifika auditiver Benutzerschnittstellen	157
2.3.3	Das Nutzerverhalten.....	169
2.3.3.1	Akzeptanz als Prämisse für die Gebrauchsentscheidung	174
2.3.3.2	Einflussgrößen auf die Akzeptanz von Sprachportalen.....	175
2.3.3.3	Abgrenzung verwandter Konstrukte.....	183
2.4	Zusammenfassung der theoretischen Grundlagen.....	186
2.5	Konzeptualisierung der Variablen.....	189
2.5.1	Unabhängige Variablen.....	190
2.5.1.1	Variation der Darbietungsanzahl	190
2.5.1.2	Variation der Darbietungskomplexität	193
2.5.1.3	Variation der Darbietungsstruktur	198
2.5.1.4	Bewertete Nutzererfahrung.....	202
2.5.2	Abhängige Variablen	203
2.5.2.1	Behaltensleistung.....	204
2.5.2.2	Emotionaler Zustand.....	205
2.5.2.3	Adaption abh. Variablen aus der Akzeptanzforschung	206
2.5.2.4	Weitere Variablen.....	225
2.6	Ableitung von Hypothesen.....	226
3.	Empirischer Teil.....	249
3.1	Forschungsmethodik	249

II



3.2	Operationalisierung der Variablen	260
3.2.1	Operationalisierung unabhängiger Variablen	264
3.2.2	Eliminierung und Kontrolle von Drittvariablen.....	279
3.2.3	Operationalisierung abhängiger Variablen	295
3.3	Grundgesamtheit und Stichprobe	327
3.3.1	Die Grundgesamtheit	328
3.3.2	Die Stichprobe.....	330
3.4	Ablauf der Untersuchung	341
3.5	Material und Methoden	344
3.6	Darstellung der Ergebnisse zu den Haupthypothesen	354
3.6.1	Prozentualer Anteil korrekt wiedergegebener Items	354
3.6.2	Prozentualer Anteil falsch wiedergegebener Items	358
3.6.3	Prozentualer Anteil nicht wiedergegebener Items	364
3.6.4	Emotionaler Zustand	370
3.6.5	Nutzerzufriedenheit.....	373
3.6.6	Subjektiv wahrgenommene Nützlichkeit.....	376
3.6.7	Subjektiv wahrgenommene Nutzerfreundlichkeit	379
3.6.8	Einstellungsakzeptanz.....	384
3.6.9	Einfluss der Nutzererfahrung	389
3.7	Darstellung der Ergebnisse zu den Nebenhypothesen	394
3.7.1	Primacy-/Recency- und Wortlängeneffekt	395
3.7.2	Wahrnehmung der Musikuntermalung	400
3.7.3	Effekte auf die potenzielle Kundenorientierung	408
3.7.4	Effekte auf die potenzielle Kaufbereitschaft.....	412
4.	Diskussion	415
4.1	Interpretation der Hauptegebnisse	416
4.2	Interpretation weiterer Befunde.....	447
4.3	Implikationen auf die Praxis.....	461



V. Anhang.....	XIV
Anhang 1: Akzeptanzmodelle.....	XIV
Anhang 2: Grafische Darstellungen der Audiofile-Spezifikationen.....	XVII
Anhang 3: Fragebogen.....	XXI
Anhang 4: Übersicht der Operationalisierung abhängiger Variablen	XXIX
Anhang 5: Stat. Hypothesen/Hypothesenprüfung	XXXI
Anhang 6: Q-Q-Plots und Streudiagramme.....	XLI
VI. Literaturverzeichnis.....	LXVI

II. Abbildungsverzeichnis

Abbildung 1: Call Flow einer exemplarischen IVR-Lösung.	9
Abbildung 2: Einflussfaktoren auf die Akzeptanz von Sprachportalen.....	16
Abbildung 3: Methodisches Vorgehensmodell der Arbeit.	24
Abbildung 4: Struktureller Aufbau der Arbeit.	26
Abbildung 5: Abgrenzung des Begriffs “Sprachtechnologie“.	39
Abbildung 6: Perspektiven auf das Kunden-Kontakt-Management.	66
Abbildung 7: IVR Callflow-Schema Vertragskunden Vodafone.	73
Abbildung 8: Warteschleifenmodell.	84
Abbildung 9: Prozessmodell zur Darstellung systembezogener Stimuli.	88
Abbildung 10: Struktur des theoretischen Teils der Arbeit.	90
Abbildung 11: Serieller Prozess zur Erzeugung von Sprache.	99
Abbildung 12: Sprechersystem und Sprachapparat.	103
Abbildung 13: Aufbau des menschlichen Ohrs.	105
Abbildung 14: Prozess der kognitiven Sprachverarbeitung.....	116
Abbildung 15: Menüstruktur einer IVR mit serieller Dialogführung.....	128
Abbildung 16: Abläufe von Sprachsyntheseprozessen.....	134
Abbildung 17: Basisarchitektur der systemseitigen Sprachverarbeitung.	141
Abbildung 18: Differenzierung des Nutzerverhaltens.	173
Abbildung 19: Das serielle Mehrspeichermodell.....	177
Abbildung 20: Cognitive Load bei der Nutzung auditiver Schnittstellen.....	182
Abbildung 21: Unzureichende Trennschärfe von Konstrukt-Definitionen.	185
Abbildung 22: Varianten des Chunkings.	199
Abbildung 23: Technical Acceptance Model.....	212
Abbildung 24: Unified Theory of Acceptance and Use of Technology.	214
Abbildung 25: Theoretisches Modell für die empirische Studie.	224
Abbildung 26: Grafische Darstellung des Audiofiles "Vertragshandys".....	269
Abbildung 27: Spezifikation Audiofile mit 5 Items und ber. Musik.	273
Abbildung 28: Spezifikation Audiofile mit 15 Items und aktiv. Musik.	274
Abbildung 29: Methoden zur Kontrolle von Störvariablen.	280
Abbildung 30: Dimension "Valenz" des Self Assessment Manikin.	308
Abbildung 31: Altersverteilung (absolut) der Grundgesamtheit.....	330

Abbildung 32: Altersverteilung (absolut) der Stichprobe.	333
Abbildung 33: Häufigkeitsverteilung Kontakte p.a.	335
Abbildung 34: Häufigkeitsverteilung der Erfahrungsbewertung	335
Abbildung 35: Ablauf der Untersuchung.	343
Abbildung 36: Boxplot-Grafik der korrekt wiedergeg. Items (in %).	357
Abbildung 37: Boxplot-Grafik der falsch wiedergeg. Items (in %).	360
Abbildung 38: Boxplot-Grafik zu nicht wiedergeg. Items (in %).	365
Abbildung 39: Boxplot-Grafik zum emotionalen Zustand.	372
Abbildung 40: Boxplot-Grafik zur Nutzerzufriedenheit.	375
Abbildung 41: Boxplot-Grafik zur subj. wahrgen. Nützlichkeit.	378
Abbildung 42: Boxplot-Grafik zur subj. wahrgen. Nutzerfreundlichkeit.	382
Abbildung 43: Antwortverteilung zur Einstellungsakzeptanz (in %).	388
Abbildung 44: Serielle Positionskurve.	397
Abbildung 45: Grafische Analyse Primacy-Recency-Effekt 5 Items.	398
Abbildung 46: Grafische Analyse Primacy-Recency-Effekt 15 Items.	400
Abbildung 47: Verteilung der Musikbekanntheit.	402
Abbildung 48: Konsolidierte Darstellung des Musikgefallens M1, M2.	405
Abbildung 49: Konsolidierte Darstellung der Musikstörung M1, M2.	407
Abbildung 50: Boxplot-Grafik zur pot. Kundenorientierung.	410
Abbildung 51: Boxplot-Grafik zur pot. Kaufbereitschaft.	413



III. Tabellenverzeichnis

Tabelle 1: Vergleich IVR vs. Sprachdialogsystem.....	13
Tabelle 2: Forschungsfragen- und Zielmatrix.	20
Tabelle 3: Abgrenzung des Begriffs "Sprachtechnologie".....	35
Tabelle 4: Exemplarische Übersicht natursprachlicher Anwendungen.	42
Tabelle 5: Tarifübersicht von Sonderrufnummern-Gassen.	50
Tabelle 6: Vergleich von IVR-Leistungsgrößen nach Automatisierungsgrad. ...	58
Tabelle 7: Exemplarische CTI-Routing-Tabelle.	69
Tabelle 8: Teilkompetenzen auditiver Wahrnehmung.	109
Tabelle 9: Sprach-, kognitions-, und verhaltenswiss. Aspekte.....	124
Tabelle 10: Menschliche Sprachproduktion vs. masch. Sprachausgabe.	136
Tabelle 11: Menschliche Sprachrezeption vs. masch. Sprachverarbeitung.	144
Tabelle 12: Menschliches- versus maschinelles Verhalten.	153
Tabelle 13: DIN-Normen zur Mensch-System-Interaktion.....	156
Tabelle 14: DIN-Normen zur Softwareergonomie.....	157
Tabelle 15: Aufgabenangemessenheit.	160
Tabelle 16: Selbstbeschreibungsfähigkeit.	161
Tabelle 17: Steuerbarkeit.....	162
Tabelle 18: Erwartungskonformität und Fehlertoleranz.....	163
Tabelle 19: Individualisierbarkeit, Lernförderlichkeit und KO-Ziel.....	164
Tabelle 20: Wahrnehmung und Verständnis.	165
Tabelle 21: Exploration, Benutzungsmotivation und Hear & Feel.	166
Tabelle 22: Darstellung unabhängiger Variablen.	202
Tabelle 23: Bewertung der Akzeptanzmodelle	221
Tabelle 24: Stat. Hypothesen zum Darbietungsumfang.	228
Tabelle 25: Darbietungskomplexität: Hypothesen zu korr. wiedergeg. Items.	229
Tabelle 26: Darbietungskomplexität: Hypothesen zu falsch wiedergeg. Items.....	230
Tabelle 27: Darbietungskomplexität: Hypothesen zu nicht wiedergeg. Items.....	231
Tabelle 28: Darbietungskomplexität: Hypothesen zum emot. Zustand.	232
Tabelle 29: Darbietungskomplexität: Hypothesen zur Nutzerzufriedenheit.	233
Tabelle 30: Darbietungskomplexität: Hyp. zur subj. wahrgen. Nützlichkeit.	234
Tabelle 31: Darb.-Komplexität: Hyp. zur subj. wahrgen. Nutzerfreundlichk...	235



Tabelle 32: Darbietungskomplexität: Hypothesen zur Einstellungsakzeptanz.	236
Tabelle 33: Darbietungsstruktur: Hypothesen zur Behaltensleistung.	238
Tabelle 34: Darbietungsstruktur: Hypothesen zum emot. Zustand.	238
Tabelle 35: Darbietungsstruktur: Hypothesen zur Zufriedenheit.	239
Tabelle 36: Darbietungsstruktur: Hyp. zur subj. wahrgen. Nützlichk.	240
Tabelle 37: Darbietungsstruktur: Hyp. zur subj. wahrgen. Nutzerfreundlichk.	240
Tabelle 38: Darbietungsstruktur: Hypothesen zur Einstellungsakzeptanz.	241
Tabelle 39: Hypothesen zur Erfahrung.	243
Tabelle 40: Hypothesen zur pot. Kundenorientierung.	246
Tabelle 41: Hypothesen zur pot. Kaufbereitschaft.	248
Tabelle 42: Darbietungsanzahl in Höhe von 5 Items.	266
Tabelle 43: Realdefinition von Terminologien.	267
Tabelle 44: Darstellung der Items 6 bis 15.	270
Tabelle 45: Realdefinition weiterer Terminologien.	271
Tabelle 46: Zuordnung von Items zu Chunks.	275
Tabelle 47: Merkmalsverteilung auf die Experimentalgruppen.	337
Tabelle 48: Verteilung Nutzererfahrungen auf die Experimentalgruppen.	340
Tabelle 49: Deskriptive Statistik korrekt wiedergeg. Items (in %).	355
Tabelle 50: Ergebnisse Post-Hoc-Tests korrekt wiedergeg. Items (in %).	358
Tabelle 51: Deskriptive Statistik falsch wiedergeg. Items (in %).	359
Tabelle 52: Ergebnisse Post-Hoc-Tests falsch wiedergeg. Items (in %).	362
Tabelle 53: Deskriptive Statistik nicht wiedergeg. Items (in %).	364
Tabelle 54: Ergebnisse Post-Hoc-Tests nicht wiedergeg. Items (in %).	367
Tabelle 55: Übersicht korrekt, falsch und nicht wiedergeg. Items.	369
Tabelle 56: Deskriptive Statistik zu emotionalen Zuständen.	371
Tabelle 57: Deskriptive Statistik zur Nutzerzufriedenheit.	374
Tabelle 58: Deskriptive Statistik zur subj. wahrgen. Nützlichkeit.	377
Tabelle 59: Ergebnisse Post-Hoc-Test zur subj. wahrgen. Nützlichkeit.	379
Tabelle 60: Trennschärfe der Items zur Nutzerfreundlichkeit.	380
Tabelle 61: Deskriptive Statistik zur subj. wahrgen. Nutzerfreundlichkeit.	381
Tabelle 62: Erg. Post-Hoc-Tests zur subj. wahrgen. Nutzerfreundlichkeit.	383
Tabelle 63: Deskriptive Statistik zur Einstellungsakzeptanz.	385



Tabelle 64: Bivariate Korr. der Erfahrung zu metrisch skalierten Variablen. ...	390
Tabelle 65: Punktbiserial Korr. "Nutzererf. vs. Einstellungsakzeptanz".	392
Tabelle 66: Häufigkeitsverteilung bei 5 Items.	395
Tabelle 67: Häufigkeitsverteilung bei 15 Items.	399
Tabelle 68: Musikdarbietung versus Musikwahrnehmung.	401
Tabelle 69: Häufigkeitsverteilung Musikgefallen (in %).	404
Tabelle 70: Häufigkeitsverteilung Musikstörung (in%).	406
Tabelle 71: Deskriptive Statistiken zur pot. Kundenorientierung.	409
Tabelle 72: Ergebnisse Post-Hoc-Test zur pot. Kundenorientierung.	411
Tabelle 73: Deskriptive Statistiken zur pot. Kaufbereitschaft.	412
Tabelle 74: Ergebnisse Post-Hoc-Test zur pot. Kaufbereitschaft.	414
Tabelle 75: Kernaussagen zu den Hauptanalysen.	460
Tabelle 76: Kernaussagen zu Nebenanalysen.	461



IV. Abkürzungsverzeichnis

ACD	Automated Call Distribution
ANOVA	Analysis of Variance
AT&T	American Telephone and Telegraph Company
AttrakDiff	Mischwort aus Attraktivität und semantischem Differential
Audi	Auto Union Deutschland Ingolstadt
AV	Abhängige Variable
Bit	Binary Digit
BMW	Bayrische Motorenwerke
bpm	Beats per Minute
c.p.	ceteris paribus
C++	Programiersprache (Eigenname)
CLIR	Calling Line Identification Restriction
CLIRSO	Calling Line Identification Restriction Override
C-OAR-SE	Construct Definition-Object Classification
.....	Attribute Classification-Rater Identification-Scale Information-
.....	Enumeration and Reportings
CRM	Customer Relationship Management
CTI.....	Computer Telephony Integration
CTS.....	Concept To Speech
d	Maß für Effektstärke (auch Cohens d)
DFKI.....	Deutsches Forschungszentrum für Künstliche Intelligenz
DIN	Deutsches Institut für Normung
DMS.....	Dialogue Management System
DTMF	Dual Tone Multiple Frequency
DTPB	Decomposed Theory of Planned Behavior
E[Lq.....	Durchschnittliche Verweildauer Warteschleife
E[W] [Sek.].....	Durchschnittliche Wartezeit
EN	Europäische Norm



FAQ	Frequently Asked Questions
Fax	Telefaksimile
FCFS.....	First Come First Serve
FIFO.....	First In First Out
fMRT	funktionelle Magnetresonanztomografie
GmbH	Gesellschaft mit beschränkter Haftung
GSM.....	Global System for Mobile Communications
GUI	Graphical User Interface
HDK.....	Hypothese zur Darbietungskomplexität
HDS	Hypothese zur Darbietungsstruktur
HDU.....	Hypothese zum Darbietungsumfang
HE	Hypothese zur Nutzererfahrung
HLT	Human Language Technology
HTML.....	Hypertext Markup Language
http.....	Hypertext Transfer Protocol
Hz.....	Herz
IAO	Institut für Arbeitswirtschaft und Organisation
IBM.....	International Business Machine
ICL.....	Intrinsic Cognitive Load
IMUT	Integrated Model of User Satisfaction and Technology Acceptance
ISDN.....	Integriertes Sprach- und Datennetz
ISO.....	International Organization for Standardization
IT	Informationstechnologie
Kfz	Kraftfahrzeug
KI.....	Künstliche Intelligenz
LDV	Linguistische Datenverarbeitung
LS.....	Levene Statistik
LSD.....	Least Significant Difference
M	Mittelwert
M1.....	Beruhigendes Musikstück
M2.....	Aktivierendes Musikstück
Max.	Maximum



MD.....	Median
Min.....	Minimum
MIS	Management Information System
MSC.....	Mobile Switching Center
n	Anzahl Elemente einer Stichprobe
N	Grundgesamtheit
NH	Nebenhypothese
NLP.....	Natural Language Processing
o.D.	ohne Datumsangabe
o.S.	ohne Seitenangabe
o.V.	ohne Verfasserangabe
PC	Personal Computer
PET	Positronen-Emissions-Tomografie
PIN.....	Personal Identification Number
PR	Public Relations
p-Wert.....	Signifikanzwert
Q-Q	Quartil-Quartil
r	Korrelation
R-E-G-W	Ryan, Einot, Gabriel, Welsch
SAM.....	Self Assessment Manikin
SD	Standardabweichung
SERVIMPERF.....	Service Importance and Performance
SERVPERF	Serviceperformance
SERVQUAL.....	Servicequalität
SIM	Subscriber Identity Module
S-N-K.....	Studentized Newman Keuls
S-O-R.....	Stimulus-Organismus-Reaktion
SPSS	Statistical Packages for the Social Sciences
SUS.....	System Usability Scale
TAM	Technology Acceptance Model
TK.....	Telekommunikation
TPB.....	Theory of Planned Behavior



TRA	Theory of Reasoned Action
TTS	Text To Speech
TV	Television
USE	Usefulness, Satisfaction, Ease of Use
UTAUT	Unified Theory of Acceptance and Use of Technology
V	Effektstärke-Maß (auch Cramers V)
VDI	Verein Deutscher Ingenieure
Vgl	Vergleiche
VIP	Very Important Person
Vodafone	Voice Data Fax Over Network
VUI	Voice User Interface
VXML	Voice Extensible Markup Language
W3C	World Wide Web Consortium
WLAN	Wireless Large Area Network
www	World Wide Web
XML	Extensible Markup Language
α	Alpha
η^2	Eta-Quadrat
λ	Lambda oder Lamda
μ	Micro
$1/\mu$ [Min.]	Formel für Gesprächsdauer
Z_0X	Forschungsziel "X"



1. Einleitender Teil

Im ersten Teil dieser Arbeit wird zunächst die Themenstellung inhaltlich einführend beschrieben. Darüber hinaus werden grundsätzliche Zusammenhänge erläutert sowie auf die Motivation und die Zielstellung der vorliegenden Arbeit eingegangen. Ferner werden der methodische Ansatz und der Aufbau der Arbeit dargestellt. Da es sich um ein interdisziplinäres Themenkonstrukt handelt, wird der klaren Abgrenzung des Untersuchungsgegenstands entsprechende Aufmerksamkeit gewidmet.

1.1 Einführende Beschreibung

Maschinelle Sprachverarbeitung ist Gegenstand von Forschungen im Bereich der Künstlichen Intelligenz und bereits in vielen Situationen des Alltags vorzufinden. Audio- und Navigationssysteme in Fahrzeugen, verschiedene Mobiltelefon- und TV-Modelle sowie PC-Softwareanwendungen und Smart-Home-Applikationen lassen sich heute per natürlicher Spracheingabe bedienen. Auch so genannte Sprachportale¹, die oftmals Call Centern als kundenbetreuende Organisationseinheiten des Marketings vorgeschaltet sind, lassen sich heutzutage zum Teil bereits per Sprache steuern. Moderne Technologien und Standards ermöglichen in diesen und weiteren Anwendungsbereichen mittlerweile den Einsatz von Sprachapplikationen.²

¹ Die Begriffe Anrufer- oder Sprachportal sowie Interactive Voice Response (IVR), werden in der Praxis häufig synonym genutzt. Zur genaueren Abgrenzung, auch zu Chatbots, s. Abschnitt 1.4.2.3.

² Eine Sprachapplikation ist ein Computerprogramm, welches eine für den Anwender nützliche Funktion ausführt und per Spracheingabe bedient wird. Vgl. hierzu o.V.: Clarity-Pilot, München 2002, S.1.

Aus technischer Sicht kann, neben der Ablösung des analogen Telefonnetzes durch digitale Technologien, die Einführung der sogenannten Voice Extensible Markup Language (VXML)³, als Meilenstein bezeichnet werden.

Ältere, so genannte “Interactive Voice Responses“ (IVR) sind im Vergleich zu modernen Sprachportalen dadurch gekennzeichnet, dass sich ein Anrufer durch Drücken von Tasten auf der Telefontastatur oder durch die Eingabe rudimentärer Sprachbefehle durch das Menü navigiert. Heutzutage sind hingegen natursprachliche Dialoge bei der Interaktion zwischen Mensch und Maschine möglich. Insbesondere in Call Centern⁴ finden Sprachportale mittlerweile zunehmend Anwendung, deren Menüfunktionen via natursprachlichem Dialog gesteuert werden können. Vor allem Unternehmen, die in Märkten mit großer Wettbewerbsintensität agieren, installieren in ihren Call Centern, an der -auch und insbesondere aus Marketing-Sicht- äußerst wichtigen Schnittstelle zum Kunden, sowohl bei eingehenden als teils auch bei ausgehenden Telefonkontakten, sprachbasierte Technologien.

Häufig beinhalten diese Anruferportale neben der Funktion, Anliegen thematisch vorzustrukturieren und einen zeitlichen Puffer zwischen der Vielzahl an Anrufen und den begrenzten Ressourcen an Call Center-Agenten zu bilden, auch sogenannte Self Care-Applikationen.⁵

Diese ermöglichen es dem Kunden, bestimmte Vorgänge im Portal zu erledigen, ohne dabei mit einem Menschen persönlich sprechen zu müssen.

³ VXML, respektive VoiceXML, ist eine XML-Applikation und setzte sich Anfang dieses Jahrtausends gegen die Programmiersprache C++ durch, die bis dahin für die im Vergleich zu Sprachportalen deutlich rudimentäreren Interactive Voice Responses (IVR) genutzt wurde. XML bedeutet “Extensible Markup Language“ und bedeutet “Auszeichnungssprache“. Vgl. Harris, R. A.: Voice Interaction Design (2005) S. 549.

⁴ Als Call Center kann sowohl eine Organisationseinheit innerhalb eines Unternehmens verstanden werden, als auch ein externes Unternehmen, welches sich auf Call Center-Dienstleistungen spezialisiert hat. Zur weiterführenden Definition siehe Abschnitt 1.4.2.1.

⁵ “Self Care“ bedeutet so viel wie “Selbstbetreuung“. Kunden können, ähnlich wie auch im Internet, in Sprachportalen z.B. Flüge buchen, Verträge wechseln, Produkte bestellen und Dienste in Anspruch nehmen, sprich bindende Rechtsgeschäfte, mit speziellen Regelungen zu den Rückgaberechten, völlig autark und ohne persönlichen Kontakt mit einem Agenten eingehen.

Diese Automatisierung von Prozessen, bzw. zumindest die Teilautomatisierung von Prozessschritten, soll nach Schneider/Emde/Rapp dazu führen, dass der kostenintensive, unmittelbare Kontakt zwischen Kunde und Kundenbetreuer vermieden oder zumindest auf ein mögliches Minimum reduziert und somit die Wirtschaftlichkeit dieses Kommunikationskanals optimiert wird.⁶

Firmen, die Sprachportale einsetzen, betreuen nach Engelbach/Meier nicht selten mehrere Millionen Kunden mit einer Anzahl an eingehenden Telefonkontakten in sechs- bis achtstelliger Höhe pro Jahr.⁷

Hierdurch kann sich ein Spannungsfeld ergeben. Unternehmen sind sich durchaus der immensen Bedeutung dieser Schnittstelle im Sinne verschiedener Aspekte des Marketing-Mix bewusst. Hervorzuheben sei hier die Kommunikation zwischen dem Unternehmen und wichtigen Kunden, bzw. Interessenten, auch und insbesondere vor dem Hintergrund der Neukundenakquise, der Bindung bestehender und der Rückgewinnung gekündigter Kunden sowie auch von Cross- und Upselling-Aktivitäten.

“In vielen Unternehmen wird die Bedeutung der Kundenbetreuung (Customer Care) als strategisches Marketinginstrument und Unterscheidungsmerkmal zum Wettbewerber erkannt und eine Verbesserung des Kundenservice angestrebt.“⁸

⁶ Vgl. Schneider, D. Emde, P., Rapp, O.: Call Center, in: Grewenig, I.: Multimedia- und Call Center (2002), S. 10.

⁷ Vgl. Engelbach, W., Meier, R.: Customer Care Management (2001), S. 20. Die Kontaktfrequenz von Call Center im Bereich eingehender Anrufe variiert stark in Abhängigkeit von der Branche und dem Erklärungsbedarf der zugrundeliegenden Produkte. Ein Versicherungskunde ruft bei einer Laufzeit eines Lebensversicherungsvertrages von 12 Jahren durchschnittlich fünfmal, ein Mobilfunkkunde bei 2-jähriger Vertragslaufzeit etwa genauso oft an. Zur Kontaktfrequenz s. auch: Menzler-Trott, E., Kundenorientierung in Menzler-Trott, E., Call Center Management (1999), S. 12, oder auch o.V.: Vocatus Syndikatsstudie, Servicequalität (2004), S. 25 oder Peissner, M., Sell, D., Steimel, B.: Akzeptanz von Sprachapplikationen (2006), S. 9.

⁸ Schnorrenberg, H.: Call Center-Technik, in: Menzler-Trott, E., Hanel, M.: Call Center Evolution (2001), S. 321.

Dieser durchaus etwas allgemein gehaltenen Aussage von Schnorrenberg steht jedoch der Umstand gegenüber, dass in zahlreichen Branchen, darunter in der oligopolistisch geprägten Telekommunikationsbranche, die Marktphase des Verdrängungswettbewerbs mit einhergehenden, sinkenden Preisen und Margen vorherrscht.⁹

Hieraus ergibt sich für die Unternehmen ein erheblicher Druck, Kosten einzusparen. Werden die Kosten im Rahmen einer groben Überschlagsrechnung bei einem Volumen von mehreren Millionen eingehenden Anrufen pro Jahr und marktüblichen Durchschnittskosten von 3,86 EUR¹⁰ pro Anruf multipliziert, wird deutlich, dass das Augenmerk der Unternehmen schnell auf die Realisierung von Kosteneinsparpotenzialen durch Prozessautomatisierungen in diesem Bereich fallen kann. Unternehmen rechtfertigen ihre Entscheidung, Sprachportale einzusetzen, neben den wirtschaftlichen Aspekten der Prozessautomatisierung, mit der Argumentation, dass der Anrufer bestimmte Standardanliegen bereits im Portal selbst erledigen kann. Somit würden Wartezeiten verkürzt, die sich ergäben, wenn der Anrufer darauf warten würde, dass er persönlich mit einem Kundenbetreuer sprechen kann.¹¹

Dies suggeriert, dass die Akzeptanz des Nutzers im Umgang mit Sprachportalen durchaus recht hoch sein könnte, da der Anrufer durchschnittlich Zeit spart. Studien des Fraunhofer Instituts (Arbeitswirtschaft und Organisation) haben bereits im Jahr 2006 ergeben, dass 70% von über 1000 interviewten Probanden, Sprachapplikationen kennen und nutzen. Im Jahr 2003 waren es noch 54%. Tatsächlich geben Nutzer im Rahmen der angeführten Studie als wichtigsten Vorteil die gewonnene Zeitersparnis an.

⁹ Zur Definition und zu den Eigenschaften oligopolistischer Märkte vgl. Fischer, K.: Oligopolistische Marktprozesse (1988), S. 36 sowie Zohnhöfer, W.: Wettbewerbspolitik im Oligopol (1968), S. 18,19 und Frambach, H.: Crashkurs Mikroökonomik (2008), S. 184.

¹⁰ Vgl. o.V.: Einkauf von Call Center Lösungen (2008). Ergebnisse aus der Datenerhebung vom Bundesverband Materialwirtschaft, Einkauf und Logistik e.V., S. 20.

¹¹ Zur Berechnung der durchschnittlichen Wartezeiten von Anrufern bei Hotlines ohne Sprachportal im Vergleich zu Wartezeiten bei Hotlines mit Sprachportal vgl. Helber, S., Stollatz, R.: Call Center Management (2004), S. 96 sowie in der vorliegenden Arbeit unter Abschnitt 1.4.2.

Davon teilten 29% der Befragten mit, telefonbasierte Sprachapplikationen sogar bereits mehrfach genutzt zu haben.¹²

Dennoch ist fraglich, ob Interessenten und Kunden die Nutzung von telefonbasierten Sprachapplikationen ihrerseits als grundsätzliche Verbesserung des Kundenservices ansehen. So geben bei der vorbenannten Befragung des Fraunhofer Instituts aus 2003 26% der Probanden, die ausgesagt haben, telefonbasierte Sprachapplikationen zu kennen und zu nutzen, an, dennoch sehr unzufrieden mit der am häufigsten genutzten Sprachapplikation zu sein.¹³

Im Gegensatz zum erhöhten Bekanntheitsgrad von 2006 gegenüber 2003, hat sich die Zufriedenheit mit der jeweils am meisten genutzten Sprachapplikation nicht geändert. Auch im Jahre 2006 sind 26% der Befragten sehr unzufrieden mit den genutzten Sprachapplikationen. Als Faktoren für eine erhöhte Akzeptanz im Umgang mit der genutzten Sprachapplikation gaben die Probanden an, dass diese übersichtlicher, komfortabler, verständlicher und weniger fehleranfällig sein müsste.¹⁴

Ist ein Anrufer mit der Nutzung einer bestimmten Sprachapplikation unzufrieden, kann nach den Ausführungen von Giering/Hentschel/Homburg konstatiert werden, dass die Erwartungen des Nutzers an die Sprachapplikation nicht erfüllt sind und daraus gegebenenfalls eine geringere Nutzerakzeptanz¹⁵ folgt.¹⁶

¹² Vgl. Biesterfeld, J., Heidmann, F., Peissner, M.: Akzeptanz und Usability von Sprachapplikationen in Deutschland (2004), S. 14.

¹³ Ebd.

¹⁴ Vgl. Biesterfeld, J., Heidmann, F., Peissner, M.: Akzeptanz und Usability von Sprachapplikationen (2004), S. 17 sowie Peissner, M., Sell, D., Steimel, B.: Akzeptanz von Sprachapplikationen (2006), S. 13.

¹⁵ Zum Konstrukt der Nutzerakzeptanz siehe Abschnitt 2.3.3.1.

¹⁶ Vgl. zum Konfirmations-/ Diskonfirmationsparadigma zur Beschreibung der Entstehung von Zufriedenheit und Unzufriedenheit von Giering, A., Hentschel, F., Homburg, C.: Kundenbindung und Kundenzufriedenheit, in: Bruhn, M., Homburg, C.: Kundenbindungsmanagement (1998), S. 84.

Wie eine Studie des Marktforschungs-Instituts Vocatus AG zur Servicequalität zeigt, ist ebenfalls der Umstand in Betracht zu ziehen, dass Interessenten und Kunden den telefonischen Kontaktweg zum Unternehmen im Vergleich zu anderen Kontaktkanälen, wie z.B. Internet/Social Media, E-Mail, Brief oder Fax bevorzugt wählen, um ihr Anliegen direkt über den persönlichen Kontakt mit einem Call-Center-Agenten lösen zu können.¹⁷

Im Zuge der Literaturoswertung zum Themenkonstrukt “Sprachportal“ entsteht, nicht selten gestützt durch mündliche Berichte aus der Praxis von Sprachportal-Nutzern, schnell der Eindruck, dass Sprachportale vorrangig als Ärgernis aufgefasst werden, die dem Ziel, ein Anliegen über den persönlichen, telefonischen Kontakt mit einem Agenten zu erledigen, störend entgegenwirken. Dieser Eindruck muss kritisch hinterfragt werden. So erscheint die erwähnte Studie des Fraunhofer Instituts hinsichtlich der Abgrenzung des dieser Studie zugrundeliegenden Untersuchungsgegenstands recht oberflächlich. Als Untersuchungsobjekt wird eher allgemein vom Konstrukt der Sprachapplikation geschrieben. Aus der Studie geht nicht hervor, um welche Ausprägungen an Sprachportalen es sich im Einzelfall konkret handelt. Grundsätzlich darf im Sinne einer Praxisannahme vor diesem Hintergrund vermutet werden, dass ein Sprachportal, welches primär der Anrufer-Vorqualifizierung sowie der Informationsvergabe dient und ggf. einen zu hohen Umfang an Menüpunkten und Inhalten sowie ein wenig gut durchdachtes Dialogdesign aufweist, auf Basis einer veralteten Technik sowie noch per Tastendruck bedient werden muss, schlechter bewertet wird, als ein kundenzentriertes, natursprachlich und einfach zu bedienendes System mit sinnvollen Self-Service-Applikationen.

¹⁷ Vgl. o.V.: Vocatus Syndikatsstudie, Servicequalität (2004), S. 15.

Den Ergebnissen einer Studie von Meuter/Ostrom/Roundtree, et al., zur Nutzung von Self Services innerhalb des Kanals Telefonie/Interactive Voice Response zufolge, können die Probanden den genutzten Services durchaus greifbare Vorteile abgewinnen, wie z.B. die Ersparnis von Zeit und Geld, die zeitliche und örtliche Unabhängigkeit, eine einfache Handhabung sowie die Vermeidung eines teilweise offensichtlich auch negativ empfundenen persönlichen Kontakts zu einem anderen Menschen.¹⁸

Als störend werden hier jedoch ebenfalls schlechte Service- oder Technology Designs¹⁹, bzw. technische und prozessuale Probleme genannt.²⁰

Insbesondere diese übergreifende Erwartungshaltung von Interessenten und Kunden beim Kontakt mit einem Unternehmen, Fragen bzw. Anliegen im persönlichen Kontakt mit einem Ansprechpartner unmittelbar klären zu wollen, bleibt vermutlich unerfüllt, wenn sich der Nutzer zunächst mit komplexen Menüstrukturen eines Anruferportals und anschließenden Warteschleifen auseinandersetzen muss.

„We point out, as have many others, that interactions between employees and customers are social exchanges. We further suggest that technology creates a barrier that can inhibit this social exchange.“²¹

¹⁸ Vgl. Meuter, M., Ostrom, A. L., Roundtree, et al.: Self Service Technologies, in: Journal of Marketing (2000), S. 50-64. Den Ergebnissen dieser Studie ist zu entnehmen, dass eine Self Care Applikation gemeint ist, über die via Telefon/Sprachportal Produkte eingekauft werden können und der Nutzer sich hier den situativ ggf. nicht gewünschten vertriebsorientierten Beratungsleistungen eines Verkäufers/einer Verkäuferin entziehen kann. Dieser Anwendungsfall weicht daher davon ab, dass ein Sprachportalnutzer über den Telefonweg ein Anliegen zu seinem Vertrag oder Produkt hat und hierzu den persönlichen Kontakt zu einem Agenten bevorzugt.

¹⁹ Schlechtes Servicedesign könnte z.B. vorliegen, wenn in einer Banking-Applikation die tägliche Summe, die überwiesen werden kann, limitiert ist, der Nutzer darüber aber nicht im Vorhinein informiert wird, sondern erst am Ende des Prozesses, wenn er bereits alle Angaben eingegeben hat. Mit Problemen im Technology-Design ist hier z.B. gemeint, dass systemseitig nicht verhindert wird, dass man ein Zusatzprodukt oder einen Service kauft, der mit dem bereits gekauften Grundprodukt nicht kompatibel ist.

²⁰ Vgl. Meuter, M., Ostrom, A. L., Roundtree, et al.: Self Service Technologies, in: Journal of Marketing (2000), S. 50-64.

²¹ Giebelhausen, M., Robinson, S. G., Sirianni, N. J. et al.: Touch versus Tech, in: Journal of Marketing (2014), S. 116.

Diese einleitende Beschreibung macht deutlich, dass Sprachportale in einem Spannungsfeld eingesetzt werden, welches einerseits primär aus unternehmerischem Kostendruck motiviert ist und der Senkung von Kommunikations- und Prozesskosten dienen soll. Andererseits werden seitens der Unternehmen etwaige Unzufriedenheiten mit der Servicequalität sowie geringe Akzeptanzwerte bezüglich der Nutzung von Sprachportalen durch wertvolle Kunden und Interessenten scheinbar in Kauf genommen, bzw. u.a. mit einer für den Anrufer vorteilhaften Verkürzung der Durchlaufzeiten begründet. Das hieraus wiederum resultierende Verhalten der Kunden und die daraus etwaig abzuleitenden, ökonomischen Implikationen für Unternehmen, spielen offensichtlich häufig eine untergeordnete Rolle in der wirtschaftlichen Gesamtbetrachtung im Sinne einer Entscheidungsgrundlage, Prozesse über ein Sprachportal zu automatisieren.

1.2 Motivation und Zielstellungen der Arbeit

In diesem Abschnitt wird die aus der Literatur sowie auch aus den berichteten Praxiserfahrungen abgeleitete Motivation zur Bearbeitung dieses Themas erläutert. Ferner werden klare Zielstellungen für diese Arbeit formuliert.

Letztere leiten sich aus der Beantwortung der im Rahmen dieser Arbeit aufgeworfenen Forschungsfragen im theoretischen und empirischen Teil ab.

1.2.1 Motivation

Wird die technische Entwicklung von IVR-Lösungen mit rudimentärer Sprachsteuerung (z.B. für Adressänderungen sagen Sie: „Adresse!“), respektive mit Steuerung per Tastendruck (z.B. „für Adressänderungen drücken Sie die Drei!“) und vergleichsweise starren Menüstrukturen bis hin zu modernen Applikationen, die per natursprachlichem Dialog gesteuert werden, betrachtet, fällt auf, dass sich damit einhergehend die Bedienung von einem systemgesteuerten zu einem nutzergesteuerten Dialog verändert hat.

Wie die folgende Abbildung exemplarisch zeigt, müssen sich Nutzer älterer IVR-Lösungen teils durch weit verzweigte Menübäume navigieren.

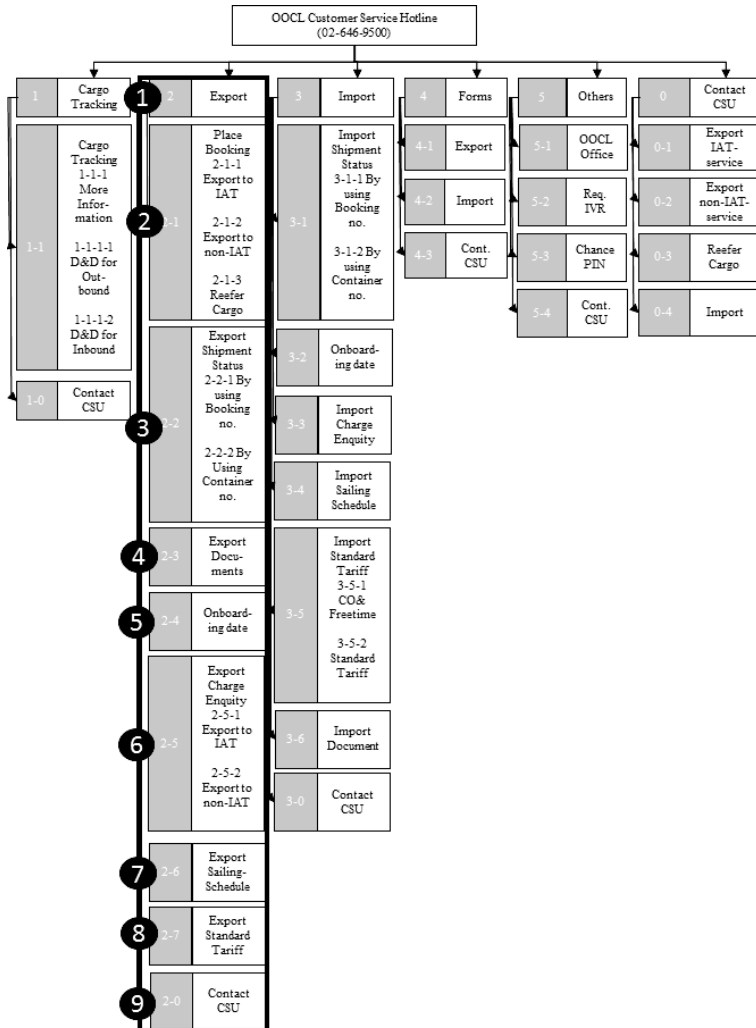


Abbildung 1: Call Flow einer exemplarischen IVR-Lösung.²²

²² Eigene Darstellung in Anlehnung an <http://www.oocl.com/thailand/eng/localinformation/localnews/2008/Pages/IVR3Flow.aspx> (zuletzt aufgerufen am 05.07.2015).

Möchte ein Kunde einen Agenten der Customer Service Hotline dieses Anbieters erstmalig zum Thema “Export“ sprechen, müsste der Nutzer beispielsweise neun²³ Menüpunkte durchlaufen, um in persönlichen Kontakt mit einem Agenten zum Thema “Export“ zu kommen. Etwaige Begrüßungsansagen und weitere Informationen zu den Navigationsoptionen des Systems, kämen noch hinzu. Würde der Anbieter stattdessen ein auf den neuesten Technologien basierendes natursprachliches Dialogsystem anbieten, wäre stattdessen ein exemplarischer Dialog wie folgt vorstellbar:

System: „Guten Tag. Was können wir für Sie tun?“

Nutzer: „Ich möchte die Hotline zum Thema “Export“ sprechen.“

System: „Vielen Dank, wir verbinden Sie mit dem nächst verfügbaren Mitarbeiter.“

Im Sinne einer Praxisannahme wird konstatiert, dass ein kurzer, natursprachlicher Dialog den Nutzer im Sinne einer höheren Benutzerfreundlichkeit und Nützlichkeit²⁴ effizienter und effektiver an sein Ziel bringt. Auf Basis der bereits einleitend beschriebenen, vergleichsweise schlechten Zufriedenheitswerte im Rahmen von Befragungen von IVR-Nutzern, konnten in den vergangenen Jahren Paradigmenwechsel bei den Servicestrategien von Unternehmen, die solche IVRs einsetzen, beobachtet werden. In Einzelfällen haben Unternehmen, wie z.B. die Firma Orange aus der Telekommunikationsbranche, sogar vollständig auf den Einsatz von IVRs verzichtet.²⁵

²³ Punkt 2 auf der Hauptmenüebene sowie dann acht weitere Untermenüpunkte.

²⁴ Zu den Konstrukten der Nutzerfreundlichkeit und der Nützlichkeit siehe auch Abschnitt 2.5.2 sowie Abschnitt 3.2.3.

²⁵ Vgl. Hasler, K.: Interview mit Michael Klötzli, Director Enterprise Service bei Orange, unter: www.cmm-magazine.ch vom 1.9.2013 (zuletzt aufgerufen am 22.08.2015).

„Noch immer setzen viele Unternehmen aus Kostengründen auf IVR Lösungen. Sie sollen helfen, Kundenanfragen effizienter zu bearbeiten und Kosten zu sparen. Dies gelingt jedoch nicht im erwarteten Maß. Eindeutiger ist das Resultat bezüglich [sic!] Kundenzufriedenheit: ein Killer. Noch nie hat mich ein IVR entzückt und zu einem Promotor einer Firma gemacht. Menschen wollen nicht mit Maschinen sprechen und sich durch komplizierte Menüs wählen. Sie möchten in ihren eigenen Worten ihr Anliegen schildern und mit einem Menschen ein Gespräch führen.“²⁶

Der überwiegende Teil der großen Telekommunikationsunternehmen hat im Rahmen der technischen Weiterentwicklungen jedoch auf moderne Sprachdialogsysteme umgestellt und ermöglicht somit natursprachliche Dialoge zwischen System und Nutzer.

Wird auch die vom Nutzer vermeintlich als subjektiv besser wahrgenommene Nützlichkeit und Nutzerfreundlichkeit augenscheinlich zunächst als vorteilhaft wahrgenommen, kann bei der genaueren Analyse unterstellt werden, dass sich auch Nachteile auf Seiten der Nutzer, jedoch auch und insbesondere auf Seiten der Unternehmen durch den Einsatz rein natursprachlich basierter Dialoge ergeben können. Ein Auszug der bedeutendsten Vor- und Nachteile von IVR-Strukturen früherer Jahre und der heutzutage immer häufiger eingesetzten, natursprachlichen Dialoge für Nutzer und Unternehmen, werden in Tabelle 1 gegenübergestellt. Dieser erste Überblick zeigt bereits auf, dass die Entwicklung von einem systemgesteuerten zu einem nutzergesteuerten Dialog nicht nur Vorteile für Unternehmen und Nutzer beinhalten. Der vollständige Verzicht auf eine für den Nutzer wahrnehmbare Struktur könnte u.a. zur Orientierungslosigkeit des Nutzers im System führen. Durch die Möglichkeit der Benennung mehrfacher Anliegen durch den Nutzer innerhalb eines Anrufs, büßt das Unternehmen den Vorteil der eindeutigen, fachlichen Vorqualifizierung im Zweifel ein.

²⁶ Klötzli, M.: Director Enterprise Service bei Orange, im Interview mit Hasler, K.: unter: www.cmm-magazine.ch vom 1.9.2013 (zuletzt aufgerufen am 22.08.2015).

Ferner geht dem Unternehmen die Option verloren, den Nutzer planmäßig durch das Menü zu steuern und gezielt zu bestimmten Angeboten oder Selfcare-Applikationen zu navigieren, die die Agenten der Call Center-Organisation von Routinevorgängen entlasten und somit im Sinne einer Prozessautomatisierung Kosten einsparen könnten. Die vorrangigen Beweggründe zur Verfassung der vorliegenden Arbeit werden daher wie folgt definiert und verdeutlichen zugleich die Praxisrelevanz des Themenkonstrukts:

Die Motivation zur Erstellung der vorliegenden Arbeit besteht vorrangig darin, auf Basis der Erforschung von Ergonomiefaktoren stimmlicher Benutzerschnittstellen von Sprachportalen, neben Implikationen auf die Theorie auch Optimierungsansätze für die Praxis ableiten zu können, die einerseits zu einer erhöhten Akzeptanz auf der Nutzerseite bei andererseits bestmöglicher Ausschöpfung der Vorqualifizierungs-, Prozessautomatisierungs- und Kosteneinsparpotenziale auf Seiten der Unternehmen beitragen können. Ferner soll eine etwaige Lücke im Hinblick auf die für die Aufnahme und Verarbeitung rein auditiver Informationen relevanten Ergonomiekriterien für stimmliche Benutzerschnittstellen von Sprachportalen in der Theorie identifiziert werden und ein wissenschaftlicher Beitrag zur Schließung dieser Lücke erfolgen.

Tabelle 1: Vergleich IVR vs. Sprachdialogsystem.²⁷

IVR Menü		Natursprachliches Dialogsystem	
Nutzer	Unternehmen	Nutzer	Unternehmen
Die vordefinierten Begrifflichkeiten des IVR-Menüs dienen als Hilfestellung, wenn der Nutzer sich fachlich nicht klar ausdrücken kann und Schlüsselwörter wählen würde, die das System nicht kennt.	Der Nutzer kann keine Mehrfachanliegen eingeben. Der Nutzer ist gezwungen, sich für einen Menüpunkt, der seinem Anliegen entspricht, zu entscheiden. Dies erleichtert dem Unternehmen die Vorqualifizierung des Anruferanliegens.	Es gibt kein breites Eingangsmenü als Benutzerschnittstelle, d.h. der Nutzer muss sich keine Vielzahl an Punkten im Hauptmenü merken, da er sein Anliegen natursprachlich eingibt.	Das natursprachliche Dialogsystem erscheint dem Nutzer leichter zu bedienen, weniger komplex und somit zielführender zu sein, sofern der Dialog reibungslos abläuft.
	Durch die thematische Vorqualifizierung kann das Unternehmen den Nutzer zu einem speziell zum Thema geschulten Team von Agenten routen. Dies kann die Gesprächsdauer verkürzen und die Effizienz erhöhen.	Der Nutzer wird durch das System effektiv und effizient zu seinem Ziel, respektive direkt zum Agenten geroutet, ohne sich Schritt für Schritt durch das Menü navigieren zu müssen.	Vermutlich höhere Akzeptanz des Portals und somit höhere Nutzungsquote des Kommunikationskanals "Telefonie". Evtl. positiver Spillover-Effekt auf das Image des Unternehmens.
	Nutzer wählt einen Menüpunkt aus. Durch die thematisch eindeutige Vorqualifizierung kann das Unternehmen den Nutzer zu einem speziell zum Thema geschulten Team von Agenten routen. Dies kann zur schnellen und zufriedenstellenden Erledigung des Anliegens aus Nutzersicht führen.	Das Unternehmen kann den Nutzer über die einzelnen Menüschritte gezielt durch den Dialog steuern, z.B. zu Angeboten oder zu Self Care Applikationen, die Agenten Routineabläufe abnehmen sollen.	Geringere Quote an Nutzern, die aufgrund der geringen Nutzerfreundlichkeit eines IVR-Menüs den Vorgang abbrechen. Es kommen vermehrt wertvolle, persönliche Kontakte zustande, in dem z.B. Cross- oder Upsellingpotentiale ausgeschöpft werden könnten.
Nutzer muss sich ggf. viele Menüpunkte nacheinander anhören, bevor er eine einzige Auswahl treffen kann.	Die Nachteile für den Nutzer können sich als Unzufriedenheit, als gering wahrgenommene Benutzerfreundlichkeit sowie als geringe Akzeptanz zeigen und bis zur Meidung des Systems und zur Ausweichung auf andere Kontaktkanäle des Unternehmens führen.	Höhere Gefahr, dass freie Spracheingabe dazu führt, dass die vom Nutzer gewählten Schlüsselwörter vom System nicht erkannt werden.	Das Unternehmen hat deutlich geringere Möglichkeiten, den Nutzer im System gezielt zu steuern, z.B. zu bestimmten Self Care-Applikationen oder zu Angeboten. Durch den direkten, natursprachlichen Dialog findet die Nutzung nicht mehr systemgesteuert sondern nutzergesteuert statt.
Nutzer muss sich meist seriel im Menü vorwärts navigieren. Es gibt keine "Abkürzungen", keine Flexibilität, im Menü vorwärts zu springen, lediglich rückwärts, z.B. zurück ins Hauptmenü. System dominiert den Nutzer, nicht umgekehrt.	Bei Meidung des Systems durch den Nutzer, respektive das Ausweichen auf andere Kanäle, wie Web, Brief, eMail oder Fax, vermindert die Anzahl persönlicher Kontakte zwischen Kunde und Unternehmen, in denen kundenbindende oder vertriebliche Potentiale ausgeschöpft werden können.	Interpretiert das System die natursprachliche Eingabe nicht im Sinne des Nutzers, gibt es im Gegensatz zum IVR-Menü keine klassische "Zurück"-Funktion zur vorherigen Menüebene. Es besteht schnell Orientierungslosigkeit seitens des Nutzers, sofern keine geeigneten Hilfefunktionen systemseitig verfügbar sind.	Nutzer wählt nicht einen eindeutigen Menüpunkt aus. Nutzer kann durchaus mehrere Anliegen in einem Satz formulieren. System muss entscheiden, in welches Team es den Anruf zunächst leitet oder es muss organisatorisch ein Multiskill-Team eingerichtet werden.

²⁷ Eigene Darstellung.

1.2.2 Zielstellungen

Insbesondere in Call Center-Organisationen sind sprachdialogbasierte Anruferportale heutzutage branchenübergreifend weit verbreitet. Nahezu jeder geschäftsfähige Mensch hat wirtschaftliche Verbindungen zu Versicherungen und Telekommunikationsanbietern, ist Kunde einer Bank, eines Versandhauses oder nutzt Verkehrsmittel, z.B. von Bahn- oder Fluggesellschaften. Innerhalb des jeweiligen Kundenlebenszyklusses kommt es in der Regel häufiger zu Kontakten zwischen Konsumenten und Unternehmen.²⁸

Die Kontaktaufnahmen erfolgen zwar grundsätzlich bidirektional, werden aber in deutlich höherem Maße seitens der Interessenten und Kunden als durch das Unternehmen initiiert. Seitens der Unternehmen werden Kunden beispielsweise im Rahmen ihres Customer Lifecycle-, bzw. Customer-Relationship-Managements (CRM)²⁹ gezielt vor dem Beginn einer vertraglichen Kündigungsfrist angerufen, um Verträge zu verlängern und die Kunden somit juristisch an das Unternehmen zu binden. Darüber hinaus sollen Zusatzverkäufe generiert, respektive Up- und Cross-Selling-Potenziale ausgeschöpft werden.³⁰

“In lediglich 5% der Unternehmen initiiert das Call Center heute auch den Kontakt mit dem Kunden, z.B. für Verkauf oder Marktforschung“.³¹

Der Großteil der telefonischen Kontakte, der zwischen Interessenten, respektive Kunden und Unternehmen zustande kommt, wird somit kundenseitig initiiert. Bestandskunden wenden sich branchenabhängig mit den unterschiedlichsten Anliegen an ihr Unternehmen.

²⁸ Zu Kontaktfrequenzen siehe Abschnitt 1.1.

²⁹ Customer Lifecycle bedeutet Kunden-Lebenszyklus. CRM (Customer-Relationship-Management) heißt soviel wie Kundenbeziehungs-Management oder Kundenpflege und beinhaltet die Dokumentation und Verwaltung von Kundendaten in einem CRM-System als Voraussetzung für ein auf den Kunden ausgerichtetes Beziehungsmarketing. Vgl. hierzu Kotler, B., Bliemel, F.: Marketing-Management (1992), S. 1011.

³⁰ Upselling wird als Verkauf einer im Vergleich zum bereits vorhandenen Produkt höherwertigen Variante verstanden. Cross-Selling beinhaltet Zusatzverkäufe ergänzender Produkte und Dienstleistungen. Einen Überblick über Cross Selling-Potentiale verschiedener Branchen bieten Bruhn, M., Meffert, H.: Dienstleistungsmarketing (1997), S. 302.

³¹ Grewenig, I.: Multimedia- und Call-Center (2002), S. 10.

Das Spektrum reicht von einfachen Anfragen zu Produkten, kleinen Anpassungen in den Stammdaten des Kunden, wie der Aktualisierung einer Adresse, über vertragsrelevante Änderungen, wie z.B. einem Tarifwechsel, bis hin zu Kündigungen und Beschwerden.

“Der Kunde ist heute anspruchsvoll und verlangt einen schnellen Zugang zu dem kompetentesten Sachbearbeiter, der augenblicklich und freundlich seinen Wunsch erfüllen soll. Call Center gestatten diese Kundennähe. Der Kunde will seine Zeit nicht in Warteschlangen verbringen und dabei endlose Sprachmenüs anhören.“³²

Diese Ausführung fasst die durchschnittlichen Erwartungshaltungen von Konsumenten im Sinne einer Praxisannahme vermutlich recht kompakt zusammen.

Die skizzierten Erwartungen sehen sich nach Grewenig in der Praxis allerdings dennoch dem Umstand gegenüber, dass in Call Centern zunehmend Sprachdialogsysteme eingesetzt werden, die zum Ziel haben, die Kosten der Kundenkommunikation zu senken, gleichzeitig den Kundennutzen durch eine thematische Vorstrukturierung und verringerte Wartezeiten zu erhöhen und dabei eine verbesserte Kundenbindung zu erzielen.³³

“Wenn ein solches Menü es dem Anrufer nicht ermöglicht, sein Anliegen schnell und sicher durch Sprach- oder Tastatureingabe zu kommunizieren, so wird er entweder auflegen oder doch einen Agenten verlangen. Dann wird aber das Ziel einer automatisierten und dadurch kostengünstigen Abwicklung des Anrufs nicht erreicht.“³⁴

Ob vor diesem Hintergrund die genannten Ziele durch den Einsatz eines Sprachdialogsystems erreicht werden können, ist von vielfältigen Einflussfaktoren abhängig. Die folgende Abbildung stellt einige Faktoren überblickartig dar.

Die abgebildeten Faktoren lassen bereits auf die Interdisziplinarität des Konstrukts “Sprachdialogsystem“ schließen.

³² Menzler-Trott, E., Kundenorientierung in Menzler-Trott, E., Call Center Management (1999), S. 11, vgl. hierzu auch die repräsentative Studie zur Analyse der Anrufmotive von Kundenhotlines, o.V.: Vocatus Syndikatsstudie, Servicequalität (2004), S. 15.

³³ Vgl. Grewenig, I.: Multimedia- und Call-Center (2002), S. 10.

³⁴ Helber, S., Stollatz, R.: Call Center Management (2004), S. 8.

So spielen sowohl technische, verhaltenspsychologische, zeitliche und prozessuale, wie auch sprachwissenschaftliche Faktoren eine große Rolle. Aufgrund dieser fachlichen Interdisziplinarität des Themenkonstrukts wird dem Kapitel zur Abgrenzung und Definition des Untersuchungsgegenstands eine entsprechend hohe Bedeutung zugemessen.

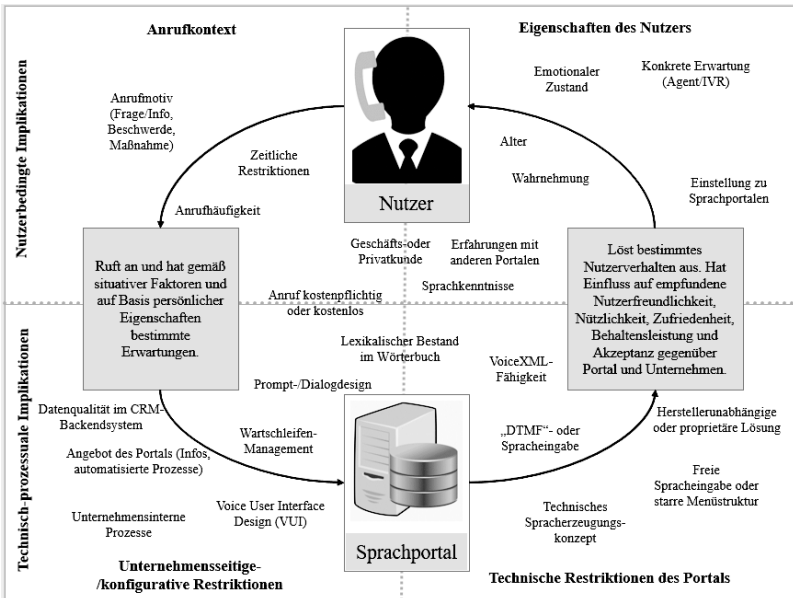


Abbildung 2: Einflussfaktoren auf die Akzeptanz von Sprachportalen.³⁵

Im Sinne einer Beschreibung der Zielstellungen dieser Arbeit können die Einflussfaktoren einerseits zunächst grob in solche unterteilt werden, die das Sprachdialogsystem direkt (z.B. technische Rahmenbedingungen oder das Design der stimmlichen Benutzerschnittstelle), bzw. mittelbar (z.B. zugrunde liegende Unternehmensprozesse und dadurch determinierte Durchlauf-, bzw. Wartezeiten für den Kunden) betreffen.

³⁵ Eigene Darstellung.

Andererseits existieren Einflussfaktoren, die eher dem Umfeld des Nutzers zuzuordnen sind (Erwartungen, Emotionen, Motive, Wahrnehmung, zeitliche Restriktionen, Alter, usw.). Jeder dieser Faktoren im Umfeld eines Sprachdialogsystems könnte zunächst als Stellschraube für den Grad der durch den Nutzer wahrgenommenen Nutzerfreundlichkeit, respektive Nützlichkeit mit entsprechenden Implikationen auf die Einstellungs- und Verhaltensakzeptanz angesehen werden. Im Sinne der empirischen Untersuchung wird diese Vielzahl an Faktoren jedoch bereits im Rahmen der Abgrenzung des Untersuchungsgegenstands im einleitenden Teil dieser Arbeit klar zwischen abhängigen und unabhängigen Variablen auf der einen Seite und Dritt-, bzw. Störvariablen auf der anderen Seite unterschieden. Die Inhalte des einleitenden Teils der Arbeit bestehen demnach neben einer thematischen Einführung, der Darstellung der Ziele und der Methodik der Arbeit, insbesondere aus der Abgrenzung und Definition des Untersuchungsgegenstandes. Im Zuge der folgenden Ausführungen soll das Hauptaugenmerk auf die Zielstellungen des theoretischen und empirischen Teils der Arbeit gerichtet werden, welche aus einer Differenzierung der Forschungsfragen abgeleitet wird. Eine erste Zielstellung im Rahmen des theoretischen Teils der Arbeit ist es, literaturbasiert herauszuarbeiten, welche Faktoren bei der Produktion und Rezeption auditiver Informationen im Rahmen des menschlichen Organismus eine bedeutende Rolle spielen und welche Implikationen auf das menschliche Verhalten bestehen können. Diese Erkenntnisse werden im Sinne eines Wissenstransfers dem aktuellen Stand der Forschung zur Entwicklung von Sprachportalen und somit einem Gegenstand des Forschungsfeldes der Künstlichen Intelligenz, gegenübergestellt. Ziel dieses Wissenstransfers ist es, Ergonomiefaktoren für die sprachlich-auditive Mensch-Maschine-Interaktion zu identifizieren und im Hinblick auf die spezifischen Belange auditiver Nutzerschnittstellen von Sprachportalen zu selektieren. Die Identifikation und Selektion relevanter Ergonomiefaktoren stellt die Basis für die weiterführenden Analysen dar.

Im Rahmen dieser wird dann eruiert, welche Einflussfaktoren, und zwar im engsten Sinne der Ergonomie auditiver Schnittstellen von Sprachportalen, gemäß dem aktuellen Stand der Forschung, bereits als solche in der Praxis Anwendung finden. Eine weitere Zielstellung des theoretischen Teils dieser Arbeit ist es demnach, aus diesem Vergleich herauszuarbeiten, ob und inwieweit spezifische Ergonomiekriterien bei der Entwicklung und Bewertung von Sprachportalen Beachtung finden. Es wird diskutiert, ob auf Basis des heutigen Stands der Forschung bereits Kriterien berücksichtigt werden, die in der Lage sind, die Interaktion zwischen Mensch und Maschine an den menschlichen Restriktionen ergonomisch auszurichten. Diese literaturbasierte Diskussion und Bewertung findet vor dem Hintergrund relevanter Erkenntnisse aus dem Wissensgebiet der Psychologie statt. Es werden Kriterien zusammengestellt, die eine hohe Relevanz im Hinblick auf die Ausrichtung der Mensch-Maschine-Interaktion an den menschlichen Limitierungen bei der Wahrnehmung und Verarbeitung von Sprachinformationen haben. Die ausgewählten Faktoren werden dann im Rahmen des empirischen Teils der Arbeit in Form unterschiedlicher Sprachportal-Modelle variiert. Die Wirkungen der unterschiedlichen Faktorstufen auf abhängige Variablen werden mit Hilfe empirisch etablierter Skalen gemessen und geeigneten Methoden statistisch analysiert. Ziel des empirischen Teils ist es somit, durch die Erhebung von Daten zu erforschen, wie sich die Variation von, für Sprachportale relevante, auditiven Ergonomiefaktoren auf die Bewertung von Nutzern im Hinblick auf die Behaltensleistung, Zufriedenheit, Nutzerfreundlichkeit, Nützlichkeit und die Einstellungsakzeptanz gegenüber dem Portal auswirkt. Darüber hinaus soll eruiert werden, inwiefern Erfahrungen von Nutzern auf die Evaluierung der einzelnen Sprachportal-Modelle wirkt. Die einzelnen Zielstellungen des theoretischen und empirischen Teils der Arbeit sind in der folgenden Übersicht kompakt zusammengefasst und zu Zwecken der verbesserten Zuordnung im Verlauf dieser Arbeit mit einer Identifikationsziffer versehen. Diese jeweilige Identifikationsziffer wird in den entsprechenden Kapiteln und Abschnitten, in denen die jeweiligen Teil-Forschungsfragen behandelt werden, im Sinne eines roten Fadens wieder aufgenommen.

Die Forschungsfragen werden sowohl für den theoretischen als auch für den empirischen Teil formuliert. Im Gegensatz zu den literaturbasierten und theoriegeleiteten Fragestellungen werden die Forschungsfragen im empirischen Teil auch als aus der Theorie sachlogisch fundiert hergeleiteten Hypothesen formuliert und in Form der Annahme oder Verwerfung dieser beantwortet. Zusammenfassend wird vereinfacht festgehalten, dass im Sinne der übergreifenden Forschungsfrage ergründet werden soll, inwiefern Sprachportalmenü-Modelle an den menschlichen Restriktionen von auditiv wahrzunehmenden und zu verarbeiten den Reizen ausgerichtet werden sollten, damit diese vom Nutzer als ergonomisch empfunden werden. Dieser Erkenntnisgewinn soll einen Beitrag zur Verbesserung der Nutzerakzeptanz von Sprachportalen leisten. Zur Beantwortung dieser Forschungsfrage werden im Verlauf des dargestellten, strukturellen Aufbaus der Arbeit, Teilaspekte analysiert in den jeweiligen Kapiteln gemäß formulierter Zielstellungen analysiert und beantwortet.

Tabelle 2: Forschungsfragen- und Zielmatrix.³⁶

	Forschungsfragen	Zielstellungen	ID
Theorie	Welche Faktoren spielen bei der Aufnahme und der Verarbeitung auditiver Informationen im menschlichen Organismus sowie der daraus resultierenden, menschlichen Verhaltensweisen eine zentrale Rolle?	Darstellung des Stands aktueller Forschungen und Herleitung relevanter Faktoren der menschlichen Aufnahme und Verarbeitung auditiver Informationen sowie der Implikationen auf menschliche Verhaltensweisen.	Z_01
	Welche Faktoren sind bei der Ein- und Ausgabe von Sprache bei und für das Verhalten von Maschinen maßgeblich und inwiefern fließen aktuelle Erkenntnisse der menschlichen Sprachproduktion und -Rezeption heutzutage in den Herstellungsprozess von Sprachdialogsystemen ein?	Darstellung des Stands aktueller Forschungen bzgl. maschineller Sprachausgabe und Sprachverarbeitungsmechanismen. Vergleichende Gegenüberstellung der Spezifika menschlicher sowie maschineller Sprachproduktion und -Rezeption und etwaiger Paralleltäten von Verhaltensweisen im Rahmen der Mensch-Maschine-Interaktion.	Z_02
	Welches sind die relevanten Ergonomiekriterien bei der Nutzung von Sprachdialogsystemen und inwiefern finden diese bei der Erstellung und Bewertung von Sprachdialogsystemen bereits heutzutage Berücksichtigung?	Darstellung des Stands aktueller Forschungen, die Ergonomie von Mensch-Maschine-Interaktionen im Allgemeinen und die Ergonomie auditiver Schnittstellen im Besonderen betreffend. Identifikation etwaiger Lücken in der Berücksichtigung bedeutender Ergonomiefaktoren bei der Herstellung und Bewertung von Sprachdialogsystemen.	Z_03
	Wie sind Ergonomiekriterien hinsichtlich ihrer Relevanz für die Interaktion zwischen Nutzer und Sprachportal zu bewerten? Welche Faktoren sollen im Sinne eines hohen Innovationsgrades im Rahmen dieser Arbeit detailliert analysiert werden und welche Hypothesen sind abzuleiten?	Bewertung von Ergonomiekriterien hinsichtlich ihrer Relevanz für die Interaktion zwischen Nutzer und Sprachportal sowie vor dem Hintergrund des Innovationsgrades für die vorliegende Arbeit. Definition und Konzeptualisierung der unabhängigen und der abhängigen Variablen. Ableitung und Verbalisierung von Hypothesen.	Z_04
	Welches Untersuchungskonzept ist zur empirischen Überprüfung der abgeleiteten Hypothesen geeignet?	Definition der übergreifenden Forschungsmethodik.	Z_05
	Wie sind die konstatierten Effekte reliabel und valide zu messen?	Operationalisierung der abhängigen und unabhängigen Variablen.	Z_06
	Wie werden die Daten erhoben?	Definition von Grundgesamtheit und Stichprobe. Beschreibung des Datenerhebungsprozesses.	Z_07
Empirie	Mit welchen Methoden und Hilfsmitteln werden die Daten statistisch ausgewertet?	Definition von Material und Methoden. Datenerfassung und Durchführung statistischer Analysen.	Z_08
	Welche Effekte gehen von der Variation der ausgewählten Ergonomiefaktoren auf die abhängigen Variablen aus?	Schriftliche, tabellarische und grafische Darstellung der statistisch analysierten Ergebnisse.	Z_09
	Wie sind die Ergebnisse zu interpretieren und wie sind die gewonnen Erkenntnisse in die Theorie einzuordnen?	Einordnung der Ergebnisse der empirischen Untersuchung in die Theorie, respektive in Studien anderer Autoren. Kritische Diskussion der Ergebnisse und der verwendeten Methoden. Ableitung potentieller Forschungsfragen für anknüpfende, weiterführende Analysen.	Z_10
	Spielen die gewonnen Erkenntnisse eine Rolle für die Praxis und wenn ja, welche?	Beschreibung etwaiger Optimierungspotentiale, die auf Basis der Erkenntnisgewinne der vorliegenden Arbeit bei der Gestaltung von Sprachportalen zu einer verbesserten Nutzerakzeptanz führen können.	Z_11

³⁶ Eigene Darstellung.

Die Wirkungszusammenhänge der beschriebenen Forschungsfragen und Zielstellungen werden nun im Rahmen der folgenden Erläuterungen zum strukturellen Aufbau der Arbeit zugeordnet und im Anschluss grafisch veranschaulicht.

1.3 Methodischer Ansatz und Aufbau der Arbeit

In diesem Abschnitt wird zunächst einleitend auf den grundsätzlichen methodischen Ansatz eingegangen. Einleitend bedeutet hier, dass, abgeleitet aus den im vorherigen Abschnitt erläuterten Zielstellungen, für den weiteren Fortgang der vorliegenden Arbeit geeignete Verfahrensweisen definiert und nachvollziehbar dargestellt werden. Die Forschungsmethodik für die empirische Untersuchung der Arbeit wird zu Beginn des empirischen Teils im Detail beschrieben.

1.3.1 Methodischer Ansatz

Der Begriff “Methode“ oder “Methodik“ steht nach Simon für einen systematischen Weg, der planmäßig zu einem definierten Ziel führt.³⁷

Bründler/Bürgisser/Lämmli postulieren, dass das wissenschaftliche Arbeiten zum Ziel hat, Wissen über die Beschaffenheit der Wirklichkeit zu erzeugen. Forscher planen demnach ihr Vorgehen und wählen entsprechend der Forschungsfragen die wissenschaftlichen Methoden zur Gewinnung von Daten und Erkenntnissen aus.³⁸

Zur Zielerreichung dieser Arbeit, respektive zur Beantwortung der gestellten Forschungsfragen, werden zwei methodische Ansätze gewählt, die aufeinander aufbauen.

³⁷ Vgl. Simon, F.: Die wissenschaftlichen Methoden unter www.stuwap.org (zuletzt aufgerufen am 29.08.2015).

³⁸ Vgl. Bründler, P., Bürgisser, D., Lämmli, D., et al., Einführung in die Psychologie und Pädagogik (2004), S. 31.

Zunächst wird im Rahmen des theoretischen Teils literaturbasiert relevantes Wissen analysiert und diskutierend gegenübergestellt, welches die Spezifika der Sprachproduktion, der Sprachrezeption und der durch die zwischenmenschlichen Sprachdialoge auslösbaren Verhaltensweisen beinhaltet. Diese Vorgehensweise erfolgt ebenfalls im Hinblick auf die maschinelle Sprachproduktion, die Sprach-eingabe sowie das maschinell vorprogrammierte Verhalten. Die Erkenntnisse werden vor dem Hintergrund der Mensch-Maschine-Interaktion im Rahmen der Nutzung von Sprachportalen mit dem Ziel vergleichend gegenübergestellt, spezifische Ergonomiekriterien für stimmliche Benutzerschnittstellen ableiten zu können. Darüber hinaus wird ebenfalls literaturbasiert überprüft, ob, bzw. welche Ergonomiekriterien bei der Entwicklung, Herstellung und Bewertung von Sprachportalen heutzutage ggf. bereits Anwendung finden. Ferner wird analysiert, ob diese, insbesondere vor dem Hintergrund der analysierten Spezifika kognitiver Verarbeitungsmechanismen rein auditiver Informationen, als hinreichend und geeignet erscheinen. Den Ergebnissen dieser Analyse zur Folge wird im engsten Sinne eingegrenzt, welche Ergonomiekriterien einer stimmlichen Benutzerschnittstelle im Hinblick auf die Wahrnehmung und Verarbeitung auditiver Informationen beim Menschen einen besonderen Stellenwert im Hinblick auf die konstatierte Wirkung auf die vom Nutzer empfundenen Konstrukte, wie die Zufriedenheit, Nutzerfreundlichkeit oder die Nützlichkeit haben könnte. Darüber hinaus werden etwaige Einflüsse auf die Einstellungsakzeptanz und der Einfluss der bisherigen Praxiserfahrungen der Probanden auf die abhängigen Variablen separat gemessen. Im Sinne eines innovativen Charakters dieser Arbeit wird dabei besonderes Augenmerk auf Faktoren gelegt, die im oben beschriebenen Sinne einerseits eine hohe ergonomische Bedeutung haben können, jedoch andererseits nach dem aktuellen Stand der Forschung ggf. noch keine sonderliche Beachtung bei der Entwicklung, Herstellung und Evaluierung von Sprachportalen gefunden haben. Der theoretische Teil schließt dementsprechend mit der Ableitung und Verbalisierung zu untersuchender Hypothesen. Die textuell abgeleiteten Hypothesen werden im Hinblick auf die konstatierte Wirkungsrichtung und die statistische Überprüfbarkeit tabellarisch dargestellt.

Im empirischen Teil der Arbeit wird das Forschungsdesign spezifiziert. Die unabhängigen und abhängigen Variablen werden operationalisiert. Anhand des Informationsbedarfs wird die geeignete Wahl der Datengewinnung diskutiert. Die Datenerfassung erfolgt anhand eines eigens entwickelten und im empirischen Teil der Arbeit, auch und insbesondere im Hinblick auf empirisch etablierte Skalierungen der Antwortoptionen, vorab getesteten und detailliert beschriebenen Fragebogens. Nach der Datenerhebung auf Basis der ebenfalls im Forschungsdesign erläuterten Stichprobe, erfolgt die statistische Auswertung der Daten. Der empirische Teil schließt auf Basis der Hypothesenprüfung mit der Interpretation der Ergebnisse sowie einer kritischen Diskussion der gewählten Methoden und der Einordnung der berichteten Befunde in die Theorie. Im Anschluss werden Implikationen auf die Praxis abgeleitet. Darüber hinaus wird die Arbeit in einem einseitigen Abstract kompakt zusammengefasst, welche der Arbeit vorangestellt wird. Die folgende Abbildung dient nicht zur Darstellung des strukturellen Aufbaus der Arbeit. Dieser wird im darauffolgenden Abschnitt hergeleitet. Die folgende Grafik verdeutlicht grafisch das in diesem Abschnitt beschriebene methodische Vorgehensmodell.

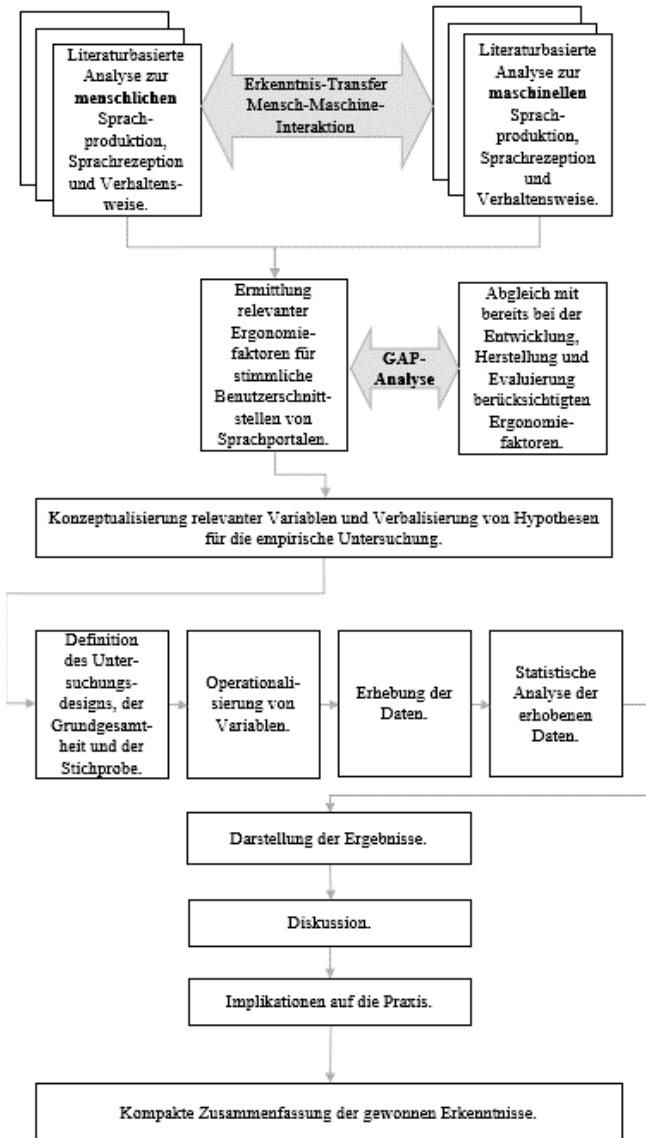


Abbildung 3: Methodisches Vorgehensmodell der Arbeit.³⁹

³⁹ Eigene Darstellung. Die finale Zusammenfassung/das Abstract der Arbeit wird dieser vorangestellt.

1.3.2 Struktureller Aufbau

Der strukturelle Aufbau der Arbeit folgt dem methodischen Vorgehensmodell und wird zunächst in Form von Abbildung 4 auf der folgenden Seite grafisch dargestellt. Im Anschluss werden die einzelnen Bestandteile des strukturellen Aufbaus kurz inhaltlich erläutert. Die unter Tabelle 2 dargestellten Zielstellungen werden in der folgenden Abbildung in Form der jeweiligen Identifikationsziffer auf der Ebene des Kapitels oder des jeweilig relevanten Abschnitts entsprechend zugeordnet. Durch diese Zuordnung wird deutlich, in welchem Kapitel, bzw. in welchem Abschnitt die Beantwortung der jeweiligen Forschungs-Teilfrage erfolgt. Die Arbeit gliedert sich auf der obersten Ebene in vier Hauptteile, und zwar in den einleitenden, den theoretischen, den empirischen und den schließenden Teil. Jeder Teil versteht sich in der Verzeichnisstruktur als erste Ebene, ist dementsprechend an der einstelligen Nummerierung zu erkennen und untergliedert sich auf der zweiten Ebene in unterschiedliche Kapitel. Sofern ein Kapitel in mindestens zwei weitere Komponenten unterteilt wird, untergliedert sich ein Kapitel auf der dritten Verzeichnisebene in mehrere Abschnitte, bzw. auf der vierten Ebene, wo aus Gründen der Übersicht sinnvoll und nötig, in entsprechende Unterabschnitte.

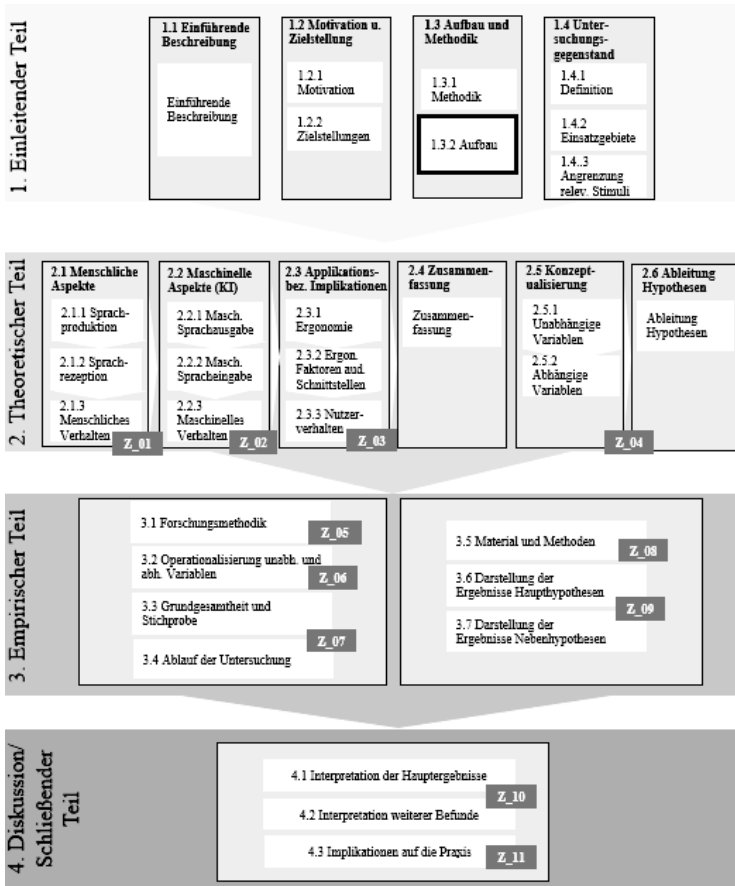


Abbildung 4: Struktureller Aufbau der Arbeit.⁴⁰

1.3.2.1 Einleitende Beschreibungen

In der einleitenden Beschreibung wurden durch die Erläuterung wichtiger Aspekte des betrachteten Themenkonstrukts die zu erwartenden Inhalte der vorliegenden Forschungsarbeit zunächst Überblickartig vermittelt.

⁴⁰ Eigene Darstellung.

In Verbindung mit den Erläuterungen im darauffolgenden Kapitel “Motivation und Zielstellungen“ wurden sowohl die Praxisrelevanz und der wissenschaftliche Mehrwert der Arbeit näher beschrieben als auch die Teilziele der jeweiligen Kapitel und Abschnitte sowie das Gesamtziel der Forschungsbestrebungen konkret erläutert. Inhaltlicher Kern der Darstellungen innerhalb dieses Kapitels ist es, neben der Erläuterung des methodischen Ansatzes auch den strukturellen Aufbau der Arbeit, unter Zuordnung definierter Zielstellungen, zu bestimmen. Der dargestellte, strukturelle Aufbau der Arbeit dient auch und insbesondere zur verbesserten Orientierung beim Lesen dieser recht umfangreichen Arbeit. Es wird jeweils anhand eines Bildausschnitts aus Abbildung 4 im weiteren Verlauf der Arbeit verdeutlicht, an welcher Stelle der Leser sich im Rahmen des strukturellen Aufbaus der Arbeit gerade befindet.⁴¹ Ferner wird die dem jeweiligen Kapitel oder Abschnitt zugeordnete Forschungsfrage beantwortet. Im darauffolgenden Kapitel 1.4 wird auf die Definition und Abgrenzung des Untersuchungsgegenstandes eingegangen. Zunächst wird der Stand der Forschung zur Sprachtechnologie und der entsprechenden Einsatzgebiete auf Basis einschlägiger Literatur hergeleitet. Auf dieser Grundlage werden dann die Spezifika des Einsatzes von Sprachtechnologie bei Sprachportalen in Call Centern detailliert verdeutlicht. Die Analyse an der Schnittstelle zwischen Mensch und Maschine ist vielschichtig, interdisziplinär und damit recht komplex. Wie bereits konstatiert, beinhaltet das Konstrukt “Sprachportal“ zahlreiche Variablen, die in Zusammenhang mit verschiedenen Prozessen innerhalb des menschlichen Organismus auf das Verhalten des Nutzers, als Reaktionen im Umgang mit einem Sprachdialogsystem, einwirken können.⁴²

Als Untersuchungsgegenstand wird daher das Konstrukt “Sprachdialogsystem“ definiert sowie die von ihm ausgehenden Stimuli herausgestellt. Es erfolgt eine Abgrenzung der für die Analyse relevanten Variablen von Dritt-, bzw. Störvariablen.

⁴¹ Auf die zusätzliche Darstellung des Oberkapitels in der Kopfzeile wird daher verzichtet.

⁴² Zum Stimulus-Reaktions-, bzw. Stimulus-Organismus-Reaktions-Paradigma vgl. Behrens, G.: Konsumentenverhalten (1991), S. 17., oder Kotler, B., Bliemel, F.: Marketing-Management (1992), S. 248, oder Peter, U.: Psychologie (1991), S. 9.

Aufgrund der beschriebenen Komplexität des interdisziplinären Untersuchungsgegenstands an der Schnittstelle zwischen Menschen und Maschinen und somit u.a. zwischen den Fachgebieten der Ingenieurs-, Wirtschafts-, Sprach- und Verhaltenswissenschaften, wird der genauen Definition des Untersuchungsgegenstandes, der entsprechenden Abgrenzung des zu analysierenden Themengebietes sowie der klaren Definition von Begrifflichkeiten innerhalb dieses Kapitels eine entsprechend hohe Bedeutung beigemessen, welche sich in der benötigten, inhaltlichen Tiefe und damit einhergehend auch in einem erhöhten, relativen Umfang dieses Kapitels widerspiegelt.

1.3.2.2 Inhalte des theoretischen Teils

Im theoretischen Teil der Arbeit werden Aspekte menschlicher Sprachproduktion und –Rezeption sowie Implikationen auf das Verhalten hergeleitet. Dies erfolgt literaturbasiert und vor dem Hintergrund der Spezifika auditiver Wahrnehmung. Dieser Erkenntnisstand wird im Sinne eines Know-How-Transfers dem ebenfalls literaturbasiert hergeleiteten Stands der Forschung, die maschinellen Aspekte der Sprachproduktion, Sprachverarbeitung sowie das maschinelle Verhalten im Rahmen des Forschungsfelds der Künstlichen Intelligenz betreffend, vergleichend gegenübergestellt. Aus diesem Vergleich wird begründet hergeleitet, welche applikationsbezogenen Ergonomiefaktoren für auditive Schnittstellen bedeutend sein könnten und inwiefern dieser Stand heute bereits bei der Erstellung von Sprachportalen Berücksichtigung findet oder nicht. Sofern die zum Stand der gegenwärtigen Forschung bereits zur Konzeption, Erstellung und Bewertung von Sprachportalen berücksichtigten Ergonomiefaktoren für die weitere Analyse als nicht hinreichend innovativ und geeignet erscheinen, werden Faktoren aus dem Forschungsgebiet der Psychologie ergänzt, denen, sachlogisch fundiert, eine hohe Bedeutung bei der Ausrichtung von Sprachapplikationen an den menschlichen Begebenheiten zugesprochen wird.

Auf dieser theoretischen Grundlage werden die unabhängigen und abhängigen Variablen abgeleitet. Im Sinne der Beantwortung der Forschungsfragen als Zielstellung für diesen Abschnitt werden neben der Spezifikation der Variablen entsprechende Hypothesen verbalisiert, die im Rahmen des empirischen Teils der Arbeit statistisch dargestellt und detailliert untersucht werden.

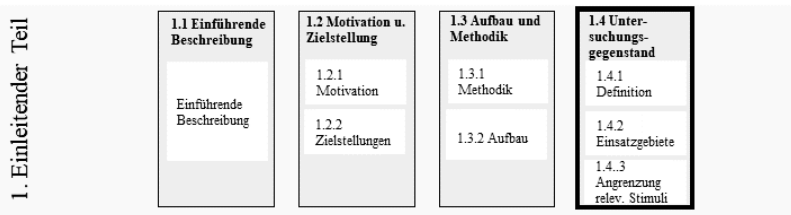
1.3.2.3 Aufbau der empirischen Analysen

Zunächst wird im Rahmen des empirischen Teils das Forschungskonzept definiert. Relevante Methoden der Primärforschung werden literaturbasiert gegenübergestellt und hinsichtlich ihrer Eignung für die empirische Analyse kritisch bewertet und ausgewählt. Die im Zuge der Feldphase erhobenen Daten werden im Statistik-Tool SPSS erfasst und mit dessen Hilfe auf Basis der definierten Methoden über die entsprechenden SPSS-Prozeduren statistisch analysiert. Die Ergebnisse werden im Anschluss strukturiert beschrieben, respektive textuell, tabellarisch und grafisch dargestellt.

1.3.2.4 Gegenstand der Diskussion

Im vierten Teil der Arbeit werden die Ergebnisse kritisch diskutiert, die Hypothesen überprüft und in Bezug zur Theorie und somit auch zu vergleichbaren Studien anderer Autoren gebracht. Es wird abgeleitet, inwiefern die im Zuge der vorliegenden Studie gewonnenen Erkenntnisse Implikationen auf die Praxis haben und somit einen Beitrag zu Verbesserung der Nutzerakzeptanz liefern können. Abschließend erfolgt eine kompakte, jeweils einseitige Zusammenfassung der Arbeit in deutscher und englischer Sprache, die, wie üblich, der Arbeit vorangestellt wird und einem potenziellen Leser als Entscheidungshilfe dienen kann, sich mit der vorliegenden Studie gemäß der jeweils individuellen Bedarfs- und Interessenslage vertiefend auseinander setzen zu wollen.

1.4 Abgrenzung des Untersuchungsobjekts



In diesem Kapitel wird das zu untersuchende Themengebiet eingegrenzt und definiert. Da die Analyse an einer Schnittstelle zwischen Mensch und Maschine durchgeführt wird, erstreckt sich die Definition des Untersuchungsgegenstands auf den genauen Umriss zweier unterschiedlicher Aspekte: Zum einen muss die systemtechnische Seite erläutert und eingegrenzt werden, da diese den Stimulusvariablen, also den unabhängigen Faktoren zugeordnet ist, die eine bestimmte Reaktion beeinflussen. Dazu wird das Konstrukt “Sprachdialogsystem“ insofern erläutert, als dass die für die vorliegende Arbeit relevanten Faktoren von nicht relevanten Aspekten abgegrenzt werden. Hierbei wird kurz auf die historisch bedingte Vielfalt von Ausprägungen an IVR-, bzw. Sprachdialogsystemen eingegangen. Dies ist im Sinne der Themeneingrenzung erforderlich, da heutzutage durchaus noch IVR-Systeme in der Praxis betrieben werden, die nicht Gegenstand der vorliegenden Untersuchung sind.

Im zweiten Teil dieses Kapitels werden die intervenierenden Determinanten im Sinne psychischer Prozesse erläutert und abgegrenzt, die im weiten Feld der Konsumentenforschung eine entsprechend hohe Relevanz für die vorliegende Analyse aufweisen. Hier existieren Interdependenzen zur Definition der Stimulusvariablen. So werden definitionsgemäß Sprachportale ausgeschlossen, die durch Internetapplikationen visuell unterstützt werden. Es finden daher ausschließlich sprachwissenschaftliche Aspekte von Sprachdialogsystemen ohne visuelle Stützung Berücksichtigung.

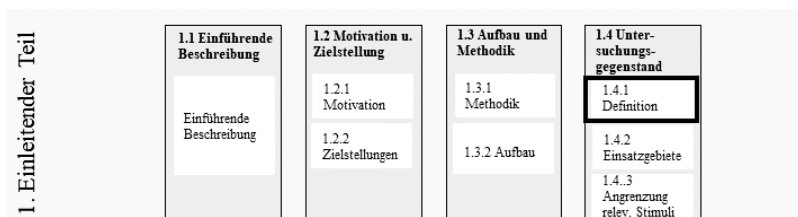
Dies führt dazu, dass im Rahmen der intervenierenden Determinanten des Konsumentenverhaltens, ausschließlich die auditive Informationsaufnahme und Informationsverarbeitung Gegenstand der Analysen sind. Aspekte der visuellen Informationsverarbeitung finden hingegen keine Berücksichtigung.

Dies ist nach Mangold/Pusler insofern wichtig, als dass auf der Rezeptionsebene die verschiedenen Sinneskanäle unterschiedliche Beiträge bei der Informationsaufnahme leisten.⁴³

So trägt das menschliche Auge in Verbindung mit dem Gehör nach Mangold/Pusler dazu bei, dass ca. 50% des Gesehenen und Gehörten aufgenommen und verarbeitet werden. Reize, die ausschließlich über das menschliche Gehör zugehen, werden lediglich zu ca. 20% gelernt. Eine mehrkanalige Ansprache kann somit die Lernleistung erhöhen.⁴⁴

Als verbindendes Element an der Schnittstelle zwischen Maschine und Mensch werden, vorrangig für diese spezielle Form der Kommunikation, ebenfalls relevante Aspekte der Sprachwissenschaften, sowie im Speziellen der linguistischen Datenverarbeitung für eine weitere, detaillierte Analyse im theoretischen Teil der Arbeit herausgestellt.

1.4.1 Definitionen und Stand der Forschung



In diesem Abschnitt wird der Stand der Forschung zur Sprachtechnologie beschrieben.

⁴³ Vgl. Mangold, M., Pusler, M., "Quality of Media", in: Häusel, H.-G., Neuromarketing (2008), S. 154.

⁴⁴ Ebd.