

Institut für Visualisierung und Interaktive Systeme
Abteilung Mensch-Computer-Interaktion
Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Diplomarbeit Nr. 3193

Multimodale Interaktion im Automobil mit Sprache und Gesten

Michael Kienast

Studiengang:	Informatik
Prüfer:	Prof. Dr. Albrecht Schmidt
Betreuer:	Dipl.-Inf. Bastian Pfleging
begonnen am:	01. Juni 2011
beendet am:	01. Dezember 2011
CR-Klassifikation:	H.5.2

Zusammenfassung

In dieser Arbeit wurde ein neuartiges Konzept zur Interaktion im Automobil aufgestellt und untersucht. Dabei wurden die beiden Modalitäten Sprache und Gestik miteinander kombiniert und daraus ein Interaktionsstil abgeleitet, der es erlaubt Objekte und Funktionen im Automobil mit Sprache auszuwählen und mit Gesten zu manipulieren.

In einer explorativen Benutzerstudie wurden Sprachkommandos und Gesten gesammelt und daraus jeweils ein konsistentes Set zusammengestellt. Diese Sprachbefehle und Gesten wurden anschließend in einen Prototypen integriert. Unter Berücksichtigung von Usability-Richtlinien erlaubt es dieser Prototyp Benutzern, vollständige Interaktionsfolgen bezüglich des Konzeptes dieser Arbeit durchzuführen.

Eine abschließende Evaluierung des prototypischen Systems und des Interaktionsstils mit einem Fahrsimulator zeigte, dass die Benutzer diesem Konzept gegenüber aufgeschlossen sind, sich durch die kurze Lernphase dieser Interaktionsform jedoch noch Schwierigkeiten in der Anwendung ergaben und sich keine eindeutige Aussage bezüglich eines verbesserten Fahrverhaltens und Fahrleistung machen ließ.

Abstract

In this work a novel automotive interaction concept was constructed and examined. In doing so, the two modalities speech and gestural interaction were combined and an interaction style was derived, which allows to select objects and functions in the car by using speech commands and to manipulate the corresponding parameters by using gestures.

In an explorative user study, speech commands and gestures were collected. Based on the findings of this study, a consistent set of gestures and appropriate speech commands was composed. Afterwards, those speech commands and gestures were used to create a prototypical implementation of the proposed interaction style. Considering usability principles, the prototype allows users to perform complete interaction sequences regarding the interaction concept of this work.

A final evaluation of the prototypical system and the interaction style using a driving simulator showed that users are receptive to this concept, however the short learning stage of this interaction style caused some difficulties in using it and no definite statement concerning improved driver behavior and driver performance could be made.

Inhaltsverzeichnis

1. Einleitung	11
2. Grundlagen und verwandte Arbeiten	15
2.1. Fahrerablenkung	15
2.1.1. Einflüsse auf die Fahrerablenkung	15
2.1.2. Messung von Fahrerablenkung	17
2.2. Interaktionsformen im Automobil	22
2.2.1. Interaktionsbereiche	22
2.2.2. Eingabemodalitäten	24
2.2.3. Ausgabemodalitäten	25
2.2.4. Sprachinteraktion	27
2.2.5. Gesteninteraktion	29
2.2.6. Multimodale Interaktion	31
3. Konzipierung eines multimodalen Systems	33
3.1. Eigenschaften von verbaler, gestikularer und manueller Interaktion	33
3.2. Multimodales Konzept dieser Arbeit	35
4. Explorative Benutzerstudie	37
4.1. Vorbereitung der Studie	37
4.2. Aufbau der Studie	38
4.3. Teilnehmer	41
4.4. Ablauf der Studie	41
4.4.1. Einführung in die Studie und erster Fragebogen	42
4.4.2. Detailliertere Erläuterung der Vorgehensweise	42
4.4.3. Erklärungen anhand einer Beispielaufgabe	43
4.4.4. Durchführung der Aufgabenreihe	44
4.4.5. Abschließender Fragebogen	45
4.5. Auswertung der Studie	45
4.5.1. Auswertung der Fragebögen	46
4.5.2. Auswertung der aufgezeichneten Sprachkommandos	50
4.5.3. Auswertung der aufgezeichneten Touch-Gesten	51

5. Multimodaler Prototyp	57
5.1. Konzipierung des Prototypen	57
5.2. Eigenschaften des Prototypen	58
5.2.1. Überblick über die Komponenten	58
5.2.2. Kommunikation der Komponenten	60
5.3. Steuerungsapplikation	61
5.4. Eingabekomponenten	63
5.4.1. Spracherkennung	63
5.4.2. Gestenerkennung	65
5.4.3. Pull-To-Talk	66
5.5. Ausgabekomponenten	66
5.5.1. Vorderer Bildschirm	66
5.5.2. Linker und rechter Bildschirm	68
5.5.3. Hinterer Bildschirm	70
6. Evaluierung des Prototypen und Interaktionsstils	71
6.1. Vorbereitung der Studie	71
6.2. Aufbau der Studie	71
6.3. Teilnehmer	74
6.4. Ablauf der Studie	75
6.4.1. Einführung in die Evaluierung und erster Fragebogen	75
6.4.2. Einweisung und Übungsfahrt mit dem Lane-Change-Test	76
6.4.3. Erste Baseline-Fahrt	77
6.4.4. Fahrten mit unterschiedlichen Interaktionsstilen	77
6.4.5. Zweite Baseline-Fahrt	80
6.4.6. Abschließender Fragebogen	80
6.5. Auswertung der Studie	81
6.5.1. Auswertung der Fragebögen	81
6.5.2. Analyse der aufgezeichneten Lane-Change-Test-Daten	86
7. Zusammenfassung und Ausblick	89
A. Anhang	95
A.1. Einverständniserklärung - Explorative Benutzerstudie	95
A.2. Eingangsfragebogen - Explorative Benutzerstudie	96
A.3. Zwischenfragebogen - Explorative Benutzerstudie	98
A.4. Abschlussfragebogen - Explorative Benutzerstudie	100
A.5. Quittung - Explorative Benutzerstudie	103
A.6. Leitfaden - Explorative Benutzerstudie	104
A.7. Sprachkommando-Set als Ergebnis der explorativen Benutzerstudie	107

A.8. Einverständniserklärung - Evaluierung	116
A.9. Eingangsfragebogen - Evaluierung	117
A.10.Zwischenfragebogen Interaktion mit Sprache und Gesten - Evaluierung	119
A.11.Zwischenfragebogen Interaktion mit Buttons und Knöpfen - Evaluierung	123
A.12.Abschlussfragebogen - Evaluierung	127
A.13.Quittung - Evaluierung	129
A.14.Leitfaden - Evaluierung	130
Literaturverzeichnis	135
Inhalt der CD-ROM	143

Abbildungsverzeichnis

2.1.	Shutterbrille zur Verwendung bei der Okklusionsmethode	18
2.2.	Fahrspurwechselanweisung während des Lane-Change-Tests	20
2.3.	Bestimmung der Fahrabweichung nach Durchführung des Lane-Change-Tests	20
2.4.	Ansicht der Simulator-Komponente des Configurable Automotive Research Simulator	21
2.5.	Karten-Editor und Analyse-Tool des Configurable Automotive Research Simulator	22
2.6.	Physische Aufteilung des Fahrzeugcockpits nach Aufgabenbereichen	23
2.7.	Bildschirm und Steuerelement des BMW iDrive-Systems	25
2.8.	Ansicht eines Head-Up-Displays auf der Windschutzscheibe	26
2.9.	Handgeste zum Umschalten von Radiosendern und Musikstücken	29
2.10.	Auf einem Lenkrad befestigtes Touchpad zur Gesteneingabe mit dem rechten Daumen	30
2.11.	Aufbau einer gestenbasierten, handschriftlichen Texteingabe auf dem Lenkrad und der Mittelkonsole	31
3.1.	Konzept der multimodalen Interaktion mit Sprache und Gesten	35
4.1.	Holzlenkrad auf einem Logitech G27 Unterbau und ein darin integrierter Tablet-Computer	40
4.2.	Positionierung der Probanden während der explorativen Benutzerstudie . . .	40
4.3.	Angezeigte Bild-Typen während der Benutzerstudie zur Sammlung von Sprachkommandos	43
4.4.	Durchführung einer Gesteneingabe während der explorativen Benutzerstudie	45
4.5.	Ablaufdiagramm der explorativen Benutzerstudie	46
4.6.	Mittlere Beurteilung einzelner Funktionen hinsichtlich der Sinnhaftigkeit einer Bedienung mit Sprache und Gesten	49
5.1.	Anordnung der Komponenten des Prototypen	59
5.2.	Architektur des EIToolkit	61
5.3.	Rückseitenansicht des Prototyp-Lenkrads mit markierten Pull-To-Talk-Hebeln	66
5.4.	Ansicht des vorderen Prototyp-Bildschirms	67

5.5. Ansichten der links und rechts des Benutzers angebrachten Bildschirme des Prototypen	69
5.6. Ansicht des hinteren Bildschirms beim Scheibenwischen mit Wischwasser . .	70
5.7. Gesamtaufbau des Prototypen	70
6.1. Physischer Aufbau der Evaluierung	73
6.2. Nahaufnahmen der in der Evaluierung verwendeten Android-Lenkrad-Applikation, Wii-Controller und Halterung	74
6.3. Ansicht eines Probanden bei der Fahrt mit dem Lane-Change-Test während der Evaluierung	80
6.4. Ablaufdiagramm der Evaluierung	81
6.5. Ergebnisse der SUS-Methode für die Evaluierung	83
6.6. Ergebnisse der DALI-Methode für die Evaluierung	84
6.7. Darstellung der Anzahl der durchgeführten Interaktionsaufgaben je Proband und Interaktionsstil während der Evaluierung	86
6.8. Mittlere Fahrabweichung aller vier Fahrttypen der Evaluierung	87

Tabellenverzeichnis

2.1. Statistik schwerer Unfälle, involvierter Fahrer und Todesopfer in den USA von 2005 bis 2009	17
4.1. Verwendete Objekte und Funktionen der explorativen Benutzerstudie	39
4.2. Allgemeine Informationen zu den Teilnehmern der explorativen Benutzerstudie	41
4.3. Gesten für die Funktionen des Fahrersitzes	53
4.4. Gesten für die Funktionen der Außenspiegel	54
4.5. Gesten für die Funktionen der Scheibenwischer	54
4.6. Gesten für die Funktionen der Seitenscheiben	55
4.7. Gesten für die Funktionen des Tempomaten, der Luftdüsen und des Head-Up-Displays	56
6.1. Übersicht der Objekte und durchzuführenden Funktionen der Evaluierung . .	72
6.2. Allgemeine Informationen zu den Teilnehmern der Evaluierung	75

Verzeichnis der Codefragmente

5.1. Codebeispiel in C# zum Anlegen und Versenden eines UDP-Paketes mit dem EIToolkit	60
5.2. Entgegennahme und Auswertung empfangener UDP-Nachrichten mit dem EIToolkit	62
5.3. Deklaration eines Sprachsynthese-Objektes und Übergabe einer Zeichenkette zur Sprachausgabe mit der SpeechSynthesizer-API	63
5.4. Deklaration eines Spracherkennungs-Objektes und Definition einer Grammatik mit der SpeechRecognizer-API	64

1. Einleitung

Viele Menschen verbringen täglich eine Stunde oder mehr in ihren Fahrzeugen [65] und werden mit immer neuen Funktionen darin konfrontiert. Diese Funktionen gehen inzwischen weit über die Bedienung der Pedale und das Lenkrades hinaus und somit über die eigentliche Fahraufgabe. Es steht außer Frage, dass der Fahraufgabe die höchste Priorität eingeräumt werden muss, doch führt die stete Zunahme an neuen Funktionen zwangsläufig auch zu einer verringerten Aufmerksamkeit bezüglich dieser primären Aufgabe.

Obwohl neue Funktionen, allen voran Infotainment- und Entertainmentsysteme, den Fahrkomfort gerade während längerer Fahrtzeiten erhöhen, ist es andererseits eine Notwendigkeit den zuvor erwähnten Konzentrationsverlust soweit nur möglich zu reduzieren. Dass dies keine einfache Aufgabe ist, legt die Tatsache nahe, dass nicht nur ein optischer Reiz, z.B. eine Anzeige des Navigationssystems die Aufmerksamkeit des Fahrers auf sich ziehen kann, auch akustische Ablenkung kann den Fokus der Augen von der Straße lenken. Des Weiteren induzieren neue, komplexe Funktionalitäten auch Denkprozesse, die den Fahrer in Anspruch nehmen. Dieser Vielzahl an negativen Auswirkungen entgegenzutreten erfordert ausgeklügelte Strategien und Interaktionsstile.

Nicht wenige Forschungsansätze suchen bereits nach neuen Möglichkeiten der Interaktion im Automobil unter Nutzung verschiedener Ein- und Ausgabemodalitäten. Dabei liegen die Ziele beispielsweise auf einer erleichterten und weniger ablenkenden Bedienung von Informations-, Kommunikations-, oder Unterhaltungssystemen. Die in diesen Fällen eingesetzten Modalitäten umfassen ein breites Spektrum: von touch-basierter Interaktion wie in [21, 25], Spracheingabe [14, 34, 67], aber auch innovativen Techniken wie Eingaben über Hand-, Head- und Eye-Tracking [2, 3, 38].

Diese Arbeit nimmt die Intention solcher Forschungsansätze auf und untersucht die Kombination von Interaktionstechniken, die jede für sich mit Nachteilen zu kämpfen haben. Diese Techniken erscheinen in einer multimodalen Interaktionsform jedoch geeignet, die Ausübung neuer Aufgaben und Funktionen in einer intuitiven und effizienten Art und Weise durchführen zu können. Würde diese Zielsetzung für eine Desktopanwendung ausreichen, so spielt im Automobilbereich die angesprochene Ablenkung des Fahrers aus Sicherheitsgründen eine sehr große Rolle. Ein derartiger Interaktionsstil ist somit in diesem Kontext nur dann von Nutzen, wenn er die Ablenkung des Fahrers nicht erhöht, sondern bestenfalls reduziert.

1. Einleitung

Konkretes Ziel dieser Arbeit ist es, die Eingabemodalitäten Sprache und Gestik so zu vereinen, dass einerseits eine intuitive Interaktion im Automobil möglich ist und andererseits die Nachteile beider Modalitäten so abgeschwächt werden, dass eine Akzeptanz unter Benutzern gegeben ist und Aufmerksamkeit wieder verstärkt der primären Fahraufgabe gewidmet werden kann.

Sprache ist längst keine Innovation mehr im automobilen Anwendungsbereich, doch kämpft sie für sich genommen in keiner leisen Umgebung und auch andere Faktoren wie technische Einschränkungen senken unter Umständen die Erkennungsgenauigkeit der verbalen Äußerungen. Der Ansatz dieser Arbeit ist nicht dieses Problem vollständig aufzulösen und eine hundertprozentige Erkennungsgenauigkeit zu bieten, doch beschränkt sich die Spracheingabe hier auf die Auswahl des Objektes mit dem interagiert werden soll in dem Sinne, dass lediglich ein kurzes Sprachkommando gegeben wird, mit dem ein Interaktionsobjekt eindeutig gekennzeichnet ist. Damit reduziert sich die Verwendung dieser Modalität auf kurze Wortfolgen und senkt die Wahrscheinlichkeit von ungenauen oder falschen Erkennungen. Zur Manipulation eines ausgewählten Objekts kommt in der Folge die Modalität Gestik ins Spiel. Über möglichst intuitiv aufgebaute Gesten soll über eine Touch-Oberfläche, die wiederum direkt auf dem Lenkrad befestigt ist, eine zentralisierte Bedienung verschiedenster Funktionen möglich werden. Die in dieser Arbeit mit dem beschriebenen multimodalen Interaktionsstil auswählbaren und manipulierbaren Objekte sind u.a. die Fensterscheiben, die Außenspiegel und die Scheibenwischer.

Nach der Erschließung der grundlegenden Sachverhalte und verwandter Arbeiten in Bezug auf Fahrerablenkung und automobilen Interaktionsformen, steht die Erstellung eines schlüssigen Konzeptes für den dargelegten multimodalen Interaktionsstil im Vordergrund. Dabei wird der Sammlung von benutzerdefinierten Sprachkommandos und Gesten eine besondere Rolle zugemessen. Um dem Rechnung zu tragen, werden in einer Studie von Benutzern geäußerte Sprachkommandos und Gesten auf ihre Brauchbarkeit und Intuitivität hin untersucht und in der Folge für eine Verwendung zusammengestellt und aufbereitet. Eine anschließende prototypische Realisierung integriert die zuvor gesammelten Sprachbefehle und Gesten und macht vollständige Interaktionsvorgänge während der Fahrt mit einem Fahrsimulator möglich. Diese Umsetzung und der zugehörige Interaktionsstil werden dann wiederum in einer abschließenden Evaluierungsphase auf ihre Eignung hin untersucht und bewertet.

Gliederung

Die Arbeit ist in folgender Weise gegliedert:

Kapitel 2 – Grundlagen und verwandte Arbeiten erläutert Interaktionsformen im Automobil sowie ihre Auswirkungen auf die Fahrerablenkung. Dabei wird die Fahrerablenkung als Ursache für schwere Unfälle hervorgehoben und Methoden angeführt, Fahrerablenkung zu messen. Des Weiteren werden Ein- und Ausgabemodalitäten betrachtet, speziell Sprach- und Gesteninteraktion. Diese Interaktionsformen werden abschließend im Rahmen multimodaler Interaktion diskutiert.

Kapitel 3 – Konzipierung eines multimodalen Systems beschreibt das verfolgte Konzept zur Umsetzung einer multimodalen Interaktionsform. Dazu werden verschiedene Modalitäten hinsichtlich Vor- und Nachteilen beurteilt und die daraus resultierende Konzeption eines multimodalen Interaktionsstils unter Beachtung von Usability-Richtlinien besprochen.

Kapitel 4 – Explorative Benutzerstudie erläutert den Aufbau und Ablauf der ersten Benutzerstudie zur Sammlung von benutzerdefinierten Sprachkommandos und Gesten. Alle Aktivitäten der Vorbereitung dieser explorativen Benutzerstudie, des physikalischen Aufbaus der Hardware, die eingesetzte Software und die Durchführung werden ausgeführt. Abschließend werden die Auswertung und die Ergebnisse der Studie dargelegt.

Kapitel 5 – Multimodaler Prototyp stellt den prototypischen Aufbau eines multimodalen Systems, das Sprach- und Gesteneingabe integriert, vor. Dieser, in einem Labor gebaute Prototyp, setzt das Konzept einer multimodalen Interaktion um, bei dem der Benutzer Sprachkommandos zur Objekt- und Funktionsauswahl und Gesten zur Parametermanipulation einsetzen kann. Sämtliche hard- und softwarespezifischen Eigenschaften dieses Prototypen werden erörtert und illustriert.

Kapitel 6 – Evaluierung des Prototypen und Interaktionsstils beschreibt die Durchführung einer abschließenden Evaluierung und deren Ergebnisse. Dabei wird der Aufbau und Ablauf dieser Vergleichsevaluierung dargelegt, die zwei verschiedene Interaktionsstile in Konkurrenz zueinander stellt. Die diesbezüglichen Analysemethoden sowie deren Resultate werden aufgezeigt.

Kapitel 7 – Zusammenfassung und Ausblick fasst abschließend die Arbeit und deren Ergebnisse zusammen und stellt Anknüpfungspunkte vor.

2. Grundlagen und verwandte Arbeiten

Dieses Kapitel erläutert die Grundlagen und verwandte Arbeiten bezüglich der verschiedenen Interaktionsmöglichkeiten im Automobil und ihre Auswirkungen auf die Fahrerablenkung. Dabei werden Ein- und Ausgabemethoden dargestellt und in diesem Zusammenhang Interaktionsformen unter dem Einsatz von Sprache und Gestik näher beschrieben. Die Gefahren der Fahrerablenkung sowie der Beitrag multimodaler Interaktion zur Reduzierung dieser werden diskutiert.

2.1. Fahrerablenkung

Abschnitt 2.1.1 stellt die Auswirkungen von Fahrerablenkung auf das Fahrverhalten und deren (teils drastische) Konsequenzen dar und zeigt so die Notwendigkeit einer Berücksichtigung des Einflusses neuer Funktionen im Automobil auf die Fahrerablenkung auf. Abschnitt 2.1.2 gibt dementsprechend einen Überblick, wie Fahrerablenkung im Vorfeld untersucht werden kann.

2.1.1. Einflüsse auf die Fahrerablenkung

Die Interaktion mit beispielsweise Desktop-Computern erlaubt es dem Benutzer, seine volle Konzentration auf den Interaktionsvorgang zu richten. Dieser Aspekt trifft für den Fahrer eines Automobils nicht zu, da seine wesentliche Aufgabe in der Steuerung seines Fahrzeugs liegt und vor allem Interaktionsvorgänge mit z.B. dem Radio oder Navigationssystem keine zentralen, sondern nebenläufige Interaktionsaufgaben darstellen und somit nur geringe Aufmerksamkeit erfordern sollen und dürfen. Um die durch nebenläufige Aufgaben entstehende Ablenkung gering zu halten, muss die Benutzeroberfläche eines Automobils entsprechend entworfen werden. Dazu stehen Entwicklern und Designern von Nutzungsschnittstellen im automobilen Kontext verschiedene Methoden, unter anderem internationale Standards wie z.B. [17] und [19] sowie Richtlinien wie [16] und [58] zur Verfügung, die Empfehlungen und zwingende Vorgaben bezüglich des Interaktionsdesigns machen. So sollten beispielsweise Interaktionselemente so entworfen und positioniert werden, dass eine unabsichtliche Aktionsauslösung vermieden wird oder dass der Fahrer während einer Interaktion jederzeit in

der Lage ist, mindestens eine Hand auf dem Lenkrad zu belassen [16]. Des Weiteren geben sie Empfehlungen über den Einsatz von spezifischen Farben, Helligkeitsunterschieden oder Schriftarten [58] zur Rückmeldung von visuellen Informationen an den Fahrer.

Um die Ablenkung zu minimieren, muss eine Benutzerschnittstelle grundlegende ergonomische Belange erfüllen, d.h. sie muss einfach zu erlernen und intuitiv zu bedienen sein. Außerdem muss sie spezifische Merkmale bieten, die die Fahrerablenkung reduzieren. Dazu gehört beispielsweise, dass dem Benutzer die Kontrolle über die Interaktionsgeschwindigkeit überlassen wird [30]. Ein solches System muss über Eigenschaften verfügen, die jede der vier Arten von Ablenkung individuell adressieren. Neben der visuellen Ablenkung zählen dazu die auditive, biomechanische und kognitive Ablenkung. Eine visuelle Ablenkung tritt auf, wenn der Fahrer den Blick auf die Straße vernachlässigt und seine visuelle Aufmerksamkeit für eine gewisse Zeit auf ein anderes Ziel richtet. Bei auditiver Ablenkung legt der Fahrer seine Aufmerksamkeit auf Töne und Signale, wie z.B. Musik aus dem Autoradio, aber auch die Unterhaltung mit einem Beifahrer kann diese Art der Ablenkung fördern. Eine biomechanische, physische oder manuelle Ablenkung tritt dann auf, wenn der Fahrer eine oder beide Hände vom Lenkrad nimmt, um eine gewisse Zeit lang z.B. einen Drehregler zu manipulieren oder einen Schalter zu bedienen. Die Ursache für kognitive Ablenkung sind sämtliche Denkvorgänge, die die Aufmerksamkeit des Fahrers in einer Art und Weise auf sich ziehen, dass ein sicheres Navigieren im Straßenverkehr nicht mehr gewährleistet und die Reaktionszeit reduziert ist [1].

Die stetige Zunahme an Funktionen innerhalb des Automobils, z.B. Infotainment- und Fahrassistenzsysteme betreffend, erhöhen einerseits den Fahrkomfort, andererseits erhöhen sie die Gefahr der Fahrerablenkung. Wie wichtig ein ausgeklügeltes Design der automobilen Benutzerschnittstellen ist, sodass die Fahrerablenkung auf ein Mindestmaß reduziert werden kann, zeigt Tabelle 2.1. Sie gibt einen Überblick der Verkehrsunfälle mit Todesfolge in den USA in den Jahren 2005 bis 2009. So gab es im Jahr 2009 30 797 Unfälle mit Todesfolge, in die 45 230 Fahrer involviert waren und bei denen 33 808 Personen starben. Für 11% der involvierten Fahrer wurde Fahrerablenkung als Unfallursache angegeben. In diesen durch Fahrerablenkung verursachten Unfällen starben 16 % aller Todesopfer. Der Anteil der Todesfälle, bei denen Fahrerablenkung eine Rolle spielte, stieg von 10% auf 16% innerhalb von nur fünf Jahren. Ebenso stieg der Anteil der Unfälle bei denen Fahrerablenkung eine Rolle spielte von 10% auf 16% innerhalb desselben Zeitraumes [66].

Die Ursachen der dabei eine Rolle spielenden Fahrerablenkung lassen sich nicht nur auf die Interaktion mit automobilen Benutzerschnittstellen zurückführen, auch Unterhaltungen mit dem Beifahrer oder das Fokussieren auf die äußere Umgebung und weg von der Straße sind für tödliche Verkehrsunfälle verantwortlich gewesen. Klauer et al. [42] führen visuelle Ablenkung als Schlüsselfaktor bei Unfällen oder Beinahe-Unfällen an. So besteht für Fahrer, die komplexere visuelle oder manuelle Aufgaben durchführen ein dreifach höheres Risiko

Jahr	Gesamt			durch Fahrerablenkung		
	Verkehrsunfälle	beteiligte Fahrer	Todesfälle	Verkehrsunfälle	beteiligte Fahrer	Todesfälle
2005	39 252	59 220	43 510	4 026 (10%)	4 217 (7%)	4 472 (10%)
2006	38 648	57 846	42 708	5 245 (14%)	5 455 (9%)	5 836 (14%)
2007	37 435	56 019	41 259	5 329 (14%)	5 552 (10%)	5 917 (14%)
2008	24 172	50 416	37 423	5 307 (16%)	5 477 (11%)	5 838 (16%)
2009	30 797	45 230	33 808	4 898 (16%)	5 084 (11%)	5 474 (16%)

Tabelle 2.1.: Statistik schwerer Unfälle, involvierter Fahrer und Todesopfer in den USA von 2005 bis 2009 (Datenquelle: [66])

eines Unfalls oder Beinahe-Unfalls als für Fahrer, die auf die Fahraufgabe konzentriert sind. Bei einer Blickablenkung, bei der die Augen nicht auf die Straße fokussiert sind und die mehr als zwei Sekunden dauert, ist die Gefahr eines Unfalls oder Beinahe-Unfalls bereits um mindestens das Zweifache erhöht [42].

Dass die Bedienung neuer Funktionen im Automobil auf Kosten der Fahrerablenkung geht, ist dementsprechend aus Sicherheitsgründen nicht akzeptabel und erfordert für jede dieser Funktionen eine Überprüfung der entstehenden Fahrerablenkung noch vor der Integration in ein Automobil.

2.1.2. Messung von Fahrerablenkung

Um die Auswirkungen der Durchführung von Aufgaben neben der primären Fahraufgabe auf die Fahrerablenkung messen zu können, wurden verschiedene Methoden entwickelt. Diese ermöglichen es, die durch Nebenaufgaben induzierten Auswirkungen auf die Aufmerksamkeit des Fahrers in einer Laborumgebung zu messen. Urbas et al. [64] geben für den Verzicht auf Messungen während Realfahrten zum einen pragmatische Gründe wie Kosten- und Risikofaktoren und testtheoretische Gründe wie Wiederholbarkeit und Kontrollierbarkeit an. Unter kontrollierten Bedingungen lassen sich so bestimmte Aspekte des Fahrens

nachstellen und Hinweise auf Abweichungen im Fahrverhalten mit und ohne Interaktion erhalten. Einige dieser Methoden werden im Folgenden näher erläutert.

Okklusionsmethode

Mit der Okklusionsmethode werden Blickunterbrechungen während der Bedienung von Nebenaufgaben simuliert. Dabei kommt eine Shutterbrille wie in Abbildung 2.1 gezeigt zum Einsatz, deren Sichtfeld abwechselnd geöffnet und geschlossen wird. Diese Öffnungs- und Verschlussintervalle können gezielt eingestellt werden. Das Ziel der Okklusionsmethode ist es, die Unterbrechung einer Nebenaufgabe zu simulieren, welche durch die Fahraufgabe hervorgerufen wird [24]. Die Durchführung der Fahraufgabe selbst entfällt damit. Die ISO-Norm 16673 [35] gibt einen zweckgemäßen Messvorgang vor, welcher den Grad der Unterbrechbarkeit misst. D.h., es wird die Dauer der Durchführung der Nebenaufgabe mit Okklusion und ohne Okklusion gemessen. Ein positives Resultat liegt dann vor, wenn der Einsatz von Okklusion im Vergleich nur wenig Einfluss auf die Durchführungszeit der Nebenaufgabe hat. Dies bedeutet, dass kurze Blicke oder geringe Mengen an visueller Aufmerksamkeit für die Durchführung der Nebenaufgabe genügen und sie nach der Unterbrechung leicht wieder aufgenommen werden kann [72]. Damit kann die Nebenaufgabe auch dann noch akzeptabel durchgeführt werden, wenn die primäre Fahraufgabe die Hauptaufmerksamkeit erfordert und somit auch die Ablenkung von der primären Fahraufgabe durch die Nebenaufgabe minimal ist.

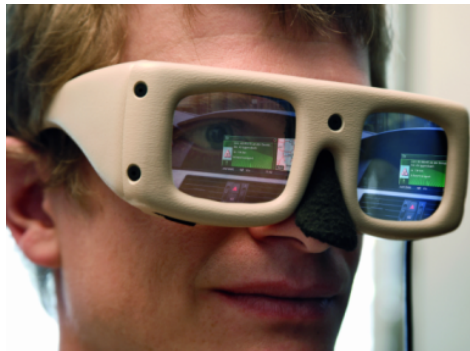


Abbildung 2.1.: Shutterbrille zur Verwendung bei der Okklusionsmethode (Quelle: [10])

Peripheral-Detection-Task

Die von Martens und van Winsum entwickelte Peripheral Detection Task-Methode (PDT) [45] wird während der Fahrt mit einem Fahr Simulator und der zeitgleichen Durchführung einer Nebenaufgabe eingesetzt. Dabei werden dem Probanden alle drei bis sechs Sekunden für maximal zwei Sekunden visuelle Reize im peripheren Blickfeld angezeigt [64]. Sobald der

Proband einen solchen Reiz erkennt, soll er so schnell wie möglich durch einen Tastendruck darauf reagieren. Ausgewertet werden dann die Reaktionszeiten und die Entdeckungshäufigkeiten der visuellen Reize. Die visuellen Reize stellen dabei wahrzunehmende Objekte oder Ereignisse im peripheren Blickfeld dar, auf die während einer Fahrt trotz Nebenaufgabe und sich damit eventuell einstellendem Tunnelblick reagiert werden muss [64].

Lane-Change-Test

Der Lane-Change-Test (LCT) ist eine von der DaimlerChrysler AG (heute Daimler AG) entwickelte PC-basierte Fahrsimulation zur Messung des Einflusses von Nebenaufgaben auf die Fahrleistung und damit der Fahrerablenkung [46]. Für den Betrieb des Simulators sind lediglich ein PC mit Monitor, ein Spiele-Lenkrad sowie Fußpedale erforderlich. Die Methode und die Hardwarespezifikationen ist inzwischen in Form der ISO-Norm 26022 [36] beschrieben.

Jede der 10 verschiedenen Strecken des LCT ist ca. 3 000 m lang und beginnt vor oder in einer Kurve. Sobald er diese Kurve verlässt, muss der Benutzer bis zum Ende einer Strecke das Gaspedal durchdrücken und mit Maximalgeschwindigkeit (60 km/h) die Fahrt absolvieren. Damit muss er dieser Dimension während der Fahrt keine Aufmerksamkeit mehr schenken. Bis auf die anfängliche Kurve hat der Rest der Strecke einen geraden Verlauf mit drei Fahrspuren und einer Streckenbreite von ca. 3,85 m. Die Fahrt auf jeder Strecke nimmt ca. 3 Minuten in Anspruch. Die zentrale Aufgabe besteht für den Probanden darin, auf die in Abbildung 2.2 dargestellten weißen Schilder zu reagieren, die durch ein Symbol anzeigen, auf welche der drei Fahrspuren gewechselt werden muss. Dieser Fahrspurwechsel muss erfolgen, sobald die Wechselanweisung wahrgenommen wurde, d.h. nicht erst nachdem ein Schild passiert wurde. Insgesamt gilt es während der Fahrt auf 18 dieser Schilder zu reagieren. Alle 140 bis 188 m (exponential)verteilt findet sich ein solches Schild, der mittlere Abstand zwischen zwei Schildern beträgt 150 m. Das bedeutet, dass ein Fahrspurwechsel im Mittel alle neun Sekunden nötig ist. Jeder der sechs möglichen Fahrspurwechsel (von Spur 1 auf Spur 2 bzw. 3, von Spur 2 auf Spur 1 bzw. 3 und von Spur 3 auf Spur 1 bzw. 2) ist dreimal pro Strecke durchzuführen. Um Lerneffekte zu vermeiden, sind die Schilder auf jeder Strecke anders angeordnet.

Die Durchführung des LCT sieht zu Beginn eine Testfahrt vor, auf der sich ein Benutzer mit der Steuerung und den notwendigen Fahrspurwechseln vertraut machen kann. Danach wird eine Baseline- oder Referenzfahrt ohne Nebenaufgaben durchgeführt und die Fahrleistung aufgezeichnet. Anschließend wird eine weitere Fahrt gestartet, dabei müssen aber neben der primären Fahraufgabe auch Nebenaufgaben durchgeführt werden. Der Versuchsleiter gibt dem Benutzer dazu hintereinander kurze Anweisungen, die dieser durchzuführen hat. Der LCT erlaubt es dabei mit einem einfachen Tastendruck den Beginn und das Ende einer Nebenaufgabe zu markieren. So lassen sich später Erkenntnisse aus der aufgezeichneten

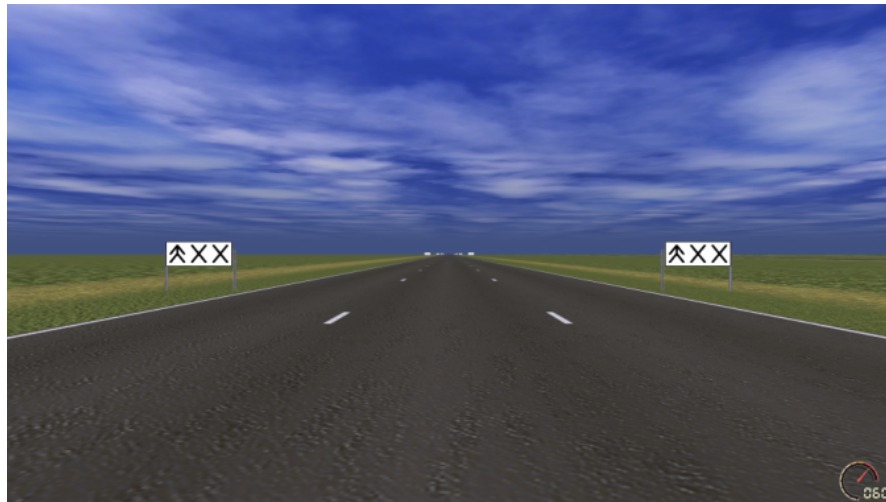


Abbildung 2.2.: Fahrspurwechselanweisung während des Lane-Change-Tests

Fahrleistung direkt einer bestimmten Nebenaufgabe zuordnen.

Zur Bewertung der Fahrleistung wird ein einfaches normatives Modell benutzt (siehe Abbildung 2.3). Die Abweichung zwischen diesem normativen Modell und der tatsächlichen Fahrleistung zeigt die Fahrqualität an. Diese wird durch die Wahrnehmung des Probanden (späte Wahrnehmung oder Verpassen eines Schildes), Reaktionsschnelle, Manövrier-Verhalten (langsame Fahrspurwechsel führen zu größerer Abweichung) und Spurhaltung beeinflusst.

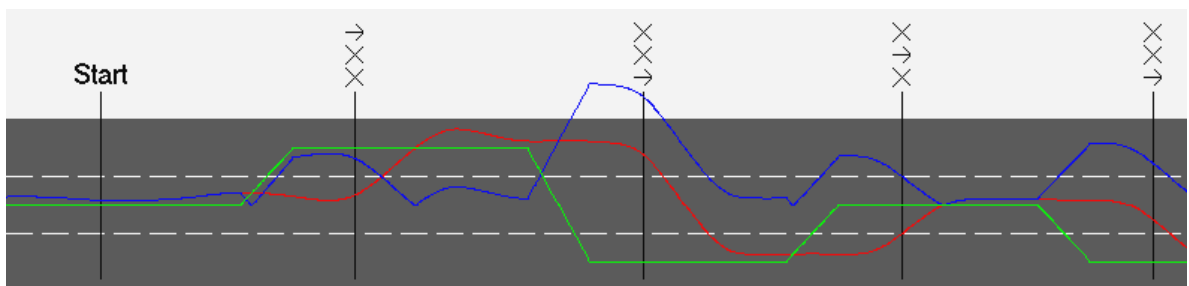


Abbildung 2.3.: Abweichung (blaue Kurve) der tatsächlichen Fahrleistung (rote Kurve) von einem normativen Modell (grüne Kurve) als Maß für die Fahrerablenkung

Configurable Automotive Research Simulator

Der Configurable Automotive Research Simulator (CARS) ist ein Open Source-Fahrsimulator, mit dessen Hilfe Fahrerablenkung gemessen werden kann [39, 63]. CARS wurde darauf ausgelegt, frühe Phasen des Entwicklungsprozesses von im Fahrzeug eingesetzten Systemen zu unterstützen. Im Gegensatz zum Lane-Change-Test erlaubt es CARS, eigene Strecken, je nach den spezifischen Anforderungen, anzulegen. CARS benötigt wie der LCT einen PC mit Monitor, ein Spiele-Lenkrad und Fußpedale (Gas und Bremse). Während der Fahrt werden die Straße, eventuelle Straßenmarkierungen und eine grüne Umgebungslandschaft angezeigt (siehe Abbildung 2.4).



Abbildung 2.4.: Ansicht der Simulator-Komponente des Configurable Automotive Research Simulator

Ein Karten-Editor mit graphischer Benutzeroberfläche erlaubt den Entwurf individueller Simulator-Karten. Dazu können verschiedene Straßenelemente wie Kurven, Geraden oder Kreuzungen, aber auch Vorfahrts- oder Geschwindigkeitsschilder eingebaut werden. Der Editor, wie in Abbildung 2.5a gezeigt, erlaubt ein schnelles und intuitives Erstellen einer Karte.

Für die Analyse der Fahrleistung werden während einer Fahrt die Koordinaten in der virtuellen Umgebung, die Geschwindigkeit, Zeit, Lenkrad- und Pedal-Position, Bremsaktivitäten und Abweichung von der Fahrspur mit einer Frequenz von 20 Hz gespeichert. Ein Analyse-Tool (siehe Abbildung 2.5b) berechnet über diese Daten verschiedene statistische Werte wie den mittleren Abstand zur Fahrlinie, die mittlere Geschwindigkeit, die Frequenz der Geschwindigkeits- und Steuerkorrekturen sowie weitere Werte. Mit diesen Ergebnissen lassen sich beispielsweise Simulator-Fahrten mit und ohne Nebenaufgaben vergleichen und daraus Schlüsse über die entstandene Fahrerablenkung ziehen.

2. Grundlagen und verwandte Arbeiten

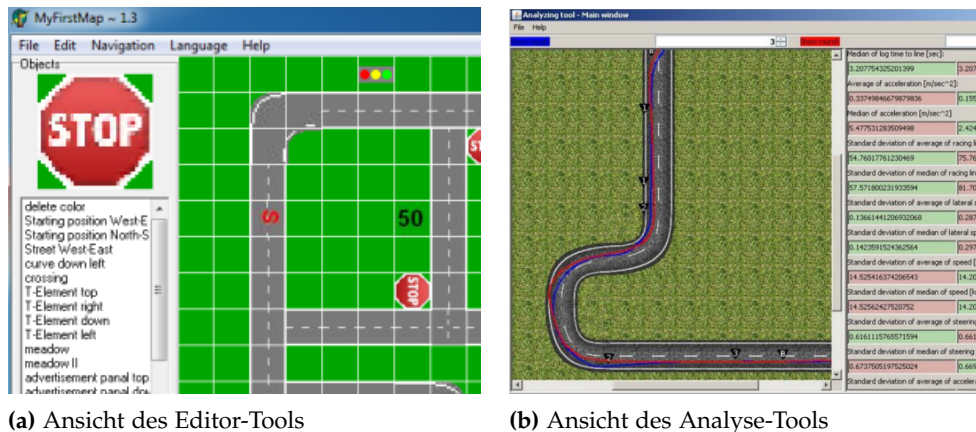


Abbildung 2.5.: Der Karten-Editor und das Analyse-Tool des Configurable Automotive Research Simulator (Quelle: [39])

2.2. Interaktionsformen im Automobil

Interaktion im Automobil findet in unterschiedlichen Bereichen und in diesen wiederum mit unterschiedlichen Modalitäten statt. In den folgenden Abschnitten werden zuerst die Interaktionsbereiche lokalisiert und verschiedene Ein- und Ausgabemodalitäten betrachtet. In den Abschnitten 2.2.4 und 2.2.5 werden dann die für diese Arbeit besonders interessanten Modalitäten Sprache und Gestik gesondert diskutiert. Abschnitt 2.2.6 erörtert abschließend die Potentiale multimodaler Interaktion.

2.2.1. Interaktionsbereiche

Geiser [23] teilt die komplexe Fahraufgabe in drei Klassen auf. Zur primären Aufgabenklasse zählt das Manövrieren des Fahrzeuges durch den Straßenverkehr. Dazu gehört z.B. das Kontrollieren der Fahrzeuggeschwindigkeit oder die Abstandskontrolle und -einhaltung zu anderen Verkehrsteilnehmern durch Einsatz der Pedale und des Lenkrads. Die sekundäre Aufgabenklasse umfasst obligatorische Funktionen, die die Sicherheit des Fahrers und der Verkehrsumgebung erhöhen. Dies betrifft beispielsweise das Betätigen des Blinkers oder das Einschalten des Scheibenwischers um eine ausreichende Sicht auf die Straße zu gewährleisten. Zur letzten Klasse, der tertiären Aufgabenklasse, zählen sämtliche Funktionen die Entertainment-, Kommunikations- und Informationssysteme wie z.B. die Radiosteuerung oder die Bedienung des Navigationssystems betreffen.

Nach dieser von Geiser vorgeschlagenen Aufteilung der im Fahrzeug durchzuführenden Funktionen, haben Tönns et al. [60] eine Zuteilung von Eingabe- und Ausgabeelementen zu diesen Funktionsklassen vorgenommen. Abbildung 2.6 illustriert diese aufgestellte physische Abgrenzung der drei Aufgabenklassen.



Abbildung 2.6.: Physische Aufteilung des Fahrzeugcockpits nach Aufgabenbereichen (Quelle: [60])

Nach Tönns et al. [60] werden Funktionen der primären Aufgabenklasse mit Eingabegeräten bedient, die in bequemer Reichweite von Armen und Beinen liegen. Dies betrifft Eingabegeräte, die einen direkten Einfluss auf die Richtung und Geschwindigkeit des Fahrzeugs haben. Zu diesen Eingabegeräten zählen die Fußpedale sowie das Lenkrad. Diese Geräte werden normalerweise eins zu eins auf ihre Funktionalität abgebildet. D.h., es besteht keine Möglichkeit, dass versehentlich ein nicht gewünschter Eingabemodus aktiv ist. Des Weiteren müssen diese Geräte ein sofortiges haptisches Feedback auf die Benutzung geben.

Funktionen der sekundären Aufgabenklasse sollten ebenfalls mehrheitlich in einer gut zu erreichenden physischen und sichtbaren Entfernung vom Fahrer liegen. Diese Funktionen lassen sich am besten hinter dem Lenkrad, unter der Windschutzscheibe positionieren. Steuerelemente für sekundäre Funktionen sind demnach u.a. Blinker- oder Scheibenwischerhebel. Auch wenn sekundäre Funktionen weniger wichtig sind als die Funktionen der primären Aufgabenklasse, die das Fahren an sich gewährleisten, sollten sie dennoch leicht zugänglich und zu verstehen sein.

Tertiäre Interaktionselemente für Entertainment- und Informationsfunktionen sind eher in der Mitte des Fahrzeugcockpits gelegen. Sie werden oftmals auf der Mittelkonsole positioniert, da ein sofortiger und direkter Zugriff nicht relevant ist, so aber dennoch für den Fahrer erreichbar sind und die Primär- und Sekundärfunktionen nicht stören. Bei immer neuen tertiären Funktionen und deren entsprechender Vielzahl, können Interaktionselemente nicht

eins zu eins auf die Funktionen abgebildet werden.

Häufig eingesetzte Funktionen der tertiären Aufgabenklasse können für einen schnelleren Zugriff auch in die Bereiche der primären und sekundären Aufgaben verlegt werden. So lassen sich z.B. Steuerelemente für die Radiokontrolle direkt auf das Lenkrad auslagern. Der auf der Lenkradvorder- und -rückseite verfügbare Platz wird inzwischen generell zur Kontrolle von Funktionen aller drei Aufgabenklassen verwendet [40].

2.2.2. Eingabemodalitäten

Kern und Schmidt [40] spannen einen Entwurfsraum für automobiler Benutzerschnittstellen auf, der sich von Entwurfsräumen für gängige Benutzerschnittstellen (siehe z.B. [13]) in der Form unterscheidet, dass alle Steuerelemente fest im Fahrzeug angebracht sind und dementsprechend die Positionierung dieser Elemente von entscheidender Bedeutung ist. Des Weiteren muss die Einschränkung in der Bewegungsfreiheit des Fahrers beachtet werden, der physische Interaktionen nur mit den Händen oder Füßen durchführen kann.

Der von Kern und Schmidt [40] vorgeschlagene Entwurfsraum beschreibt mehrere Eingabemöglichkeiten, die im Folgenden zusammengefasst werden.

- Die Gruppe der **Schaltflächen, Knöpfe und Tasten** umfasst die in verschiedenen Formen am häufigsten eingesetzten Eingabeelemente. Sie lässt sich unterteilen in Elemente, die kein dauerhaftes haptisches Feedback über ihren Zustand geben und mechanische Elemente, die ein solches Feedback ermöglichen. Elemente, die kein dauerhaftes haptisches Feedback geben, zeigen ihren Zustand oft über eine visuelle Rückmeldung, z.B. per Schaltflächenbeleuchtung. Bei mechanischen Elementen hingegen kann der Zustand z.B. über einen eingerasteten Knopf ohne Blickkontakt „erfüht“ werden.
- **Schieberegler** werden beispielsweise eingesetzt um die Luftrichtung der Luftdüsen einzustellen.
- **Drehregler** lassen sich in kontinuierlicher und diskreter Eingabe unterscheiden. Eine kontinuierliche Umsetzung der Bedienung dieser Steuerelemente eignet sich z.B. für die Einstellung der Radio-Lautstärke, eine diskrete für die Einstellung der Innenraumtemperatur.
- Lenkrad-**Hebel** eignen sich beispielsweise für die Bedienung der Scheibenwischerfunktionen oder der Blinker.
- **Daumenräder** für z.B. die Lautstärkekontrolle von Entertainmentssystemen können sich auch direkt auf dem Lenkrad wiederfinden, wenn diese eigentlich tertiäre Funktion für einen schnelleren Zugriff in den primären Interaktionsbereich ausgelagert wird, wie in Abschnitt 2.2.1 beschrieben.

- **Fußpedale** bestehend aus Gaspedal, Bremse und in Fahrzeugen mit manuellem Getriebe, der Kupplung.
- **Multifunktionale Steuerelemente**, die Drehreglern ähneln und dementsprechend gedreht aber auch gedrückt werden können, integrieren mehrere tertiäre Funktionen, die aufgrund ihrer zunehmenden Zahl nicht eins zu eins auf Steuerelemente abgebildet werden können. Zum Einsatz kommen solche multifunktionalen Steuerelemente z.B. im BMW iDrive [30, 73] (siehe Abbildung 2.7b) oder Mercedes-Benz COMMAND-System [30], die direkt mit einem zentral angebrachten Bildschirm verbunden sind wie auf Abbildung 2.7a zu sehen.
- **Touchscreens** finden sich ebenfalls im Bereich der tertiären Aufgabenklasse, d.h. für Entertainment-, Komfort- und Infotainmentsysteme.
- **Sprach- und Gesteneingabe** siehe Abschnitte 2.2.4 und 2.2.5.



(a) iDrive Control Display



(b) iDrive Controller

Abbildung 2.7.: Oberhalb der Mittelkonsole angebrachter Bildschirm und Steuerelement des BMW iDrive (Quelle: [10])

2.2.3. Ausgabemodalitäten

Die menschlichen Sinne können Ausgaben in visueller, auditiver, haptischer und olfaktorischer Form [40] wahrnehmen. Bezüglich visueller Ausgabe besteht die Einschränkung darin, dass die visuelle Aufmerksamkeit auf der primären Fahraufgabe liegen muss und darum nur wenig ablenken darf. Dementsprechend finden sich die wichtigsten Anzeigen im Bereich der sekundären Aufgaben [60] (siehe Abbildung 2.6). Dies umfasst z.B. die Geschwindigkeits- oder Tankstandsanzeige. Diese Anzeigen können entweder auf analoger oder digitaler Basis beruhen. Weitere Rückmeldungen über Funktionen wie Blinker oder Abblendlicht werden

2. Grundlagen und verwandte Arbeiten

über Lämpchen oder LEDs gegeben.

(Touch-)Bildschirme werden für das Feedback von hauptsächlich tertiären Funktionen verwendet und müssen unter Beachtung gewisser Einschränkungen positioniert werden. So dürfen sie nicht mehr als 30 Grad unter der horizontalen Sichtlinie des Fahrers liegen [16]. Die Möglichkeiten beschränken sich deshalb auf die Anbringung direkt hinter dem Lenkrad oder im tertiären Aufgabenbereich (siehe Abbildung 2.6). Doch auch im primären Aufgabenbereich hat sich diese Ausgabemöglichkeit niedergeschlagen. Semi-immersive Head-Up-Displays wie in Abbildung 2.8 gezeigt, geben Informationen wie Geschwindigkeit oder Navigationshinweise direkt auf der Windschutzscheibe aus. Der Vorteil von Head-Up-Displays liegt in den kürzeren Blickabweichungen von der Straße, da die Augen nur minimal bewegt werden müssen [60].



Abbildung 2.8.: Head-Up-Display auf der Windschutzscheibe (Quelle: [10])

Auditive Ausgabe über Lautsprecher entlastet den visuellen Aufnahmekanal und kann z.B. in kritischen Situationen zum Einsatz kommen um Alarmsignale auszugeben. Diese auditiven Ausgaben weisen den Fahrer auf sicherheitskritische Ereignisse hin, so kann ein Einparkassistent akustische Warntöne ausgeben, sobald ein Abstandssensor Alarm schlägt [60]. Auditive Hinweise ziehen zwar sofort die Aufmerksamkeit des Fahrers auf sich, jedoch besteht die Gefahr, dass der Fahrer entweder zwischen verschiedenen Tönen nicht unterscheiden kann bzw. ein Warnhinweis von Motorenlärm, Radiogeräuschen oder Stimmen von anderen Fahrzeuginsassen überlagert wird [56].

Die auditive Ausgabe kann auch in Form von sprachbasierter Kommunikation stattfinden. Dies kann dann von Nutzen sein, wenn große Informationsmengen überbracht werden müssen [60]. Dies findet vor allem im tertiären Funktionsspektrum anklang, wenn beispielsweise

eingehende SMS- oder E-Mail-Nachrichten wiedergegeben werden sollen. Für Alarmhinweise ist Sprachausgabe allerdings weniger geeignet, da das Ende der Wiedergabe abgewartet werden muss und für kritische Situationen, in denen eine schnelle Reaktion nötig ist, somit ungeeignet ist. Zur technischen Umsetzung der Sprachausgabe kann dabei entweder die Ausgabe im Voraus aufgezeichnet und dann wiedergegeben werden oder über Text-To-Speech-Funktionalität (TTS) erfolgen [61]. Damit wird eine Sprachsynthese beschrieben, die eine menschliche Stimme künstlich erzeugt und so beliebige Zeichenketten akustisch wiedergeben kann (siehe z.B. [59]).

Durch haptische Ausgabe kann der Fahrer beispielsweise über Vibrationen auf dem Lenkrad oder am Fahrersitz Warnhinweise erhalten wenn die Fahrspur überschritten und der Blinker nicht gesetzt wurde [15, 40]. Die in Abschnitt 2.2.2 auf Seite 24 genannten mechanischen Knöpfe sind wie dort angegeben ebenso ein Beispiel für haptisches Feedback.

2.2.4. Sprachinteraktion

Sprachinteraktion definiert die Möglichkeit Funktionen im Fahrzeug mit verbalen Kommandos zu steuern. Im Automobil könnte sie eine Alternative zur Bedienung mit physischen Steuerelementen bieten, deren Lokalisierung und Handhabung visuelle Aufmerksamkeit erfordern und damit Fahrerablenkung begünstigen. Der Fahrer kann sich so der primären Fahraufgabe widmen, die Fahrleistung erhöht sich, die mentale Beanspruchung reduziert sich hingegen und die Blicke weg von der Straße nehmen ebenso ab [6]. Gärtner et al. [28] folgern eine bessere Fahrqualität bei der Durchführung von komplexen Aufgaben wie Interaktion mit dem Navigationsgerät oder Mobiltelefon mit Sprach- statt mit manueller Eingabe.

Natürlichsprachliche Formulierungen scheinen bei dieser Interaktionsform geeignet, um den Benutzer kognitiv zu entlasten, der unter Umständen viele Sprachkommandos aus dem Gedächtnis abrufen muss. Eine solche Umsetzung hat allerdings noch keinen Einzug in aktuelle Automobile gefunden. Winter et al. [68] untersuchten diesbezüglich die Art und Weise wie Nutzer ihre natürlichsprachlichen Formulierungen wählen, um ihre Absichten auszudrücken. Sie klassifizierten diese Formulierungen und kamen zu dem Ergebnis, dass Benutzer ähnliche Sprachmuster, die aus einer sehr begrenzten Menge an verschiedenen Wörtern zusammengesetzt waren, oft wiederholten. Die meisten Bestandteile längerer natürlichsprachlicher Formulierungen stellten sich als nicht-interaktionsrelevant heraus. Ein weiteres Ergebnis dieser Untersuchung war, dass Benutzer anhand eingesetzter Sprachmuster und -komplexität unterschieden werden können und sich eine natürlichsprachliche automobiler Spracherkennung so auf einzelne (und im Automobilkontext pro Fahrzeug auch wenige) Nutzer einstellen und damit eine verbesserte Genauigkeit bieten kann.

Technische Probleme erschweren den Einsatz von Spracheingabe, Hataoka et al. [32] beschreiben Interferenzen durch Umgebungsgeräusche, ebenso wirkte sich das eingeschränkte Vokabular negativ auf die Bedienung aus. Chang et al. [14] führen weitere Fallstricke bei der Interaktion zwischen System und Fahrer an. So geben sie das Unwissen der Benutzer an, welche Sprachkommandos in bestimmten Situationen verwendet werden können. Fehlende Undo-Funktionalität zur Korrektur einer Eingabe oder die Rückkehr in einen vorherigen Systemzustand kann Benutzer ebenso Aufmerksamkeit kosten. Feedback sollte je nach Aufgabenbereich individuell passend entworfen werden. So wurden in der von Chang et al. [14] durchgeführten Studie stets alle geäußerten Sprachkommandos vom System verbal wiederholt. Dieses Vorgehen stellte sich bei der Eingabe von Telefonnummern als ungeeignet heraus, wenn das System auch hier jede Ziffer wiederholt und die Benutzer teilweise im Unklaren darüber waren, wie und ob sie ihre Eingabe fortsetzen können. Wenn es (aufgrund technischer Unzulänglichkeiten) zu Fehlerkennungen kommt, müssen diese entsprechend behandelt werden. Eine sprachbasierte Interaktion muss somit den Benutzer darüber in Kenntnis setzen, ob es sein Fehler war und er ein falsches Sprachkommando verwendete oder das System eine Spracheingabe nicht erkannt hat. Damit der Benutzer eigene Eingabefehler minimieren kann, ist es wichtig möglichst intuitive Sprachkommandos bereitzustellen, die leicht aus dem Gedächtnis abgerufen werden können [14].

In einer von Hua und Ng [34] durchgeführten Fallstudie wird die Problematik unnatürlicher Sprachkommandos nochmals unterstrichen. Auch können mehrere Sprachbefehle für dieselbe Funktion die Wahrscheinlichkeit erhöhen, dass sich ein Benutzer an eines dieser korrekten Sprachkommandos erinnern kann. Des Weiteren geben Hua und Ng [34] konkrete Richtlinien für eine benutzerfreundliche Spracherkennungs-Schnittstelle an. So ist z.B. eine breite und flache Hierarchiestruktur der Sprachkommandos wünschenswert und auch visuelles Feedback und Gedächtnisstützen sollten angeboten werden.

Die technische Umsetzung einer Sprachinteraktion sollte so gestaltet sein, dass es ersichtlich ist, ob das Mikrofon aktiviert ist und Spracheingaben aufnimmt [14]. Dazu kann eine Push-To-Talk-Funktionalität (PTT) verwendet werden, die das Drücken eines Knopfes erfordert, der sich direkt auf dem Lenkrad oder der Mittelkonsole befindet und in der Folge das Mikrofon aktiviert [31]. Der Benutzer sollte über die Aktivierung zusätzlich durch ein akustisches Signal oder eine visuelle Anzeige informiert werden. Dies kann auch implizit geschehen, nämlich dann wenn der Benutzer grundsätzlich während der gesamten Spracheingabe den PTT-Knopf gedrückt halten muss [14]. Weinberg et al. [67] schlagen vor, nicht nur einen PTT-Knopf zur Verfügung zu stellen, sondern mehrere (Multi-PTT). Dies hat zum Ziel, ein kontextbezogenes PTT-Verfahren anzubieten, sodass sich der Benutzer direkt in den jeweiligen Spracherkennungskontext wie z.B. Musikwiedergabe schalten kann und so die kognitive Auslastung aufgrund des direkten Zugangs zu einem bestimmten Kontextvokabular reduziert wird.

2.2.5. Gesteninteraktion

Der Duden definiert eine Geste als eine „spontane oder bewusst eingesetzte Bewegung des Körpers, besonders der Hände und des Kopfes, die jemandes Worte begleitet oder ersetzt.“ [9]. Althoff et al. [2] sehen ein Einsatzpotential von Kopfgesten im Fahrzeug darin, dass sie als intuitive Alternative zur Beantwortung von Ja-/Nein-Entscheidungen, die ein System erfordert, eingesetzt werden können. Dazu ist nur ein Kopfnicken oder -schütteln erforderlich. Denkbar wäre z.B. die Annahme eines Anrufs auf dem Mobiltelefon während der Fahrt durch ein Nicken mit dem Kopf. Ein Ablehnen des Anrufes würde nur ein einfaches Kopfschütteln erfordern. Für einfache Gesten mit der gesamten Hand geben Althoff et al. [2] ein nahtloses Umschalten von Radiosendern oder Audioquellen an (siehe Abbildung 2.9). Die Durchführung der Handgesten soll dabei über der Mittelkonsole stattfinden. Die Hand kann dabei z.B. unter Infrarotlichtbestrahlung als hellstes Objekt identifiziert werden. Außerdem könnte man so, im Vergleich zur Interaktion mit Knöpfen, schneller und intuitiver in Untermenüs von Infotainment-Systemen im Fahrzeug blättern. Für kurze Kommandos wie das Springen in ein anderes Untermenü sind Gesten demnach sehr effizient einsetzbar [2].



Abbildung 2.9.: Umschalten von Radiosendern und Musikstücken mit Handgesten (Quelle: [10])

Gesten benötigen zur Durchführung nicht die gesamte Hand, auch einzelne Finger können Gesten vollführen. Dazu eignen sich im Automobilkontext Touchscreens, die sich durch eine intuitive und effiziente Bedienung auszeichnen [21]. Eine von Bach et al. [5] durchgeführte Studie verglich die Bedienung eines Autoradios über normale punkt-basierte Berührung eines Touchscreens (z.B. Berührung eines virtuellen Knopfes) und gestenbasierte Interaktion auf einem Touchscreen. Die Studie zeigte, dass der Einsatz von Touch-Gesten visuelle Ablenkung reduzieren und die Leistung der Fahrer verbessern kann. Ecker et al. [21] untersuchten Möglichkeiten der Optimierung eines solchen gestenbasierten Interaktionsstiles für ein

2. Grundlagen und verwandte Arbeiten

Navigationssystem. Sie zeigten, dass visuelles Feedback beim Einsatz von Touch-Gesten Ausführungszeiten signifikant verringern kann und Benutzer sich dadurch sicherer fühlen.

Interaktion mit Gesten, die direkt auf dem Lenkrad stattfindet, wurde u.a. von González et al. [25] umgesetzt, die ein kleines Touchpad auf einem Lenkrad anbrachten. Dieses Touchpad wurde an der rechten Lenkradseite befestigt (siehe Abbildung 2.10) und sollte Gesten des rechten Daumes aufnehmen können. Verschiedene Gesteninteraktionstechniken zur Auswahl von Einträgen aus einer Liste wurden hinsichtlich ihrer Eignung untersucht. Dazu zählte z.B. eine Interaktionsform, die der auf einem Laptop-Touchpad nahekommt und bei der mit dem Daumen ein Auswahlzeiger horizontal und vertikal bis zu dem gewünschten Eintrag gezogen wird. Die im Schnitt schnellsten Ergebnisse ließen sich allerdings mit der EdgeWrite-Methode [70] erzielen, bei der Buchstaben auf einem Touchpad mit einer Geste dadurch beschrieben werden, in welcher Reihenfolge die vier Ecken eines Quadrates beim Ziehen eines Fingers berührt werden.



Abbildung 2.10.: Auf dem Lenkrad befestigtes Touchpad zur Gesteneingabe mit dem rechten Daumen (Quelle: [25])

Nicht nur zur Auswahl, auch zur Eingabe kann sich eine Touch-Fläche auf dem Lenkrad eignen. Kern et al. [41] verglichen verschiedene Positionen für die Eingabe von handschriftlichem Text (siehe Abbildung 2.11) und zeigten, dass die Eingabe mit dem Finger auf einem Touchscreen, der direkt auf dem Lenkrad angebracht ist, zum einen von Benutzern gut angenommen wird und zum anderen die Korrektur- und Fehlerhäufigkeit im Vergleich zur handschriftlichen (Touch-)Eingabe auf der Mittelkonsole um 25% niedriger ist.

Döring et al. [20] entwickelten ein Multi-Touch-Lenkrad, das eine gestenbasierte Interaktion mit Infotainmentsystemen direkt auf der Lenkradoberfläche ermöglicht. Dazu wurde in einer ersten Studie ein Gestenset zusammengestellt, wobei sich zeigte, dass es nicht einfach ist, intuitive Gesten zur Auswahl von Funktionen zu finden. In einer zweiten Studie konnte gezeigt werden, dass sich die notwendige visuelle Aufmerksamkeit bei dieser Interaktionsform spürbar reduzieren lässt.

Die Zusammenstellung von benutzerdefinierten Gesten beschreiben Wobbrock et al. [69] für Gesteninteraktion im Bereich des Surface-Computing. Sie gaben Benutzern Funktionen



Abbildung 2.11.: Gestenbasierte, handschriftliche Texteingabe auf dem Lenkrad und der Mittelkonsole (Quelle: [41])

animiert vor, welche dann eine Geste angeben sollten, die ihrer Meinung nach diese Funktion durchführen kann. Die für eine Funktion jeweils am häufigsten eingesetzte(n) Geste(n) wurden dieser später zugewiesen. Wobbrock et al. [69] geben außerdem eine Taxonomie an, mit welcher sich solche Gesten klassifizieren lassen. Benutzerdefinierte Gesten wurden ebenso von Kray et al. [43] gesammelt. In deren Fall sollten mittels Gesten Interaktionen von Mobiltelefonen mit anderen Geräten wie Public Displays durchgeführt werden.

2.2.6. Multimodale Interaktion

Nach Bernsen und Dybkjaer [8] definiert sich ein multimodales, interaktives System dadurch, dass es sowohl Eingaben als auch Ausgaben zulässt und mindestens zwei verschiedene Modalitäten für Eingabe und/oder Ausgabe einsetzt. Bei dieser Definition werden bereits Systeme mit jeweils einer Eingabemodalität und einer Ausgabemodalität (die sich unterscheiden) als multimodal angesehen. Oviatt [53] verlangt mindestens zwei unterschiedliche Eingabemodalitäten als Charakteristik eines multimodalen Systems. Dafür können beispielsweise Sprache, explizite Gesten, Berührung, Blicke, Kopf- und Körperbewegungen in Frage kommen. Diese Eingabemodalitäten werden mit einer multimedialen Systemausgabe koordiniert.

Der Vorteil von multimodalen Benutzerschnittstellen liegt darin, dass die Schwächen von einzelnen Interaktionsmodalitäten reduziert werden können und sich die Vorteile dieser Einzelmodalitäten verbinden lassen. Ein Beispiel für eine solche vorteilhafte Verknüpfung ist der Einparkassistent von BMW. Die erste Version, die ausschließlich auf auditiven Warnsignalen basierte war nicht wirkungsvoll genug, um den Abstand zu möglichen Hindernissen an den Fahrer zu übermitteln. Deswegen wurde eine visuelle Komponente hinzugefügt, die das Fahrzeug und seine durch Sensoren erfasste nahe Umgebung aus der Vogelperspektive

anzeigt und Hindernisse farblich darstellt [60]. Dies führte zu einer deutlichen Verbesserung des Systems.

Müller und Weinberg [50] beschreiben drei Methoden, wie einzelne Modalitäten im Automobilkontext verknüpft werden können.

Eine Umsetzungsmöglichkeit sind *zeitlich hintereinander folgende Modalitäten*. Dies bedeutet, dass zwei oder mehr Modalitäten in einer bestimmten zeitlichen Reihenfolge nacheinander ausgeführt werden und der dabei von der jeweils vorangehenden Modalität definierte Zustand die Interpretation der nachfolgenden Modalität beschränkt [53]. Müller und Weinberg [50] geben als Beispiel für diese Methode die Verbindung von Sprache und manueller Interaktion an. So kann mit dem Sprachbefehl „*vorderes rechtes Fenster*“ der Interaktionskontext auf dieses spezielle Fenster eingengt und daraufhin mit dem Drücken eines Knopfes dieses Fenster geöffnet oder geschlossen werden. Ein weiteres Beispiel das in aktuellen Fahrzeugen umgesetzt wird ist die Nutzung von kontextbezogenen Sprachbefehlen. Der Kontext wird zuvor z.B. durch den Sprachaktivierungsknopf für das Telefonbuch eingeschränkt. In der Folge sind nur Sprachkommandos zulässig, die auf die Interaktion mit den Kontaktdaten ausgelegt sind.

Redundante Modalitäten sind eine spezielle Form zeitlich hintereinander folgender Modalitäten. Sie gestatten bei jedem Interaktionsschritt die freie Auswahl jeder verfügbaren Modalität [50]. So kann jeder Benutzer nach seinen eigenen Wünschen und Bedürfnissen entscheiden, mit welcher Modalität er einen Interaktionsschritt durchführt. Ein Fahrer kann die Wahl der einzusetzenden Modalität nach den aktuellen Verkehrsgegebenheiten treffen, d.h. beispielsweise bevorzugt Sprache in Situationen einsetzen, bei denen beide Hände besser auf dem Lenkrad verweilen sollten. Reale Umsetzungen dieser Methode gestatten z.B. die Auswahl von Menüeinträgen sowohl durch Sprache als auch über einen multifunktionalen Steuerknopf [50].

Fusionierte Modalitäten können in bestimmten Situationen Nachteile der jeweils anderen Modalitäten ausgleichen. So können z.B. im selben Interaktionsschritt zur Bedienung des Mobiltelefons der Name einer anzurufenden Person mit Sprache geäußert und mit einem Finger derselbe Name in die „Luft geschrieben“ [22] werden. Der Nachteil der schlechten Spracherkennungsgenauigkeit bei lauten Geräuschkulissen kann damit von der Gesteneingabe aufgehoben werden, bei dynamisch wechselnden Lichtverhältnissen wird hingegen die optische Gestenerkennung negativ beeinflusst, was wiederum durch die Spracheingabe ausgeglichen werden kann [50]. Im nicht-automotiven Umfeld wurden fusionierte Modalitäten beispielsweise von Bolt [11] eingesetzt um virtuelle Objekte zum einen mit Zeigegesten und zum anderen mit Sprachanweisungen, die Begriffe wie „das“ oder „dort“ enthielten, auf einem Bildschirm zu positionieren.

3. Konzipierung eines multimodalen Systems

In diesem Kapitel wird das gewählte Konzept für die Erstellung eines multimodalen Systems erläutert. Dazu werden zuerst Modalitäten hinsichtlich ihrer Eignung für die Interaktion im Automobil beurteilt, danach die in dieser Arbeit verfolgte Modalitätenkombination und deren Abfolge geschildert.

3.1. Eigenschaften von verbaler, gestikularer und manueller Interaktion

Der primäre Vorteil von Spracheingabe im Automobilkontext ist der, dass die Hände auf dem Lenkrad verweilen und die Augen auf die Straße gerichtet bleiben können. Verbale Unterhaltung zwischen Personen ist ein natürlicher und alltäglicher Vorgang und dem Menschen somit unmittelbar zugänglich. Im Rahmen einer Mensch-Maschine-Kommunikation kann die Bedienung mittels Sprache demnach einfacher lernbar sein und eine hohe Akzeptanz erfahren. Mit Sprache lässt sich z.B. die Menüstruktur des Bordcomputers aufbrechen und Funktionen und Unterfunktionen können direkt angesprochen werden. Diese Aufgaben würden so weniger Aufmerksamkeit als bei manueller Auswahl auf sich ziehen. [27]. Dies führt in der Folge zu einer besseren Fahrleistung und geringerer mentaler Beanspruchung [6]. Da verschiedene Funktionen zunehmend in besagten hierarchischen Menüstrukturen kombiniert werden und das Durchsuchen dieser entweder visuelle oder auditive Ablenkung schafft oder die angesprochene kognitive Auslastung erhöht [40], kann Sprachauswahl demnach hier eine Entlastung nach sich ziehen.

Die Achillesferse der Sprachinteraktion [50] liegt in der Schwierigkeit Feineinstellungen vorzunehmen. So könnte man eine Seitenscheibe des Fahrzeugs mit einem Sprachkommando wie „Fenster öffnen“ oder „Fenster schließen“ problemlos vollständig öffnen oder schließen. Doch möchte man die Seitenscheibe z.B. nur ein kleines Stück öffnen wird es schwieriger oder zumindest umständlich diesen Vorgang mit Sprache zu beschreiben. Überhaupt eine (natürlichsprachliche) Beschreibung für solch einen Vorgang aus dem Gedächtnis zu formulieren würde bei bestimmten Funktionen eine höhere kognitive Auslastung nach sich

ziehen [50]. Und auch die Technologie der Spracherkennung an sich kann dann von Nachteil sein, wenn Umgebungsgeräusche wie die des Motors oder Radios diese beeinflussen [32]. Technische Einschränkungen der automatischen Spracherkennung bringen gerade bei einem genügend großen Vokabular Fehlerkennungen mit sich. So kann ein Spracherkennner beispielsweise bei der Abgrenzung zwischen akustisch ähnlich klingenden Wörtern in Schwierigkeiten geraten [4]. Dem Benutzer wiederum könnte selbiges widerfahren, falls ein inkonsistentes oder irreführendes Befehlsset vorliegt [14]. Zu tiefe oder komplexe Dialogstrukturen können dem Benutzer unnötig Arbeit bereiten oder gar verwirren, vor allem wenn ein Sprachkommando in einem bestimmten Kontext nicht erkannt, andere hingegen erkannt werden [14, 34]. Da ebenfalls aufgrund dieser technischen Einschränkungen eine limitierte Menge an formelhaften Befehlen natürlichsprachlichen Kommandos vorgezogen werden, müssen Spracheingaben oft stückweise nacheinander gemacht werden, statt alles in einem Kommando zu formulieren [29].

Wie Sprachinteraktion ist auch Gesteninteraktion in der Lage die Bequemlichkeit und Fahr-sicherheit positiv zu beeinflussen, da auch hier der Fokus der Augen auf der Straße verweilen kann [20]. Laute Umgebungsgeräusche beeinflussen die Gesteneingabe im Vergleich zur Spracheingabe nicht. Analoge Einstellungen, die das Erhöhen oder Reduzieren eines Wertes oder Pegels erfordern, lassen sich mit kontinuierlichen Gesten bequem ausführen [55].

Individuelle Gesten und Gestensets sind schwer zu entwerfen, da sie wie Spracherkennung mit technischen Unzulänglichkeiten bei der Erkennung zu kämpfen haben. Des Weiteren müssen auch Gesten von Benutzern erlernt werden und mit möglichst geringer kognitiver Auslastung aus dem Gedächtnis abgerufen werden können [44].

Manuelle konventionelle Interaktion bietet dem Benutzer eine vertraute Interaktionsform bei der er durch Verwendung von Hebeln, Reglern und Knöpfen Einstellungen intuitiv und fein durchführen kann. So kann beispielsweise ein Seitenfenster stufenlos geschlossen oder geöffnet werden [50]. Erkennungsschwierigkeiten oder -fehler wie sie bei Sprach- oder Gesteneingabe anfallen und entsprechende Notwendigkeit zur Fehlerkorrektur gibt es in diesem Sinne bei manueller Bedienung nicht. Für Aufgaben, die schnell und vor allem präzise durchzuführen sind, kann manuelle Interaktion somit die bessere Alternative bilden [27].

Um manuelle Einstellungen vornehmen zu können, muss zuallererst der Bedienhebel, -knopf, Drehregler oder Touch-Button lokalisiert werden. Gerade in Situationen, bei denen man mit dem Automobil nicht oder wenig vertraut ist, kann diese Suche die Fahrerablenkung begünstigen [50], weil die Augen kurzzeitig nicht mehr auf den Straßenverkehr fokussiert werden. Bei der Verwendung eines Drehreglers auf der Mittelkonsole muss sich der Benutzer dann auch oftmals noch durch hierarchische Menüs arbeiten, bis er zum gewünschten Interaktionskontext gelangt ist und dort eine Einstellung vornehmen kann [50]. Im Falle

eines Drehreglers auf der Mittelkonsole oder anderen Bedienelementen, die sich nicht auf dem Lenkrad befinden, muss zwangsläufig eine Hand vom Lenkrad genommen werden, was sich negativ auf die Fahrsicherheit auswirken kann.

3.2. Multimodales Konzept dieser Arbeit

Das Konzept dieser Arbeit beruht auf der Kombination von Sprach- und Gesteninteraktion im Automobil. Die Modalität *Sprache* wird dabei in einem ersten Schritt eingesetzt, um ein Objekt und seine zu manipulierende Funktion auszuwählen. Nach der Auswahl des Interaktionsobjektes bzw. der Funktion ändert sich die Eingabemodalität von *Sprache* zu *Gestik*. Durch den Einsatz von Gesten können nun die Parameter der Funktion eingestellt werden (siehe Abbildung 3.1).

Die Integration dieser beiden Modalitäten soll die in Abschnitt 3.1 geschilderten Nachteile der einzelnen Modalitäten verringern oder beseitigen indem deren Vorteile kombiniert werden. Des Weiteren sollen verschiedene Usability-Gesichtspunkte berücksichtigt werden, die die Akzeptanz bei Benutzern begünstigen. Shneiderman's Goldene Regeln [57], Nielsen's Zehn Usability Heuristiken [51] oder die DIN EN ISO 9241 - 110 Normen [18] sind nur einige von vielen Usability-Richtlinien.

Das Sichtbarkeitsprinzip lässt sich dabei durch die Verwendung von Sprache dadurch umsetzen, dass der Benutzer die sichtbaren Objekte im Fahrzeug und damit deren Funktionen direkt mit einem Sprachkommando ansprechen kann. Dies begünstigt das Abrufen von Sprachbefehlen aus dem Gedächtnis, da dem Benutzer das zu verwendende Vokabular direkt vor Augen gehalten wird. Mit *Sprache* lassen sich so einzelne Kommandos aus einer großen Befehlsmenge verwenden, ohne dass man beispielsweise durch Menüstrukturen navigieren muss. Bei Objekten, die nur eine einzige Funktion anbieten, genügt es bereits nur das Objekt mit einem Sprachkommando zu benennen und damit implizit die einzige

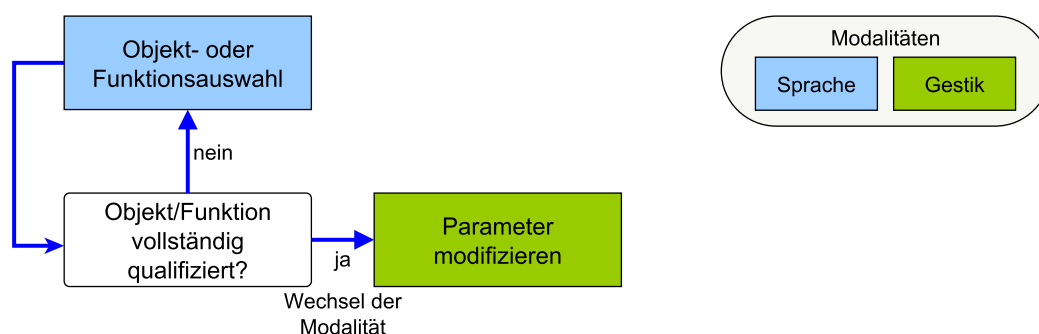


Abbildung 3.1.: Konzept der multimodalen Interaktion mit Sprache und Gesten

3. Konzipierung eines multimodalen Systems

Funktion auszuwählen. So wäre die Auswahl „Tempomat“ bereits vollständig qualifiziert: die Funktion wäre damit implizit die Festlegung der Tempomat-Geschwindigkeit. Für den Fall, dass ein Objekt mehr als eine Funktion anbietet, der Benutzer beispielsweise das Sprachkommando „Head-Up-Display“ verwendet, reagiert das System und bietet ein direktes Feedback, das den Benutzer darauf hinweist, dass für dieses Objekt mehrere Funktionen vorliegen (z.B. Helligkeit oder Position auf der Windschutzscheibe). Eine einfache Fehlerbehandlung erlaubt nun das erneute Formulieren eines (vollständig qualifizierten) Sprachkommandos. Dies geschieht, indem das System per Sprachausgabe, einem kurzen Audiosignal und/oder mit einer visuellen Darbietung der möglichen Funktionen auf den Fehler hinweist und Hilfestellung anbietet.

Falls ein Benutzer ein Objekt auswählen möchte, von dem es mehrere Instanzen im Fahrzeug gibt (z.B. Fensterscheiben), dann muss das Sprachkommando ebenfalls eindeutig formuliert werden (z.B. „Fahrerfenster“). Auch hier gilt, beim Sprachkommando „Fenster“ reagiert das System umgehend mit Feedback und weist auf das unzureichend formulierte Kommando hin. Dies kann auch als generelle Hilfe gesehen werden, denn falls sich der Benutzer an ein konkretes Sprachkommando nicht mehr erinnert, kann er im Falle einer visuellen Darbietung sich z.B. mit dem nicht-eindeutigen Kommando „Fenster“ die gültigen Sprachkommandos bezüglich der Fensterscheiben anzeigen lassen. Dabei ist beim Design des Benutzerinterfaces darauf zu achten, dass diese visuelle Darstellung die Fahrerablenkung nicht grundsätzlich negativ beeinflusst. Falls der Benutzer bereits ein Sprachkommando eingegeben hat, sich aber entscheidet doch eine Funktion auf einem anderen Objekt durchzuführen, dann kann er diese erste Kommandoingabe einfach rückgängig machen, indem er diese „überschreibt“ und ein neues Kommando äußert.

Nachdem ein Objekt und dessen Funktion eindeutig mittels *Sprache* ausgewählt wurde, bekommt der Benutzer ein direktes Feedback darüber, dass das Kommando erfolgreich erkannt wurde. Damit ist die Eingabe mit dieser Modalität abgeschlossen und die Modalität *Gestik* kommt zum Zuge. Durch den Einsatz von Gesten können nun die Funktion ausgeführt und Einstellungen verändert werden. Da mit Spracheingabe Feineinstellungen nur schwer durchzuführen sind, ergibt sich ein Vorteil der Gesteneingabe direkt dadurch, dass eine fein-granulare Interaktion möglich ist. So erlauben kontinuierliche Gesten es z.B. die Intensität einer Luftdüse in kleinsten Schritten zu erhöhen oder zu reduzieren. Dies ist auch ein Beispiel dafür, dass sich Änderungen leicht rückgängig machen lassen, da man nur die Geste zur jeweils konträren Funktion ausführen muss und so eine Einstellung zurücknehmen kann. Der Benutzer bekommt nach Durchführung einer Geste sofort ein Feedback auf die eben vorgenommene Manipulation, da sich der Zustand des Objektes direkt ändert.

Von dieser Modalitätenkombination wird erwartet, dass sie in der Lage ist, den Grad der notwendigen Aufmerksamkeit bei der Interaktion zu verringern und damit die Fahrerablenkung zu reduzieren oder mindestens nicht zu verstärken.

4. Explorative Benutzerstudie

In diesem Kapitel werden die Durchführung und die Ergebnisse einer explorativen Benutzerstudie erläutert, die zum einen die Hypothese validieren sollte, dass es Benutzern leicht fällt, ohne vorheriges Training Objekte und Funktionen im Automobil durch Spracheingabe auszuwählen. Zum anderen sollte die Hypothese validiert werden, dass Benutzer eine ähnliche Erwartungshaltung in Bezug auf die Steuerung mit Gesten haben. Dazu sollte die Benutzerstudie zwei grundsätzliche Fragestellungen beantworten:

- I Welche Sprachkommandos werden von Benutzern geäußert, um bestimmte Objekte und Funktionen auszuwählen?
- II Welche Gesten setzen Benutzer ein, um die zuvor ausgewählte Funktion auf einem Objekt durchzuführen?

Die gesammelten Sprachkommandos und Gesten wurden nach dieser eingehenden Studie ausgewertet und daraus sowohl ein logisches und konsistentes Set aus Sprachkommandos, als auch aus Gesten zusammengestellt. Diese bilden wiederum die Basis für die Erstellung des in Kapitel 5 beschriebenen Prototypen.

4.1. Vorbereitung der Studie

Zur Vorbereitung der Studie mussten zuallererst diejenigen Objekte und Funktionen formiert werden, die sich von den Probanden in der Studie auswählen lassen und deren Parameter manipuliert werden sollten. Hierzu wurden zur Recherche verschiedene Fahrzeughandbücher (u.a. Mercedes-Benz E-Klasse, Baureihe 211, Mercedes-Benz CLS-Klasse, Baureihe C219 und BMW 5er-Klasse, Baureihe F10) herangezogen, um in einem ersten Schritt möglichst viele in Frage kommende Objekte und Funktionen zu erfassen. Primäre Fahrzeugfunktionen (siehe Kapitelabschnitt 2.2.1 auf Seite 22) wurden hier außer Acht gelassen, nur Funktionen aus dem sekundären und tertiären Bereich kamen in die Auswahl.

Es ergab sich ein breites Spektrum an möglichen Objekten und Funktionen, das u.a. durch Anwendung folgender Kriterien und Randbedingungen weiter verfeinert wurde.

- Auswahl von Funktionen, deren Bedienelemente in einer gewissen Distanz zum Fahrer angebracht sind.

4. Explorative Benutzerstudie

- Auswahl von Funktionen, deren Bedienelemente Bewegungen erfordern, die Aufmerksamkeit von der Straße nehmen (z.B. Drehung des Kopfes).
- Auswahl von Funktionen, deren Bedienelemente ein gewisses höheres Maß an kognitiver Auslastung erfordern.
- Auswahl von Funktionen, von denen auszugehen ist, dass sie den Teilnehmern der Studie auch bekannt bzw. vertraut sind.
- Eingeschränkte Auswahl von einfachen Funktionen, deren Zustände nur zwischen ‚Ein‘ und ‚Aus‘ switchen.
- Eingeschränkte Auswahl von Funktionen, bei denen davon auszugehen ist, dass sie in Zukunft durch den Einsatz von Sensoren auch automatisiert, ohne Benutzereinflussnahme, durchgeführt werden können.

Die endgültige Auswahl bestand schließlich aus 9 verschiedenen Objekten, auf denen eine gewisse Zahl an Funktionen durchzuführen war. Um die Anzahl der Aufgaben je Proband in einem gewissen Rahmen zu halten, wurde bei Objekten, die mehrfach im Fahrzeug vorkommen nicht für jedes einzelne ein Sprachkommando abgefragt und für mehrfach vorkommende Objekte auch nur einmal für eine bestimmte Funktion eine Geste gesucht. So wurden beispielsweise statt der vier im Fahrzeug befindlichen Seitenscheiben nur zwei davon in die Aufgabenreihe übernommen. Dies waren die Seitenscheibe vorne rechts und die Seitenscheibe hinten links. Zu ersterer wurden dann Gesten für das teilweise Öffnen und Schließen der Scheibe gesucht, zu letzterer für das vollständige Öffnen und Schließen. Die fehlenden Sprachkommandos sollten sich aus den Ergebnissen für die abgefragten Seitenscheiben ableiten lassen, genauso auch die erschlossenen Funktionsgesten für alle diese Seitenscheiben gelten.

Dieses Vorgehen wurde auch für andere Objekte und Funktionen angewandt die mehrfach im Fahrzeug auftauchen. Tabelle 4.1 gibt einen Überblick über sämtliche für die Benutzerstudie ausgewählten Objekte und Funktionen.

4.2. Aufbau der Studie

Für die Studie, welche in einem Labor durchgeführt wurde, kam ein Logitech G27 Racing Wheel zum Einsatz, von dem das Original-Lenkrad entfernt und durch eine Eigenfertigung ersetzt wurde. In diesem aus Holz gefertigten Lenkrad wurde ein Motorola Xoom Android-Tablet-Computer passgenau in die Oberfläche integriert (Abbildung 4.1). Diese Lenkrad-Konstruktion wurde vor einem 24"-Monitor aufgebaut, auf dem die Probanden den Open Source Fahrsimulator CARS [39] (siehe Kapitelabschnitt 2.1.2 auf Seite 21) präsentiert bekamen.

Objekt	Funktion
Fahrersitz	Sitz nach vorne verschieben Sitz nach hinten verschieben
	Sitz nach oben verstellen Sitz nach unten verstellen
	Sitzlehne nach vorne neigen Sitzlehne nach hinten neigen
	Kopfstütze höher stellen Kopfstütze tiefer stellen
	Sitztemperatur senken Sitztemperatur erhöhen
Außenspiegel rechts	Horizontale Ausrichtung nach links Vertikale Ausrichtung nach oben
Frontscheibenwischer	Abstellen Dauerwischen Intervallwischen
Heckscheibenwischer	Wischen mit Scheibenwischwasser
Seitenscheibe vorne rechts	Teilweise öffnen Teilweise schließen
Seitenscheibe hinten links	Vollständig öffnen Vollständig schließen
Tempomat	Geschwindigkeit reduzieren Geschwindigkeit erhöhen
Luftdüse links	Intensität reduzieren Intensität erhöhen
Head-Up-Display	Display heller Display dunkler

Tabelle 4.1.: In der Benutzerstudie verwendete Objekte und Funktionen

4. Explorative Benutzerstudie



Abbildung 4.1.: Eigengefertigtes Holzlenkrad auf einem Logitech G27 Unterbau und der darin integrierte Android-Tablet-Computer

Das Lenkrad wurde mit dem PC verbunden und die Probanden konnten so mit dem Fahrsimulator interagieren. Abbildung 4.2 gibt den Gesamtüberblick über den Aufbau und die Positionierung der Probanden während der Studie.

Auf dem Tablet-Computer lief eine Android-Applikation, die den Probanden die Aufgaben anzeigte und für die spätere Analyse sowohl Sprachkommandos, als auch Gesten aufzeichnete. Die Sprachkommandos wurden per Audiorekorder aufgezeichnet, für die Gesten wurde jeder einzelne Berührungspunkt eines Fingers in einer Datenbank gespeichert, ebenso die Information wo eine Gestenspur begonnen und wo sie geendet hat. Da der verwendete Android-Tablet-Computer multitouch-fähig ist, waren potentiell Gesteneingaben mit mehreren Fingern möglich.



Abbildung 4.2.: Positionierung der Probanden während der Studie

4.3. Teilnehmer

Das Teilnehmerfeld bestand aus 12 Personen, die über E-Mail-Verteiler oder persönlich angeworben wurden. Die einzige Voraussetzung für die Teilnahme war der Besitz eines Führerscheines. Alle Teilnehmer erhielten eine Studienvergütung in Höhe von 5 € für durchschnittlich ca. 35 Minuten Studiendauer. Die in Tabelle 4.2 aufgelisteten allgemeinen Informationen über die Probanden wurden aus einem eingehenden Fragebogen (siehe Anhang auf Seite 96) entnommen.

Alter	zwischen 20 und 39 Jahre (im Durchschnitt 28,25 Jahre)
Automobilnutzung	6 Teilnehmer gelegentlich, d.h. mehr als zweimal im Monat 4 Teilnehmer täglich 2 Teilnehmer kaum, d.h. maximal zweimal im Monat
Beruf/Studiengang	4 Mitarbeiter der Universität 7 Studenten aus mathematisch-technischen Fachrichtungen 1 Student der Fachrichtung Umweltschutz
Fahrzeugeigner	5 Teilnehmer besaßen ein eigenes Auto
Führerscheinbesitzdauer	zwischen 2 und 22 Jahre (im Durchschnitt 10,58 Jahre)
Geschlecht	10 männliche Teilnehmer 2 weibliche Teilnehmer
Händigkeit	10 Rechtshänder 2 Linkshänder

Tabelle 4.2.: Allgemeine Informationen zu den Teilnehmern der Benutzerstudie

4.4. Ablauf der Studie

Im Folgenden werden die einzelnen Schritte während der Studie jeweils gesondert behandelt, die tatsächliche Reihenfolge während der Studie stimmt mit der Reihenfolge der Abschnitte überein.

4.4.1. Einführung in die Studie und erster Fragebogen

Zu Beginn wurde jeder Teilnehmer in die Thematik der Studie eingeführt, d.h. es gab eine kurze Erklärung des Interaktionskonzeptes, den Sinn und Zweck der Studie sowie erste Hinweise zum Ablauf. Hierbei wurde explizit darauf hingewiesen, dass während der Studie Sprach- und Videoaufzeichnungen gemacht werden. Der Proband musste daraufhin eine zuvor vorgelesene Einverständniserklärung unterschreiben (siehe Anhang auf Seite 95). Danach bekamen die Probanden einen Eingangsfragebogen ausgehändigt, in dem sie allgemeine Fragen zur Person und zu Vorkenntnissen in Hinblick auf Sprach- und Gesteninteraktion beantworten sollten.

4.4.2. Detailliertere Erläuterung der Vorgehensweise

Die Teilnehmer wurden nun vor den Fahrsimulator und das Lenkrad positioniert und ihnen wurde der Ablauf im Detail erläutert. So bestand jede Aufgabe aus einem Paar: Im ersten Aufgabenteil wurde entweder ein Einzelbild oder ein Vorher-Nachher-Bild auf dem Android-Tablet-Computer angezeigt (siehe Abbildung 4.3) und das Objekt oder die Objekteigenschaft/-funktion, die darauf markiert ist, sollte mit einem Sprachkommando eindeutig ausgewählt werden. Auf die Eindeutigkeit eines Kommandos wurde explizit Wert gelegt, da ein Objekt auch mehrfach im Fahrzeug vorhanden sein kann.

Im zweiten Teil eines Aufgabenpaares wurde jeweils eine Geste zur Durchführung der Funktion gesucht, welche auf dem Tablet-Computer auszuführen war. Welche konkrete Einstellung der Funktion durchgeführt werden sollte, gab der Tablet-Computer über eine Audioausgabe bekannt.

Im Anschluss an jedes Aufgabenpaar wurden die Teilnehmer dann noch darum gebeten stets dieselbe Frage zu beantworten: *„Die Kombination aus dem eben genannten Sprachkommando und der Touch-Geste ist während einer realen Autofahrt [...] auszuführen.“* Die Lücke konnte jeweils mit den Antwortmöglichkeiten *„sehr einfach“*, *„einfach“*, *„mittelschwer“*, *„schwer“* oder *„sehr schwer“* beantwortet werden.

Durch die Nutzung von Vorher-Nachher-Bildern sollten die Probanden leichter in die Lage versetzt werden, direkt die Objektfunktionen in Sprachkommandos auszudrücken und nicht nur das Objekt zu benennen. Die Scheibenwischerfunktionen bilden hierbei eine Ausnahme, diese ließen sich nicht mit Vorher-Nachher-Bildern darstellen und wurden deshalb nur in Form von Einzelbildern präsentiert.

Die Methoden der Studiendurchführung sind hierbei auch an Wobbrock et al. [69] zur Erstellung eines benutzerdefinierten Gestensets angelehnt. So wurde auch in dieser Benutzerstudie kein direktes Feedback auf Sprach- und Gesteneingaben der Probanden gegeben. Der Tablet-Computer zeichnete diese lediglich auf. Ebenfalls wurde hier in gleicher Weise die



(a) Beispielaufgabe der Benutzerstudie als Vorher-Nachher-Bild



(b) Einzelbildaufgabe

Abbildung 4.3.: Die zwei verschiedenen Typen von angezeigten Bildern zur Sammlung von Sprachkommandos

Methode des lauten Denkens eingesetzt, d.h. die Teilnehmer sollten ihr Vorgehen, speziell bei der Wahl und Ausführung der Geste, erklären. So konnten die mentalen Modelle der Teilnehmer erfasst und später in der Auswertung herangezogen werden.

Die Probanden hatten auch keine Möglichkeit den Dialogablauf selbst zu steuern. Der Übergang von Aufgabenteil Spracheingabe zu Aufgabenteil Gesteneingabe und dann zum nächsten Aufgabenpaar geschah ferngesteuert von einem zweiten Computer aus. So war z.B. sichergestellt, dass die Probanden nicht versehentlich schon zur Gesteneingabe übergingen, bevor das Sprachkommando überhaupt eindeutig war.

4.4.3. Erklärungen anhand einer Beispielaufgabe

Nach dem Eingangsfragebogen und bevor die richtige Aufgabenreihe begann, hatten die Probanden noch Gelegenheit sich in den Fahr Simulator einzuüben. Im Anschluss darauf wurde ihnen die oben beschriebene Vorgehensweise an einer Beispielaufgabe verdeutlicht. Diese Beispielaufgabe hatte im ersten Aufgabenteil die Auswahl eines Fahrzeuginnenlichtes

mit einem Sprachkommando zum Ziel (siehe Abbildung 4.3a). Dies war ein Vorher-Nachher-Bild, das auf der linken Seite ein helleres Licht zeigte, als rechts. Den Probanden wurde erklärt, dass man hier beispielsweise das Objekt „Deckenlicht“ oder die Objektfunktion „Deckenlichthelligkeit“ als Sprachkommando zur Auswahl verwenden könnte. Auch ganz andere Bezeichnungen könnten möglich sein, so sollten alle Sprachkommandos möglichst spontan aber natürlich eindeutig geäußert werden.

Für den Teil der Gesteneingabe der Beispielaufgabe sollte die Helligkeit dieses Deckenlichtes reduziert werden. Dies wurde wie beschrieben durch eine Audioausgabe kommuniziert. Wichtig war hier der Hinweis, dass es sich um einen Multi-Touch-Tablet-Computer handelt und mehrere Finger eingesetzt werden können. Eine Hand musste allerdings stets auf dem Lenkrad verweilen, da mit dem Fahrsimulator interagiert werden sollte. Da die Teilnehmer sich aber für den ersten Aufgabenteil auf das angezeigte Bild auf dem Tablet-Computer konzentrieren sollten, durften sie hier auch nur geradeaus fahren und die Hindernisse im Fahrsimulator ignorieren. Für die Gesteneingabe wurde aber darauf hingewiesen, dass die Spur gehalten und den Hindernissen ausgewichen werden sollte, d.h. die Gesten dann auszuführen, wenn die Verkehrslage es zulässt.

Gas zu geben war bei der Fahrt mit dem Simulator nicht nötig, es wurde von vornherein eine feste Geschwindigkeit eingestellt, so dass nur Lenken erforderlich war. Es bestand für die gesamte Studie und die Einzelaufgaben auch kein Zeitlimit für die Teilnehmer. Sie hatten ausreichend Gelegenheit, die angezeigten Bilder zu betrachten und die markierten Objekte bzw. Objektfunktionen zu identifizieren und sich danach eine Geste auszudenken. Gerade bei den Vorher-Nachher-Bildern war für eine genaue Identifizierung mitunter eine längere Betrachtung der Bilder notwendig, um die intendierte Funktion bzw. das Objekt zu erkennen, das im Fall der Vorher-Nachher-Bilder in unterschiedlichen Zuständen angezeigt war.

4.4.4. Durchführung der Aufgabenreihe

Nach den einführenden Erläuterungen, dem Beispiel und nachdem keine weiteren Fragen der Probanden mehr offen waren, begann die eigentliche Aufgabenreihe. Aus den Objekten und Funktionen aus Tabelle 4.1 wurden die oben angesprochenen (insgesamt 26) Aufgabenpaare gebildet. D.h. es wurden jeweils Sprachkommandos und Gesten abgefragt und danach eine Bewertung des Schwierigkeitsgrades verlangt. Die Aufgaben wurden in zufälliger Reihenfolge präsentiert. Da die Probanden nicht beeinflusst werden sollten, gab es außer dem angezeigten Bild keine weitere verbale und textuelle Erklärung, um das resultierende Sprachkommando in keine bestimmte Richtung zu drängen. Dies musste im Übrigen bereits beim Design der angezeigten Bilder berücksichtigt werden.

Vor allem zu Beginn fehlte bei einigen Probanden die Eindeutigkeit in ihren Kommandos. So entstanden Situationen wie: Proband: „Fenster!“; Versuchsleiter „Dies ist noch nicht eindeutig

genug, versuche eine genaue Lokalisierung in Dein Sprachkommando einzubauen.“; Proband: „Fenster hinten links!“. Nach nur wenigen Aufgaben hatten die Teilnehmer das Prinzip allerdings verinnerlicht und es musste nur noch selten auf die Eindeutigkeit hingewiesen werden. Die Gesteneingabe wurde wie in Abbildung 4.4 zu sehen, jeweils auf einem weißen Hintergrund ausgeführt. Die Spur der Geste wurde den Teilnehmern nicht angezeigt, waren sie allerdings der Meinung die Geste falsch eingegeben zu haben, konnten sie sie nochmals wiederholen.



Abbildung 4.4.: Gesteneingabe während der explorativen Benutzerstudie

4.4.5. Abschließender Fragebogen

Nach Abschluss der 26 Aufgaben bekamen die Probanden noch einen Fragebogen ausgehändigt, in dem sie die gewählte Modalitätenkombination Sprach-/Gesteneingabe hinsichtlich der Akzeptanz, Vor- sowie Nachteilen einschätzen sollten. Die in der Aufgabenreihe auszuführenden Funktionen mussten jeweils einzeln auf ihre Durchführbarkeit mit genannter multimodaler Interaktionsform beurteilt werden. Nachdem der Abschlussfragebogen ausgefüllt war, bekamen die Probanden die Studienvergütung ausbezahlt und mussten den Empfang quittieren.

Für eine graphische Veranschaulichung des vollständigen Ablaufs der Studie sei auf Abbildung 4.5 verwiesen.

4.5. Auswertung der Studie

Die folgenden Abschnitte erörtern zunächst die Auswertung der Eingangs-, Zwischen- und Abschlussfragebögen. Die anschließenden beiden Abschnitte geben die Ergebnisse der Auswertung der aufgezeichneten Sprachkommandos und der Touch-Gesten wieder.

4. Explorative Benutzerstudie

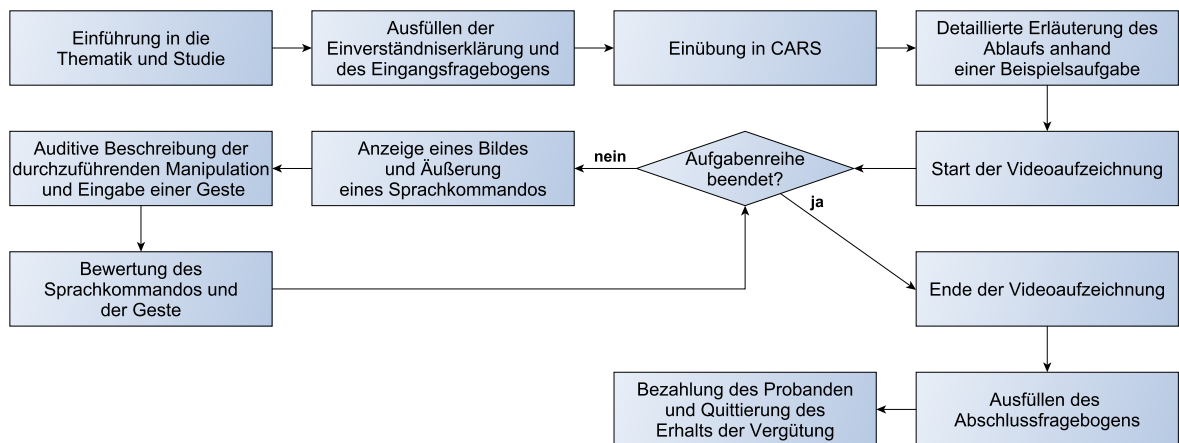


Abbildung 4.5.: Ablaufdiagramm der explorativen Benutzerstudie

4.5.1. Auswertung der Fragebögen

Eingangsfragebögen

Der Eingangsfragebogen erfasste allgemeine Informationen zur Person, die bereits in Tabelle 4.2 auf Seite 41 aufgelistet wurden.

Auch wurden im ersten Fragebogen Vorerfahrungen der Teilnehmer bezüglich Spracheingabe und Spracherkennung abgefragt. Grundsätzlich gaben 4 der 12 Teilnehmer an, überhaupt schon Erfahrungen mit dieser Interaktionsform gemacht zu haben, davon u.a. alle mit Handys/Smartphones.

Erfahrungen mit (Multi)-Touch-Systemen haben hingegen bereits mehr Teilnehmer machen können, so waren es hier 9 von 12 Probanden. Auch hier waren wiederum Handys/Smartphones die Hauptidequationsquelle (8x genannt), gefolgt von Tablet-Computern (2x genannt). Zu den nur einmal genannten Erfahrungsquellen zählten u.a. Bank- und Fahrkartenautomaten, wobei hier davon auszugehen ist, dass wesentlich mehr Probanden Erfahrungen mit der Touch-Bedienung dieser Geräte haben, als angegeben wurde.

Bewertungsfragebögen der individuellen Sprachkommandos und Gesten

Die Auswertung der Bewertungen, die die Probanden nach jedem Aufgabenpaar der Studie abzugeben hatten („Die Kombination aus dem eben genannten Sprachkommando und der Touch-Geste ist während einer realen Autofahrt [...] auszuführen.“) gestaltete sich vor allem durch die zufällige Anordnung der einzelnen Aufgaben als schwierig. Dadurch war jeder Teilnehmer durch jeweils andere vorhergehende Aufgaben beeinflusst und hatte beispielsweise andere Voraussetzungen, um die Konsistenz zu wahren. Auch wenn sich diese Bewertungsfragebögen (siehe Anhang auf Seite 98) aus diesem Grunde nur eingeschränkt untereinander

vergleichen lassen, kann doch ein Trend festgestellt werden. Von 311 abgegebenen Bewertungen überwiegt über alle Probanden summiert, die Bewertung „*einfach*“ (145x) die anderen vier Bewertungsmöglichkeiten. Es folgen „*sehr einfach*“ (94x), „*mittelschwer*“ (49x), „*schwer*“ (19x) und „*sehr schwer*“ (4x). D.h., trotz der unterschiedlichen Reihenfolgen der Aufgaben tendieren die Probanden dazu, den Einsatz von Sprache und Gesten während einer realen Autofahrt als eher einfach einzuschätzen.

Abschlussfragebögen

Im abschließenden Fragebogen waren die Probanden aufgefordert, die zuvor von ihnen eingesetzten Modalitäten bzw. deren Kombinationspotential zu bewerten. Die erste Frage betraf die Einschätzung der allgemeinen Akzeptanz der Kombination Sprachkommandos und Touch-Gesten. Hier konnte zwischen den Antwortmöglichkeiten „*gar keine Akzeptanz*“ bis „*sehr hohe Akzeptanz*“ gewählt werden (siehe Anhang auf Seite 100). Weist man diesen Antwortmöglichkeiten im Rahmen einer Likert-Skala die Werte von 1 (gar keine Akzeptanz) bis 5 (sehr hohe Akzeptanz) zu, dann ergibt sich im Mittel der Wert 3. D.h., die Probanden schätzen die allgemeine Akzeptanz als eher durchschnittlich ein.

Auf die Frage, ob sie für die Eingabe der Gesten die Hand vom Lenkrad genommen haben, antworteten 9 Teilnehmer mit „*ja*“. Darunter ist explizit die zweite Hand zu verstehen, denn eine musste nach Vorgabe stets auf dem Lenkrad bleiben. An diese Vorgabe hielten sich auch sämtliche Teilnehmer.

Alle Probanden gaben an, einen oder mehrere Finger bevorzugt für die Gesten genutzt zu haben. Dies waren drei unterschiedliche Finger: Zeigefinger (8x genannt), Daumen (5x) und der Mittelfinger (2x).

Was bereits während der Studie durch die Betrachtung der Teilnehmer, ihres lauten Denkens und damit ihrer mentalen Modelle auffiel, schlug sich auch im Abschlussfragebogen nieder: 11 Probanden gaben an, auf Konsistenz in Bezug auf zuvor eingegebene Gesten geachtet zu haben. Wurde beispielsweise für die Erhöhung der Geschwindigkeit des Tempomaten eine Richtungsgeste nach oben ausgeführt, so wurde für die Reduzierung dieser Geschwindigkeit die entgegengesetzte Richtungsgeste nach unten eingesetzt. Selbst bei den komplexeren Funktionen für die Bedienung des Scheibenwischers versuchten viele Teilnehmer möglichst zusammenpassende Gesten zu finden. Dabei ist anzumerken, dass die Reihenfolge der Aufgaben zufällig war und die Einhaltung der Konsistenz auch davon abhing, welche Aufgaben zuvor bereits mit Gesten zu beantworten waren. So fiel dies den einen schwerer, den anderen leichter.

4 der 12 Probanden würden für die Gesteneingabe maximal einen Finger einsetzen wollen, mehr als fünf Finger konnte sich keiner vorstellen. Im Durchschnitt würden sie maximal 2,42 Finger für eine Geste einsetzen wollen.

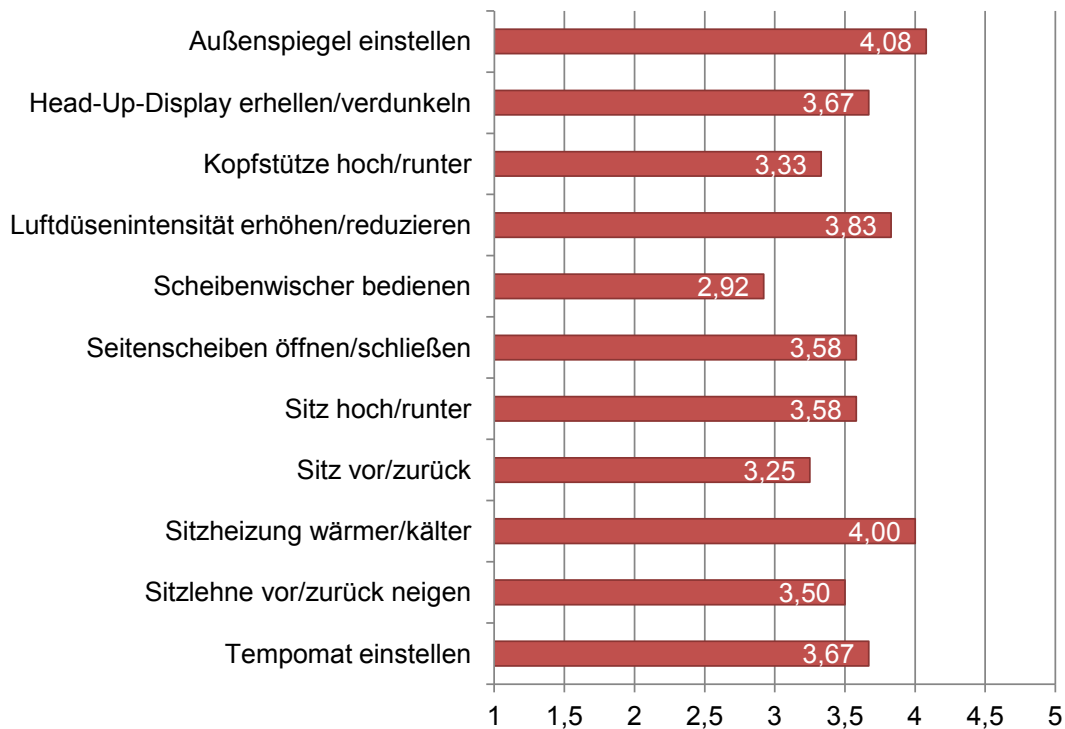
4. Explorative Benutzerstudie

Nachdem die einzelnen Probanden bereits während der Studie ihre Meinungen zu einzelnen Funktionen und den eingesetzten Modalitäten mitteilten, konnten sie dies nun in schriftlicher Form nochmals genauer fixieren. So sollten sie nun Stellung zu folgender Aussage nehmen: *„Ich halte eine Bedienung über die Kombination Sprachkommando und Touch-Geste für diese Aufgabe für sinnvoll.“* Auch hier standen wiederum fünf Antwortmöglichkeiten in Form einer Likert-Skala zur Auswahl, denen jeweils wieder ein Wert zugeordnet wurde. *„stimme voll und ganz zu“*(5), *„stimme zu“*(4), *„unentschieden“*(3), *„stimme nicht zu“*(2) und *„stimme gar nicht zu“*(1). Abbildung 4.6 zeigt das Ergebnis dieser Beurteilung durch jeweilige Angabe des Mittelwertes zu den Funktionen aus der Studie. Die höchsten mittleren Bewertungen erhielten hierbei die Funktionen zum Einstellen des Außenspiegels und der Sitzheizung. Dies kann so interpretiert werden, dass die Probanden hier eher zustimmen würden (*„stimme zu“*(4)), dass die Bedienung dieser Funktionen mit Sprachkommandos und Gesten sinnvoll wäre. Zu erkennen ist ebenfalls, dass die Probanden im Mittel keine Funktion zur Bedienung mit Sprache und Gesten explizit als nicht sinnvoll erachteten. Sie waren mindestens eher unentschieden. Zurückhaltender wurden die Funktionen zur Bedienung der Scheibenwischer betrachtet. Bereits während der Studie fiel auf, dass die Probanden bei diesen Aufgaben längere Zeit nachdenken mussten, um eine für sie geeignete Geste zu finden als bei anderen Aufgaben. Dies wird hier durch die Einschätzungen der Teilnehmer nochmals deutlich.

Zusätzlich zu den in der Studie abgefragten Funktionen konnten die Probanden noch weitere angeben, von denen sie der Meinung waren, dass sie sich zur Bedienung mit Sprache und Gesten eignen würden. So wurden aus dem Bereich der tertiären Automobilfunktionen besonders häufig die Bedienung des Radios und des Navigationsgerätes genannt. Bei den sekundären Automobilfunktionen kam u.a. die Steuerung des Schiebedaches genauso in Betracht wie die Bedienung der Scheinwerfer.

Abschließend gab es für die Teilnehmer der Studie noch die Möglichkeit ihre Meinung zu Vor- und Nachteilen bei dieser Form der Interaktion mit Sprache und Gesten anzuführen. Mehrfach als Vorteile genannt wurde die Erleichterung der Bedienung, da zum Beispiel alle Einstellungen zentral vom Lenkrad aus durchgeführt werden können und nicht erst die zugehörigen Schalter oder Hebel gesucht und betätigt oder sogar Menüs zur Einstellungsmanipulation durchsucht werden müssen. Des Weiteren wird erwartet, dass die Übersichtlichkeit aufgrund einer Reduktion von physischen Schaltern und Hebeln zunimmt. Auch wurde das Konzept insoweit bestätigt, dass Probanden angaben, dass mit Touch-Gesten Einstellungen exakter, einfacher und schneller machbar sind, als mit Sprache. Sprache eigne sich aber wiederum gut für die Auswahl der Funktion oder des Objektes.

Eine Angabe, die sowohl bei den Vor-, als auch bei den Nachteilen auftauchte, betraf die Fahrerablenkung. Als Vorteil wurde es vermerkt, weil von einer geringeren Ablenkung mit diesem Interaktionsdesign ausgegangen wird. Als Nachteil, weil befürchtet wird, dass eine ungenaue oder falsche Erkennung von Sprachkommandos und Gesten Aufmerksamkeit



■ Mittlere Beurteilung bezüglich der Aussage: „Ich halte eine Bedienung über die Kombination Sprachkommando und Touch-Geste für diese Aufgabe für sinnvoll.“

Mit 1 = "stimme gar nicht zu" bis 5 = "stimme voll und ganz zu".

Abbildung 4.6.: Mittlere Beurteilung einzelner Funktionen hinsichtlich der Sinnhaftigkeit einer Bedienung mit Sprache und Gesten

auf sich zieht. Ebenso könnten eventuelle Rückmeldungen und Feedback des Systems für Ablenkung sorgen. Problematisch wird die Erkennungsgenauigkeit allgemein gesehen. So wurde als störende Einflussquelle laute Musik und Akustik angegeben, eine Fehlerkennung während der Fahrt sogar als potentiell gefährlich eingestuft. Eine gute Fehlerbehandlung, vor allem bei der Spracheingabe, sehen die Probanden deswegen als Notwendigkeit an. Negativ wurde auch die zu erbringende Lern- und kognitive Leistung gesehen. Die Kommandos und Gesten müssten zwangsläufig vor der Nutzung erlernt werden und ein Umdenken zur klassischen Bedienung stattfinden. Gelernte Sprachkommandos und Gesten müssen vor Anwendung auch erst wieder aus dem Gedächtnis abgerufen werden. Als entscheidend für die Akzeptanz wurde auch die Wahl der Funktionen, die mit dieser Modalitätenkombination durchgeführt werden können, eingestuft. Zum einem sollte es keine Funktion sein, die sich auch ohne Geste bereits nur durch ein Sprachkommando durchführen lässt und zum anderen nur bei Funktionen zum Einsatz kommen, bei denen dies möglichst intuitiv und

damit effizient durchführbar ist. Ein Proband gab darüber hinaus an, dass es grundsätzlich merkwürdig und ungewohnt wäre, mit einem Computersystem verbal zu kommunizieren.

4.5.2. Auswertung der aufgezeichneten Sprachkommandos

Während der Studie wurden die geäußerten Sprachkommandos direkt vom Tablet-Computer aufgezeichnet. Für jede Aufgabe, bestehend aus Sprach- und Gesteneingabe, wurde jeweils eine separate Audiodatei gespeichert, so dass für einen ersten Schritt zur Auswertung der aufgezeichneten Sprachkommandos die jeweiligen Audiodateien herangezogen werden konnten. In manchen Fällen war die Qualität der Aufzeichnung allerdings nicht ausreichend, um den genauen Wortlaut der Probanden zu verstehen. In diesem Fall wurde die Videoaufzeichnung, die während der gesamten Studie lief, zu Hilfe genommen. So konnten in diesem ersten Schritt sämtliche Sprachkommandos, die für ein bestimmtes Objekt oder eine Funktion genutzt wurden, aufgelistet werden. Über alle Objekte und Funktionen gesehen, konnten so bei 312 Aufgaben insgesamt 305 verwertbare Sprachkommandos extrahiert werden. Es fiel den Probanden demnach in hohem Maße leicht, Objekte und Funktionen mit einem Sprachkommando zu benennen. Das einzige Objekt, das von einigen Teilnehmern ein- oder mehrmals nicht erkannt wurde, war das Head-Up-Display. Dies war siebenmal der Fall. Das Head-Up-Display und der Tempomat waren die Objekte, die generell für die Probanden am schwierigsten zu identifizieren waren. Da während der Studie für ein bestimmtes Objekt Gesten für mindestens zwei verschiedene Manipulationen gesucht wurden (z.B. Kopfstütze höher und Kopfstütze tiefer), wurde damit auch bei jedem Mal ein zugehöriges Sprachkommando gesammelt. So ergab sich in der Auswertung auch, dass ein Proband ein und dasselbe Objekt beim ersten Mal mit einem anderen Sprachkommando auswählte als beim zweiten Mal. Teilweise gaben die Probanden als Grund dafür an, dass ihnen ihr erstes Kommando nicht gefiel oder sie sich einfach nicht mehr an das erstgenannte Kommando erinnern konnten.

Zwar wurde den Teilnehmern durch die Verwendung von Vorher-Nachher-Bildern die Möglichkeit gegeben, direkt die Objektfunktion (z.B. „*Head-Up-Display Helligkeit*“ oder „*Head-Up-Display Helligkeit einstellen*“) mit einem Sprachkommando auszuwählen, doch wurde davon nur in geringem Maße Gebrauch gemacht. Von den 305 verwertbaren Kommandos traf dies nur auf 16,1% zu. 83,9% der verwertbaren Sprachbefehle benannten das Objekt selbst („z.B. *Head-Up-Display*“).

Während der Studie wurde ein breites Spektrum an Sprachkommandos erfasst. Es reicht demnach nicht aus, nur die jeweils für ein Objekt oder eine Funktion am häufigsten genannten Sprachbefehle in ein Befehlsset zu übernehmen. Für eine spätere Akzeptanz durch Benutzer und hinsichtlich Fehlertoleranz ist eine solch große Fülle an Befehlen entscheidend und so wurde mit den aufgezeichneten Sprachkommandos der Studie, deren Anpassung,

Erweiterung und Modifikation, ein möglichst logisches Befehlsset erstellt. Aber auch wenn ein derart breites Befehlsspektrum in der Lage ist, möglichst viele Eingaben als gültig zu akzeptieren, so sollte dem Benutzer doch nur ein kleiner Auszug dieser Befehle (in z.B. einem PKW-Handbuch) als Referenz angeboten werden, um so die kognitive Leistung für das Abrufen der jeweiligen Befehle aus dem Gedächtnis so minimal wie möglich zu halten. Das gesamte Set an gültigen Sprachkommandos findet sich im Anhang ab Seite 107 wieder. Wie in Abschnitt 4.1 auf Seite 38 angegeben, wurden fehlende Sprachkommandos für Objekte, die wie z.B. die Seitenscheibe mehrfach im Fahrzeug vorkommen, aus den gesammelten Sprachkommandos für andere Instanzen dieses Objektes abgeleitet.

Diese Auflistung an Sprachkommandos erhebt keinen Anspruch auf Vollständigkeit, die Anzahl der Teilnehmer der Studie sind nicht repräsentativ. Obwohl bereits viele Wortkombinationen und Permutationen als gültige Sprachkommandos erfasst wurden, ließe sich dies noch weiter ausbauen und letztendlich evaluieren. Ob und welche Art von Wortkombinationen in das Sprachkommando-Set übernommen wurde, hing in diesem Fall letztendlich auch davon ab, wie hoch die Varianz der Sprachbefehle bei den einzelnen Objekten und Funktionen war. Bei hoher Varianz (z.B. bei den Seitenscheiben) wurden entsprechend viele Permutationen ins Auge gefasst, bei geringer (z.B. bei den Luftdüsen) jedoch weniger. Die vielen verschiedenen Sprachkommandos und die hohe Identifikationsrate für die einzelnen Objekte und Funktionen legen nahe, dass die Auswahl von sichtbaren Objekten ein großes Potential für intuitive Sprachkommandos bietet.

4.5.3. Auswertung der aufgezeichneten Touch-Gesten

Die von den Teilnehmern der Studie eingesetzten Touch-Gesten wurden punktweise in einer Datenbank auf dem Tablet-Computer gespeichert. Innerhalb der Datenbank wurde für jede der 26 Aufgaben der Studie eine eigene Relation vorgesehen, die wiederum die Touch-Punkte der Gesten verwahrten. Eine Android-Applikation, die für die Auswertung geschrieben wurde, erlaubte es diese Gesten wieder aus der Datenbank auszulesen und als Bilddatei zu speichern. Damit konnte ein grafischer Überblick pro Objekt bzw. Funktion über die verwendeten Gesten erstellt werden. Dies war für eine Auswertung allerdings noch nicht ausreichend, um die mentalen Modelle der Probanden nachvollziehen zu können und somit das laute Denken auszuwerten, wurden zusätzlich noch die Videoaufzeichnungen der Studie herangezogen. Damit war es möglich einen Gesamteindruck zu erlangen und die Intentionen der einzelnen Probanden während der Gesteneingabe besser nachvollziehen zu können.

Bei insgesamt 312 Aufgaben konnten während der Studie 304 verwertbare Gesten gesammelt werden. Acht Gesten waren nicht verwertbar, da den Teilnehmern entweder keine passende Geste einfiel (1x) oder sie sich imaginäre Buttons auf dem Tablet-Computer vorstellten (2x),

die sie für die Durchführung der Funktion nutzen wollten. Diese Situationen traten nur bei Scheibenwischerfunktionen auf und widersprechen im Fall der imaginären Buttons dem Konzept dieses multimodalen Interaktionsstils. So würde sich die Fahrerablenkung wiederum erhöhen, da ein Blick auf den Tablet-Computer und damit auf das Lenkrad nötig wäre. Fünfmal waren Gesten nicht verwertbar, weil sich in der Auswertung zeigte, dass die Probanden die Richtung der Gesten offenbar verwechselten und dieselbe Geste schon bei der jeweils konträren Einstellung verwendeten. Diese geschah ausschließlich bei den Gesten für das teilweise/vollständige Öffnen/Schließen einer Seitenscheibe (d.h., Öffnen und Schließen wurde mit derselben Geste durchgeführt).

Eine Mehrheit von 78% der verwertbaren Gesten wurde mit nur einem Finger ausgeführt. 13,2% mit zwei Fingern, 6,9% mit drei Fingern, 1,6% mit vier Fingern und 0,3% mit fünf Fingern. Mehr als fünf Finger wurden von keinem Teilnehmer eingesetzt. Nur zweimal wurde für eine Geste sowohl die linke als auch die rechte Hand verwendet. Ansonsten wurde die rechte Hand doppelt so oft eingesetzt wie die linke.

Von den verwertbaren Gesten waren 86,5% einfache Richtungsgesten (nach links/rechts/oben/unten). Die kreativsten Gesten erdachten sich die Probanden bei den Scheibenwischerfunktionen, bei denen es in der Folge auch nur relativ wenig Übereinstimmung unter den Gesten der Probanden gab. Bei 11 der 26 verschiedenen Aufgaben existierten aber auch große Übereinstimmungen, so gab es bei diesen nur Abweichungen von maximal einer unterschiedlichen Geste unter allen Probanden.

Nach Erstellung eines Überblickes über alle aufgezeichneten Gesten, wurden diese pro Aufgabe nach Gleichheit gruppiert und die für eine Aufgabe bzw. Funktion dann jeweils am häufigsten verwendete Geste der Funktion zugewiesen. Diese Vorgehensweise zur Erstellung eines benutzerdefinierten Gestensets findet sich auch in [69].

Tabelle 4.3 listet die Gesten für Funktionen des Fahrersitzes auf. Um den Sitz nach oben/vorne bzw. nach unten/hinten zu verschieben, wurden jeweils die gleichen Gesten verwendet. Für die Probanden war dies vor allem beim Einstellen der Sitzlängsrichtung häufig eine Frage der Perspektive. So wurden die zugehörigen Vorher-Nachher-Bilder so präsentiert, dass nicht wenige Probanden für das nach vorne Verschieben auch eine Richtungsgeste von rechts nach links und eine von links nach rechts für das nach hinten Verschieben verwendeten.

Einige Probanden formulierten den Wunsch, direkt auf dem Tablet-Computer beispielsweise Regler oder Höhenanzeigen zur Verfügung zu haben, so dass sie wie u.a. bei der Sitzheizung gewünscht, mit dem Finger eine Geste bis zu einer bestimmten Reglerstufe ziehen könnten. Manche Probanden versuchten des Weiteren für verschiedene Funktionen jeweils eine verschiedene Anzahl an Fingern zu verwenden. Aufgrund der Vielzahl an Aufgaben war dies aber nur bedingt erfolgreich. Ein Proband versuchte in elf Fällen zuerst mit einer Geste zusätzlich das Objekt zu malen und dann erst mit einer weiteren Geste die Funktion durchzuführen.











Funktion	Geste	Häufigkeit	Funktion	Geste	Häufigkeit
Sitz nach vorne verschieben		8 von 12	Sitz nach hinten verschieben		7 von 12
Sitz höher stellen		12 von 12	Sitz tiefer stellen		11 von 12
Sitzlehne nach vorne neigen		8 von 12	Sitzlehne nach hinten neigen		8 von 12
Kopfstütze höher stellen		11 von 12	Kopfstütze tiefer stellen		12 von 12
Sitzheizung Temperatur erhöhen		12 von 12	Sitzheizung Temperatur reduzieren		12 von 12

Tabelle 4.3.: Gesten für die Funktionen des Fahrersitzes

Wie aus der Tabelle ersichtlich, konnten für alle Funktionen des Fahrersitzes ausschließlich Richtungsgesten von unten nach oben und von oben nach unten in das endgültige Gestenset aufgenommen werden. Generell dominieren bei allen Funktionen, bei denen Werte erhöht oder reduziert werden müssen, diese beiden Richtungsgesten.

Bei den Gesten für die Manipulation der Außenspiegel aus Tabelle 4.4 ist anzumerken, dass das Drehen der Spiegel nach rechts und nach unten nicht in der Studie abgefragt und diese Gesten jeweils aus den Gesten für die Gegenrichtung abgeleitet wurden.

Nur ein einziger Teilnehmer gab an, die Gesten für den linken Spiegel auf der linken Seite des Tablet-Computers durchgeführt zu haben und analog für den rechten Spiegel auf der rechten Seite. Als Anzeige auf dem Tablet-Computer-Display wurde vereinzelt ein Richtungskreuz gewünscht, entlang dessen man die Achseneinstellungen der Spiegel „nachfahren“ könnte.

Eine Abweichung von der Methode, einer Funktion die jeweils am häufigsten aufgetretene Geste zuzuweisen war bei den Scheibenwischerfunktionen nötig. So war die Vielzahl an verschiedenen Gesten bei keinen anderen Aufgaben so groß wie hier. Dies hatte zur Folge, dass gleich mehrere Gesten am häufigsten eingesetzt wurden. Nun wäre es grundsätzlich

4. Explorative Benutzerstudie





Funktion	Geste	Häufigkeit	Funktion	Geste	Häufigkeit
Außenspiegel nach links drehen		10 von 12	Außenspiegel nach rechts drehen		abgeleitet
Außenspiegel nach oben drehen		11 von 12	Außenspiegel nach unten drehen		abgeleitet

Tabelle 4.4.: Gesten für die Funktionen der Außenspiegel

möglich gewesen, auch mehrere Gesten für eine Funktion zuzulassen, doch bestand die Hauptproblematik darin, dass diese Gesten für die Funktionen Dauerwischen, Intervallwischen, Wischen mit Wischwasser und Abstellen logisch nicht zusammenpassten. Um dieses Ziel weiterverfolgen zu können, wurden die Gesten all dieser Funktionen miteinander verglichen und daraus ein möglichst logisches Teil-Gestenset entwickelt. Wie aus Tabelle 4.5 ersichtlich, besteht die Geste für das Intervallwischen aus einer Richtungsgeste von links nach rechts. Für das Dauerwischen muss dieselbe Geste wieder zum Ausgangspunkt zurückgezogen werden. Die wellenförmige Richtungsgeste von oben nach unten basiert auf einem mentalen Modell eines Probanden, der angab, so die herunterfließenden Tropfen des Wischwassers assoziieren zu können. Die Geste für das Abschalten der Scheibenwischer ist ein einfacher Fingertipp auf dem Tablet-Computer-Display.


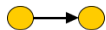


Funktion	Geste	Häufigkeit	Funktion	Geste	Häufigkeit
Scheibenwischer auf Dauerwischen		2 von 12	Scheibenwischer auf Intervallwischen		3 von 12
Scheibenwischen mit Wischwasser		1 von 12	Scheibenwischer abstellen		2 von 12

Tabelle 4.5.: Gesten für die Funktionen der Scheibenwischer






Funktion	Geste	Bemerkung	Häufigkeit	Funktion	Geste	Bemerkung	Häufigkeit
Seitenscheibe teilweise öffnen			8 von 12	Seitenscheibe teilweise schließen			5 von 12
Seitenscheibe vollständig öffnen		ganz nach unten gezogen	7 von 12	Seitenscheibe vollständig schließen		ganz nach oben gezogen	10 von 12
Seitenscheiben-automatik unterbrechen			abgeleitet				

Tabelle 4.6.: Gesten für die Funktionen der Seitenscheiben

Auch für die Ausführung von Funktionen der Seitenscheiben waren die Richtungsgesten von oben nach unten und von unten nach oben wieder am häufigsten vertreten (siehe Tabelle 4.6). Zwei Probanden gaben an, sich eine Anzeige bezüglich der Fensterhöhe auf dem Tablet-Computer-Display zu wünschen, auf der sie dann die Geste zum gewünschten Höhenwert ziehen könnten. Hohe Übereinstimmung bestand bei den Probanden auch darin, dass für das vollständige Öffnen oder Schließen einer Scheibe, die Geste bis zum unteren oder oberen Rand des Tablet-Computer-Displays gezogen werden muss.

Zur Vervollständigung dieses Teil-Gestensets wurde eine neue Funktion zum Unterbrechen des automatischen vollständigen Öffnen/Schließens eingeführt. Diese Funktion war nicht Teil der Studie. Die Geste hierzu wurde zur Einhaltung der Konsistenz von der Geste zum Abstellen der Scheibenwischer abgeleitet und ist demzufolge ein einfacher Fingertipp.

Tabelle 4.7 fasst abschließend die Gesten für die Manipulation des Tempomaten, der Luftdüse und des Head-Up-Displays zusammen. Dies sind wiederum ausschließlich Richtungsgesten von oben nach unten oder von unten nach oben gezogen. Ein Proband ging hier davon aus, dass es eine direkte Anzeige auf dem Tablet-Computer-Display gäbe, auf der eine Geste bis zum gewünschten Tempomat-Geschwindigkeitswert hin gezogen werden könnte. Derselbe Proband wünschte sich dies auch für die Intensität der Luftdüse und für die Helligkeit des Head-Up-Displays.

Es lässt sich festhalten, dass die Studie eine große Übereinstimmung an Gesten unter den Probanden zeigte und sie dabei ähnliche mentale Modelle verfolgten. Sie hatten keine größeren Schwierigkeiten damit, in kurzer Zeit Touch-Gesten zu ersinnen und wählten dabei meist ähnliche und einfache Gesten. Die größten Unterschiede zwischen den Gesten der Probanden traten dabei für die Funktionen der Scheibenwischer auf, hier waren die mentalen

4. Explorative Benutzerstudie







Funktion	Geste	Häufigkeit	Funktion	Geste	Häufigkeit
Tempomat Geschwindigkeit erhöhen		11 von 12	Tempomat Geschwindigkeit reduzieren		11 von 12
Luftdüse Intensität erhöhen		11 von 12	Luftdüse Intensität reduzieren		10 von 12
Head-Up-Display erhellen		10 von 12	Head-Up-Display verdunkeln		11 von 12

Tabelle 4.7.: Gesten für die Funktionen des Tempomaten, der Luftdüsen und des Head-Up-Displays

Modelle nur weniger Probanden jeweils ähnlich und selbst die einzelnen Probanden für sich genommen, hatten Schwierigkeiten ein vollständiges, d.h. konsistentes mentales Modell für Gesten für alle vier Scheibenwischerfunktionen zu bilden.

5. Multimodaler Prototyp

Das in Kapitelabschnitt 3.2 auf Seite 35 vorgestellte Konzept sowie die Ergebnisse der Benutzerstudie aus Kapitel 4 bildeten die Basis für den Entwurf und die Realisierung eines Prototypen für die multimodale Interaktion mit Sprache und Gesten. In diesem Kapitel wird der Aufbau dieses Prototyps erläutert und beschrieben, wie die beiden Eingabemodalitäten verarbeitet werden, welche Ausgaben dem Benutzer präsentiert werden und wie die einzelnen Komponenten zusammenspielen und miteinander verbunden sind.

5.1. Konzipierung des Prototypen

Die prototypische Umsetzung sollte nicht nur Benutzereingaben in Form von Sprache und Gesten entgegennehmen und verarbeiten, sondern dem Benutzer auch in adäquater Form Feedback geben können. In einer Laborumgebung sollten dem Benutzer, welcher die Aufgaben des Fahrers übernimmt, möglichst viele Aspekte einer realen Autofahrt vermittelt werden. Dazu zählen die primären Fahraufgaben, die durch einen Fahrsimulator abverlangt werden sollten sowie die sekundären und tertiären Aufgaben, für die innerhalb der Benutzerstudie aus Kapitel 4 Erkenntnisse in Form von Sprachbefehlen und Gesten gesammelt werden konnten.

Der Benutzer sollte in die Lage versetzt werden, die gesamte Aktionskette durchführen zu können, d.h. mittels eines Sprachkommandos ein Objekt bzw. eine Funktion auswählen, diese mit Gesten zu manipulieren und ein direktes Feedback über diese Einstellungen zu erhalten. Ein weiteres Ziel war es, dem Benutzer sowohl Hilfe bei Fehleingaben anzubieten, andererseits aber auch Fehler präventiv zu vermeiden, indem das Befehlsspektrum möglichst breit gefächert wurde. Kann sich ein Benutzer an ein vollständig qualifiziertes Sprachkommando für ein Objekt nicht erinnern, so kann er sich durch die alleinige Nennung eines Nomens (z.B. „*Fenster*“) eine Übersicht über die zu dem Nomen bzw. Objekt gehörigen Sprachkommandos anzeigen lassen (siehe Übersicht über sämtliche Sprachkommandos im Anhang ab Seite 107). Dies eignet sich nicht nur zur Hilfestellung, sondern kann auch dazu beitragen, eventuelle Fehleingaben des Benutzers abzufangen.

Um vor allem die Art der Rückmeldungen annähernd realistisch zu gestalten, sollten diese rund um den Fahrer stattfinden, d.h. nicht nur auf einem Bildschirm vor dem Benutzer,

sondern auch links und rechts von ihm sowie hinter ihm. Die Objekte und Funktionen aus Tabelle 4.1 auf Seite 39 sollten, soweit möglich, in passender Form präsentiert werden und manipulierbar sein, so dass dem Sichtbarkeitsprinzip wie in Kapitelabschnitt 3.2 auf Seite 35 geschildert, Rechnung getragen werden kann.

Da kein Sitz zur Verfügung stand, mit dem es möglich gewesen wäre, eine reale Änderung der Sitzhöhe oder Sitzlängsrichtung durchzuführen, wurden diese beiden Funktionen des Fahrersitzes nur insofern berücksichtigt, dass sich lediglich die Bilder der Seitenscheiben und -spiegel horizontal und vertikal bewegen (siehe Abschnitt 5.5.2 auf Seite 69) und dadurch ein Verschieben des Sitzes simuliert werden sollte.

Keine Berücksichtigung in der Konzipierung des Prototypen fanden hingegen die Funktionen zur Einstellung der Sitzlehnenneigung und der Kopfstützhöhe. Beide Funktionen ließen sich ohne einen wirklich funktionsfähigen Sitz nicht passend umsetzen.

Die Frontscheibenwischer wurden ebenfalls nicht berücksichtigt, da für die Umsetzung auf einem vorderen Bildschirm ein Fahrsimulator genutzt werden sollte, der diese Funktion nicht zur Verfügung stellte. Die aus Tabelle 4.1 auf Seite 39 ersichtlichen Funktionen für den Frontscheibenwischer wurden gänzlich auf den hinteren Scheibenwischer übertragen.

Die beiden hinteren Seitenscheiben wurden ebenfalls außer Acht gelassen und nur die beiden vorderen Seitenscheiben mit voller Funktionalität implementiert. Auch wenn in der Benutzerstudie aus Kapitel 4 nur für die vordere rechte Seitenscheibe Sprachkommandos und Gesten gesammelt wurden, so wurde für diesen Prototypen auch auf die daraus abgeleiteten Sprachkommandos und Gesten (siehe Auswertung selbiger ab Seite 50) für die vordere linke Seitenscheibe zurückgegriffen. Gleiches gilt für die Außenspiegel und die Luftdüsen.

Aus technischen Gründen musste die Funktionalität der beiden Außenspiegel getrennt werden. Ein Außenspiegel sollte somit nur vertikale Rotation, der andere nur horizontale Rotation zulassen. Dies lag daran, dass sonst für jeden vertikalen Zustand, jeder horizontale Zustand und umgekehrt in Bildern hätte erfasst werden müssen.

5.2. Eigenschaften des Prototypen

Das grundlegende Zusammenspiel der Komponenten des Prototypen wird in Abschnitt 5.2.1 angesprochen. Abschnitt 5.2.2 stellt Kommunikationsaspekte der prototypischen Umsetzung vor und zeigt, wie die einzelnen Komponenten dadurch miteinander kommunizieren.

5.2.1. Überblick über die Komponenten

Die in der Konzipierung des Prototypen angesprochenen Bildschirme wurden um den Benutzer herum verteilt, sodass dieser zu seiner linken Seite, zu seiner rechten Seite, vor und

hinter sich ein Ausgabegerät in Form eines Computer-Bildschirms vorfindet. Die programmiertechnische Umsetzung der Anzeigen auf diesen Bildschirmen und die verschiedenen Rückmeldungen, die der Benutzer von diesen Bildschirmen bekommt, werden in Abschnitt 5.5 für jede Komponente erläutert. Vervollständigt wird die Interaktion durch Eingabetechniken, die bei dieser prototypischen Umsetzung ein Mikrophon, das aus der Benutzerstudie aus Kapitel 4 bekannte, eigengefertigte Lenkrad und den Android-Tablet-Computer (siehe Abbildung 4.1 auf Seite 40) verwenden. Abschnitt 5.4 erläutert die Implementierung und Funktionsweise dieser Eingabekomponenten.

Durch die im folgenden Abschnitt beschriebene Form der Kommunikation, wird aus Implementierungssicht jedes Ausgabegerät und jedes Eingabegerät als separate Komponente verstanden, die physikalisch getrennt voneinander aktiv werden können und miteinander kommunizieren können. Aufgrund der Verwendung von vier 24"-Bildschirmen mussten zwei Computer eingesetzt werden, an die jeweils zwei der Bildschirme angeschlossen wurden. Abbildung 5.1 skizziert den physikalischen Aufbau und illustriert, dass an den ersten Computer die Bildschirme, die links und hinter dem Benutzer stehen angeschlossen werden, der vordere und rechte Bildschirm an den zweiten Computer. Ein anderer Aufbau ist durch die im Folgenden erläuterte Kommunikation ohne weiteres möglich. Der in Abbildung 5.1 skizzierte Anschluss der Eingabegeräte ist deshalb ebenso wenig zwingend in dieser Form vorzunehmen.

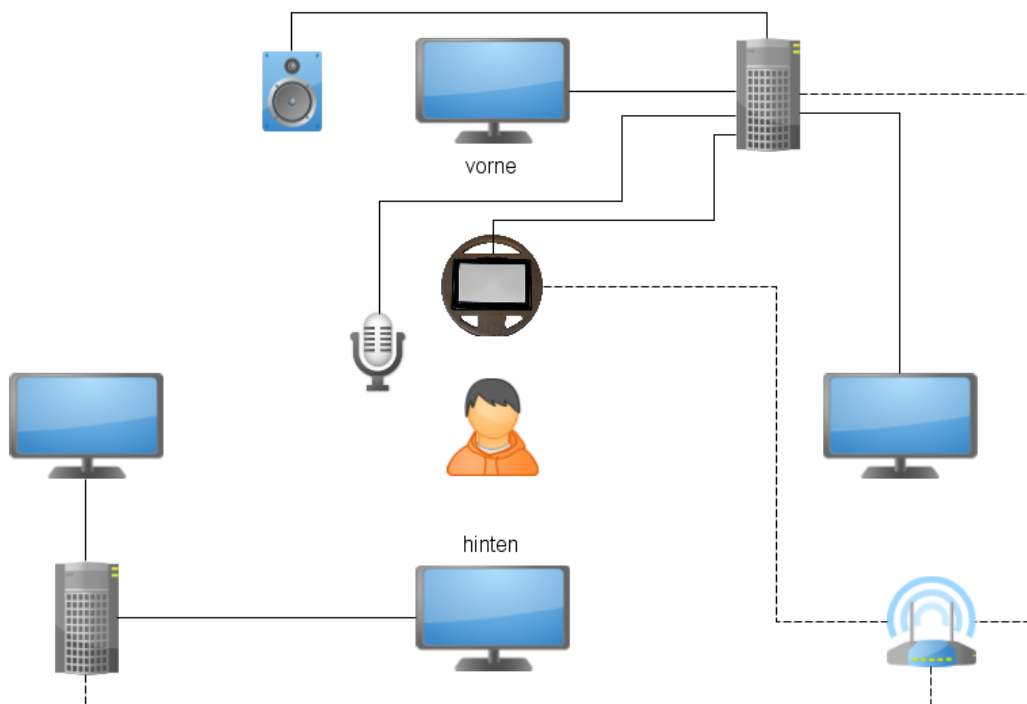


Abbildung 5.1.: Anordnung der Komponenten des Prototypen

5.2.2. Kommunikation der Komponenten

Für die Kommunikation der einzelnen Komponenten mit der Steuerungsapplikation (siehe Abschnitt 5.3), wurde das Embedded Interaction Toolkit (EIToolkit) von Holleis [33] eingesetzt. Das EIToolkit ist eine generische Plattform, die Hard- und Softwarekomponenten über ein einheitliches Kommunikationsprotokoll vernetzen soll. Dadurch ist es z.B. im Sinne dieses Prototypen möglich, die verwendete Form der Sprach- oder Gestenerkennung ohne Änderung der anderen Komponenten auszutauschen, da das Format der versendeten Nachrichten sich selbst nicht ändert.

Das Konzept des EIToolkit sieht sogenannte *Stubs* vor, deren Rolle der eines Proxies oder eines Treibers, der mit einem bestimmten Gerät kommuniziert, ähnelt. *Applikationen* können die von den *Stubs* erhaltenen Daten zur Implementierung der Semantik verwenden und über die *Stubs* wiederum selbst Daten an die Geräte versenden. Die grundlegende Architektur des EIToolkit ist in Abbildung 5.2 dargestellt. Zur Kommunikation werden User Datagram Protocol (UDP)-Pakete verwendet, die von den Komponenten über das Netz gebroadcastet und empfangen werden können. Man kann entweder eine konkrete Komponente als Empfänger des Paketes angeben oder alle Komponenten per Wildcard (*) adressieren. Diese Architektur erlaubt einen einfachen Austausch der separierten Komponenten und eine größtmögliche Unabhängigkeit der Applikationen von der Implementierung der Gerätesteuerung. Das EIToolkit kann mit verschiedenen Programmiersprachen wie Java, C++ oder C# verwendet werden. Für die Implementierung dieses Prototypen wurde auf die Programmiersprache C# zurückgegriffen. Diese wird nicht nur vom EIToolkit unterstützt, sondern bietet auch große Vorteile bezüglich der Spracherkennung (siehe Abschnitt 5.4.1 auf Seite 63).

Das Kommunikationsprotokoll sieht folgenden Aufbau der UDP-Pakete vor:

```
<senderId>:<msgType>:<message>:<receiverId>
```

Wobei <senderId> die ID des Absenders ist, <msgType> Auskunft über den Typ der Nachricht gibt, <message> die Nachricht selbst ist und <receiverId> den Empfänger kennzeichnet. Codefragment 5.1 zeigt die notwendigen Anweisungen in der Programmiersprache C# für das Anlegen und Versenden eines solchen Paketes.

Codefragment 5.1 Codebeispiel in C# zum Anlegen und Versenden eines UDP-Paketes

```
Packet pkt = new Packet(); // Neues Paket anlegen
pkt.SenderId = "SpeechRecognitionStub";
pkt.MsgType = "Sprachkommando";
pkt.Message = "Fahrrersitz";
pkt.ReceiverId = "*"; // Paket ist fuer alle Komponenten bestimmt
EIProperties.PacketSender.sendPacket(pkt); // Paket abschicken
```

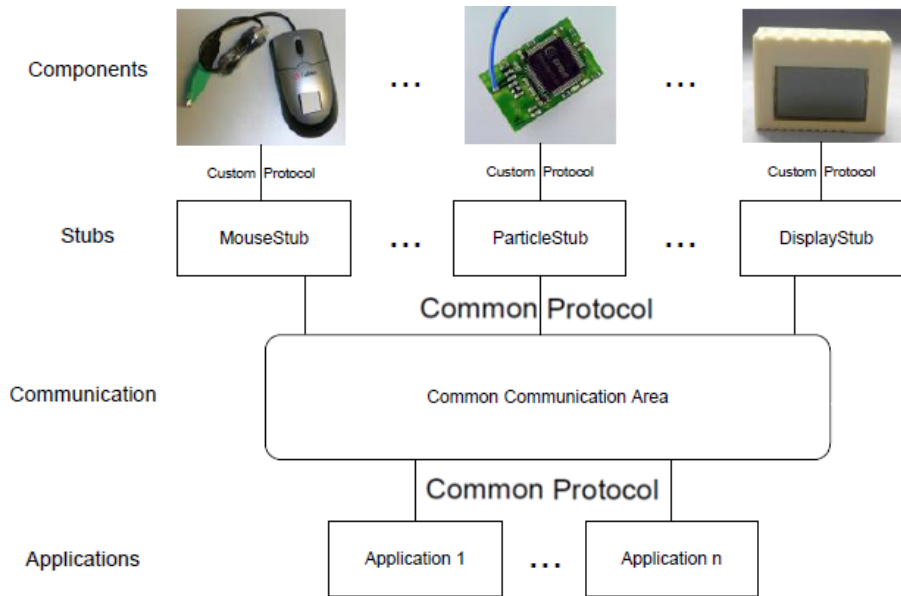


Abbildung 5.2.: Architektur des EIToolkit (Quelle: [33])

5.3. Steuerungsapplikation

Zur Verbindung der einzelnen Komponenten wurde eine Steuerungsapplikation entworfen und implementiert, die im Kontext des EIToolkit mit den verschiedenen Stubs über die in Abschnitt 5.2.2 beschriebenen UDP-Nachrichten kommuniziert. Ihre Aufgabe ist es, die Eingabedaten der Stubs für das Lenkrad, Sprach- und Gestenerkennung zu verarbeiten, als auch die Bildschirm-Stubs zu veranlassen bestimmte Ausgaben darzustellen.

Jeder Interaktionsvorgang, wie er in Abbildung 3.1 auf Seite 35 dargestellt ist, beginnt mit einem eingehenden Sprachkommando. Jeder Stub und jede Applikation, die beim EIToolkit als Empfänger von UDP-Nachrichten registriert wurde, muss eine bestimmte Funktion implementieren, um diese Nachrichten entgegenzunehmen und auswerten zu können. Codefragment 5.2 zeigt den Aufbau dieser Funktion sowie einige Nachrichtentypen, die von der Steuerungsapplikation ausgewertet werden. Dies betrifft insbesondere auch die eingegebenen Sprachkommandos, die vom zugehörigen Stub an die Steuerungsapplikation versendet werden. Nachdem ein Sprachkommando empfangen wurde, wird dies weiter untersucht und verarbeitet.

Für den Beginn eines Interaktionsvorgangs und damit der Spracheingabe, ist es allerdings notwendig diese erst zu aktivieren. Dies geschieht, indem der Benutzer entweder das initiiierende Sprachkommando „Kommando“ äußert oder einen Pull-To-Talk-Mechanismus einsetzt, der bei diesem Prototypen das Ziehen eines Hebels am Lenkrad erfordert. Wird eine dieser beiden Möglichkeiten eingesetzt, akzeptiert das System weitere Spracheingaben und ist in

Codefragment 5.2 Entgegennahme und Auswertung empfangener UDP-Nachrichten

```
void PacketObserver.packetReceived(PacketEvent pktEvent)
{
    Packet pkt = pktEvent.Packet;

    string receiverId = pkt.ReceiverId;
    if (receiverId == InteractionMasterApplication
        || receiverId == EIProperties.PacketFormatter.WildcardAll)
    {
        // Je nach Nachrichtentyp werden die Nachrichten dann weiter an die Handler verteilt
        switch (pkt.MsgType)
        {
            // PTT ausgeloeset, Spracheingabe freischalten
            case "AktiviereSpracheingabe": AktiviereSpracheingabe();
            break;

            // Behandelt ein eingehendes Sprachkommando
            case "Sprachkommando":      UntersucheSprachkommando(pkt.Message.ToString());
            break;

            // Behandelt ein eingehendes Touch-Event
            case "UpdateTouchPointer":   VerarbeiteTouchEvent(pkt.Message.ToString());
            break;

            .
            .
            .
        }
    }
}
```

der Lage die als Ergebnisse der Benutzerstudie aus Kapitel 4 erstellten Sprachkommandos (im Anhang ab Seite 107) zu verarbeiten. Ist die Aktivierung erfolgt, wird der Benutzer darüber durch ein akustisches Signal informiert und erhält so ein direktes Feedback. Das System erwartet nun innerhalb eines kurzen Intervalls ein Sprachkommando. Geht ein solches Sprachkommando ein, so wird dies gemäß dem Konzept aus Kapitelabschnitt 3.2 auf Seite 35 daraufhin untersucht, ob es ein vollständig qualifizierter Sprachbefehl ist. Ist dies nicht der Fall, so wird dem Benutzer dies ebenfalls durch ein akustisches Signal und eine Computerstimme gemeldet und eine knappe visuelle Darstellung der in Frage kommenden Sprachkommandos aufgelistet. So erhält der Benutzer eine direkte Hilfestellung und kann daraufhin seine Fehl- oder unvollständige Eingabe innerhalb eines Zeitfensters mit einem neuen Sprachbefehl korrigieren. Im Falle eines vollständig qualifizierten und gültigen Sprachkommandos, wird dies dem Benutzer wiederum durch ein akustisches Signal

gemeldet und das gültige Kommando von einer Computerstimme wiederholt. Danach wird die Eingabemodalität von Sprache auf Gestik geändert und innerhalb eines bestimmten Intervalls wird nun eine Gesteneingabe erwartet. Jeder einzelne Punkt, der während der Touch-Geste auf dem Tablet-Computer berührt wird, wird in einer UDP-Nachricht von der Steuerungsapplikation entgegengenommen. Je nach aktuellem Interaktionsobjekt werden die entsprechenden Ausgabe-Stubs über diese Gestendurchführung dann ebenfalls per UDP-Nachricht von der Steuerungsapplikation informiert. Diese Stubs können dann ihrerseits die Anzeige auf den zugehörigen Geräten entsprechend setzen.

Im Übrigen wird das System jeweils in den Ausgangszustand zurückgesetzt und ein akustisches Hinweissignal ausgegeben, falls entweder ein Sprachkommando oder eine Gesteneingabe nicht innerhalb der angesprochenen Zeitfenster einging. Nach Abschluss einer Geste geht das System ebenfalls wieder in den Startzustand zurück und es kann ein neuer Interaktionsvorgang begonnen werden.

Die angesprochene akustische Wiederholung eines gültigen Sprachbefehls und der Fehlerhinweis per Computerstimme wurden durch eine Text-To-Speech-Integration umgesetzt. Dazu wurde die Sprachsynthese-API [49] von Windows 7 und eine deutsche Stimme eingesetzt. Die deutsche Stimme musste auf den Webseiten von Microsoft [47] nachgeladen werden, da eine solche bei dieser Betriebssystemversion nicht mitgeliefert wurde. Die Sprachsynthese-API ließ sich auf sehr einfache Weise verwenden, Codefragment 5.3 zeigt eine solche Implementierung.

Codefragment 5.3 Deklaration eines Sprachsynthese-Objektes und Übergabe einer Zeichenkette zur Sprachausgabe

```
SpeechSynthesizer synth = new SpeechSynthesizer();  
synth.SpeakAsync("Keine Geste eingegeben!");
```

5.4. Eingabekomponenten

Drei wesentliche Eingabekomponenten sind Bestandteil des Prototypen. Diese ermöglichen Sprach- und Gesteneingabe sowie eine Aktivierung der Spracherkennung mit Schalthebeln des Lenkrads. Diese Komponenten, die zugehörige Hardware und Implementierungsaspekte werden in den Abschnitten 5.4.1 bis 5.4.3 veranschaulicht.

5.4.1. Spracherkennung

Für die Spracheingabe wird generell ein Mikrofon verwendet. Dies kann z.B. ein Tischmikrofon sein, Versuche zeigten allerdings, dass ein Headset sich besser eignet, da die

Erkennungsrate der Sprachbefehle damit höher ist. Für die Evaluierung aus Kapitel 6 wurde deshalb auch auf ein Headset zurückgegriffen.

Auf Implementierungsseite wurde für die Spracherkennung ein Stub erstellt, der in erster Linie dafür verantwortlich ist, eine Grammatik mit allen akzeptierten Sprachkommandos zu erzeugen und in der Folge diese Sprachkommandos zu erkennen und weiter zu verarbeiten. Dazu wurde die Spracherkennungs-API [48] von Windows 7 eingesetzt, die es ohne größeren Aufwand erlaubt, Grammatiken so zu erzeugen, dass auch Permutationen von Wörtern und Wortfolgen erkannt werden können. Codefragment 5.4 zeigt beispielhaft, wie eine Grammatik angelegt und ein Event-Handler zur weiteren Verarbeitung von erkannten Sprachkommandos angehängt werden kann.

Codefragment 5.4 Deklaration eines Spracherkennungs-Objektes und Definition einer Grammatik

```
SpeechRecognitionEngine sre = new SpeechRecognitionEngine();

// Grammatik definieren
Choices initialesKommando = new Choices(

    "Kommando" // Aktivierungsbefehl fuer Freischaltung der weiteren Spracheingabe
);

// Grammatikobjekt anlegen
Grammar gr = new Grammar(new GrammarBuilder(initialesKommando));

//Event-Handler anhaengen
gr.SpeechRecognized += new EventHandler<SpeechRecognizedEventArgs>(Grammar_SpeechRecognized);

// Grammatik in die SpeechRecognitionEngine laden
sre.LoadGrammar(gr);
```

Es muss betont werden, dass auf diese Art nur die in der Grammatik definierten Sprachkommandos erkannt werden können. Falsche Sprachkommandos erkennt das System als solche nicht. Aus diesem Grunde wurde der in Abschnitt 5.3 auf Seite 62 beschriebende Timeout-Mechanismus eingeführt, sodass nach Aktivierung der Spracheingabe z.B. durch Pull-To-Talk, ein Sprachkommando innerhalb eines bestimmten Intervalls eingegeben werden muss. Ein nicht-definiertes Sprachkommando wird damit so behandelt, als wäre kein Sprachkommando eingegeben worden. Wird andererseits ein in der Grammatik definiertes Sprachkommando erkannt, dann versendet der zugehörige Event-Handler eine UDP-Nachricht an die Steuerungsapplikation, die ihrerseits eine weitere Verarbeitung durchführt und beispielsweise damit das Interaktionsobjekt festlegt.

5.4.2. Gestenerkennung

Das physische Eingabegerät für die Gesten ist der Android-Tablet-Computer, der auch schon bei der Benutzerstudie aus Kapitel 4 zum Einsatz kam und in das selbstgebaute Lenkrad eingebettet wird. Für die Aufnahme der Gesten auf diesem Tablet-Computer wurde eine bereits bestehende Android-Applikation namens TUIOdroid [26] eingesetzt. TUIOdroid ist eine multitouch-fähige Open Source Anwendung, die auf dem TUIO-Protokoll [37] basiert. TUIO ist ein offenes Framework, das Protokoll und API für berührbare Multitouch-Oberflächen definiert und auf dem Open Sound Control-Protokoll [71] basiert. Das TUIO-Protokoll wurde erstmals als Bestandteil des reactIVision-Projektes [7] implementiert und erlaubt den Versand einer abstrakten Beschreibung von interaktiven Oberflächen, darunter u.a. von Touch-Events. Dazu codiert das TUIO-Protokoll Daten einer Tracker-Applikation und sendet sie als TUIO/UDP-Nachricht an jede Client-Anwendung, die in der Lage ist, diese Daten zu dekodieren. Das TUIO-Protokoll findet sich inzwischen in vielen Projekten wieder [62].

Für die Verwendung in dieser Arbeit wurde die originale TUIOdroid-Applikation leicht in der Hinsicht verändert, dass der Hintergrund gänzlich schwarz gesetzt wurde und auch sonst (mit Ausnahme der Systemleiste) keinerlei Ausgaben und Anzeigen auf dem Tablet-Computer zu sehen sind. Die Applikation versendet bei einem Touch-Event auf dem Tablet-Computer eine TUIO/UDP-Nachricht (keine Anbindung an das EIToolkit) für jeden Berührungspunkt mit dessen Koordinaten (die horizontalen und vertikalen Koordinatenachsen werden von der Applikation dabei jeweils auf das Intervall $[0, 1]$ normiert). Ein für die Gestenverarbeitung implementierter Stub empfängt in der Folge diese Nachricht. Dazu instantiiert dieser Stub ein TUIO-Client-Objekt und kann dadurch für jeden einzelnen Finger auf dem Tablet-Computer nachvollziehen, von welcher Koordinate zu welcher Koordinate er gezogen wird, ob ein Finger von der Oberfläche genommen oder ein neuer Finger auf die Oberfläche gesetzt wird. Über eine weitere UDP-Nachricht im Format des EIToolkit-Kommunikationsprotokolls, wird nun die Steuerungsapplikation über ein solches Event in Kenntnis gesetzt. Je nach aktuellem Interaktionsobjekt wird daraufhin ein für ein Ausgabegerät zuständiger Stub aufgefordert, eine bestimmte Manipulation des Objektes (für den Benutzer sichtbar) durchzuführen.

Die Oberfläche der TUIOdroid-Applikation schaltet sich je nach Rotationswinkel automatisch in Hoch- oder Querformat, so lassen sich auch bei Lenkbewegungen Gesten aus derselben Perspektive ausführen. Eine Geste endet, wenn der auslösende Finger von der Oberfläche genommen wird. Eine denkbare Erweiterung wäre, dies ebenfalls mit einem Timeout-Mechanismus zu verbinden, sodass der Benutzer ohne weitere Eingabe desselben Sprachkommandos den Finger kurz von der Oberfläche nehmen und im Anschluss wieder aufsetzen kann, um eine Folgemanipulation durchzuführen.

5.4.3. Pull-To-Talk

Auf der Rückseite des Lenkrades befinden sich zur linken und rechten Seite zwei Hebel. Diese Hebel wurden für einen Pull-To-Talk-Mechanismus verwendet, sodass ein Heranziehen einer dieser Hebel die Spracheingabe aktiviert. Ein dafür implementierter Stub baut über die Microsoft DirectInput-Schnittstelle eine Verbindung mit dem Lenkrad auf und überwacht mittels Polling Zustandsänderungen der Hebel. Ändert sich deren Zustand, d.h. ein Hebel wird herangezogen, dann wird die Steuerungsapplikation mit einer UDP-Nachricht kontaktiert, welche daraufhin die Spracheingabe aktiviert. Abbildung 5.3 zeigt eine Rückansicht des Lenkrades mit den markierten PTT-Hebeln.



Abbildung 5.3.: Rückseite des Lenkrads mit markierten Pull-To-Talk-Hebeln

5.5. Ausgabekomponenten

In den nachfolgenden drei Abschnitten werden die auf den vier Bildschirmen für den Benutzer sichtbaren Anzeigen bzw. Ausgaben beschrieben. Des Weiteren werden relevante Eigenschaften der Implementierung hervorgehoben.

5.5.1. Vorderer Bildschirm

Für die Ausgabe auf dem vor dem Benutzer aufgebauten Bildschirm wurde der bereits in der Benutzerstudie aus Kapitel 4 verwendete Fahrsimulator CARS [39] (siehe auch Kapitelabschnitt 2.1.2 auf Seite 21) abermals eingesetzt. Der Fahrsimulator füllt dabei ca. zwei Drittel des Platzes auf dem Bildschirm aus, im unteren Drittel bekommt der Benutzer noch ein virtuelles Armaturenbrett zu sehen. Auf diesem virtuellen Armaturenbrett sind von links

nach rechts Anzeigen für die Objekte linke Luftdüse, Sitzheizung, Tempomat und rechte Luftdüse angeordnet. Außerdem gibt es zwischen der Anzeige der linken Luftdüse und der Sitzheizung noch einen Bereich, in dem dem Benutzer kurze textuelle Ausgaben angezeigt werden. Abbildung 5.4 zeigt eine Ansicht des vorderen Bildschirms.



Abbildung 5.4.: Ansicht des vorderen Bildschirms

Wie in Abschnitt 5.3 auf Seite 62 angeführt, kann die Steuerungsapplikation in diesem Bereich textuelle Feedback-Informationen an den Benutzer ausgeben lassen, die beispielsweise unvollständig qualifizierte Sprachkommandos betreffen oder dem Benutzer einen Timeout anzeigen. Im Falle von unvollständig qualifizierten Befehlen bekommt der Nutzer die aufgrund seines genannten Sprachkommandos in Betracht kommenden Optionen aufgelistet, aus denen er ein vollständig qualifiziertes Kommando wählen und damit die Spracheingabe wiederholen kann. Ist das eingegebene Sprachkommando gültig, wird in diesem Ausgabebereich das gewählte Interaktionsobjekt textuell angezeigt, so dass dem Benutzer der aktuelle Zustand des Systems jederzeit ersichtlich ist.

Die Kontrolle über die Anzeige auf dem vorderen Bildschirm obliegt einem dafür implementierten Stub. Dieser empfängt von der Steuerungsapplikation Anweisungen, für welches Objekt sich dessen Zustand in welcher Form (bedingt durch die Eingabe einer Geste) ändern soll. Bezüglich der linken und rechten Luftdüse sind die Auswirkungen einer Geste dabei durch einen horizontalen Balken unter den Bildern der Luftdüsen sichtbar. Diese Balken füllen sich nach rechts, wenn die Intensität der Luftdüse erhöht wird und nehmen nach links ab, wenn die Intensität reduziert wird. Da in der Benutzerstudie aus Kapitel 4 ein Sprachkommando für das Objekt Luftdüse und die Funktion Intensität gesucht wurden, sich dies in den Ergebnissen allerdings nicht widerspiegelte (die Funktionsbezeichnung

„Intensität“ fehlte bei allen genannten Kommandos), musste dies für ein verwendbares Set an Sprachkommandos noch entsprechend erweitert werden (siehe Anhang auf Seite 110). Da für die Luftdüsen grundsätzlich noch andere Funktionen außer der Intensität denkbar sind, wurde exemplarisch die (nicht manipulierbare) Funktion „Luftrichtung“ als Bestandteil der gültigen Sprachbefehle aufgenommen.

Die Sitzheizung kann auf sieben verschiedene Stufen eingestellt werden, dazu wird jeweils ein zu einer Stufe gehörendes Bild auf dem Armaturenbrett angezeigt. Der Tempomat ist wiederum mit dem Fahrsimulator CARS gekoppelt. D.h., wird mittels einer Geste die Geschwindigkeit des Tempomaten erhöht oder reduziert, übermittelt der Stub für den vorderen Bildschirm dies per UDP-Nachricht an einen zu CARS gehörenden Stub (das EIToolkit ist ebenfalls in CARS integriert worden). Innerhalb von CARS wird dann die gewünschte Geschwindigkeit des Tempomaten direkt eingestellt. Wird die Tempomat-Geschwindigkeit wieder auf Null gestellt oder das Bremspedal gedrückt wird der Tempomat entsprechend wieder deaktiviert. Über diese Anzeige wird im Übrigen auch die aktuelle Geschwindigkeit des Fahrsimulators angezeigt, denn dieser versendet wiederum seinerseits Nachrichten mit seiner aktuellen Geschwindigkeit.

Das letzte manipulierbare Objekt auf dem vorderen Bildschirm ist das im Fahrsimulator integrierte, textbasierte Head-Up-Display. Dessen Helligkeit kann mit Gesten erhöht oder reduziert werden. Diese Änderungen werden ebenfalls mit UDP-Nachrichten an den an CARS gebundenen Stub geschickt, welcher die Helligkeitsänderungen veranlasst. Anders als bei der Luftdüsenfunktion „Intensität“ war die Funktionsbezeichnung „Helligkeit“ in einigen wenigen Fällen Bestandteil der gesammelten Sprachkommandos. Doch hat das Head-Up-Display ebenfalls noch andere Funktionen, die nicht in der Benutzerstudie verwendet wurden und so ist die Funktionsbezeichnung „Helligkeit“ im erstellten Sprachkommando-Set grundsätzlich Bestandteil des zu diesem Objekt und dieser Funktion gehörenden Sprachbefehls. Exemplarisch wurde für das Head-Up-Display noch die Funktion „Position“ als Bestandteil valider Sprachkommandos aufgenommen. Auch diese Funktion ist mit diesem Prototyp nicht manipulierbar.

5.5.2. Linker und rechter Bildschirm

Für die Ausgabe auf dem linken und rechten Bildschirm war ebenfalls jeweils ein Stub verantwortlich und die Objekte auf beiden Bildschirmen waren jeweils dieselben. So wird auf dem einen das linke vordere Fenster und der linke Außenspiegel, auf dem anderen das rechte vordere Fenster und der rechte Außenspiegel angezeigt (siehe Abbildung 5.5). Für die verschiedenen Zustände dieser Objekte wurden jeweils einzelne Frames hinterlegt, die je nach Zustandsänderung wechseln. So kann der Benutzer mit Gesten die Fensterscheiben nach oben und nach unten fahren, den linken Außenspiegel auf der vertikalen Achse

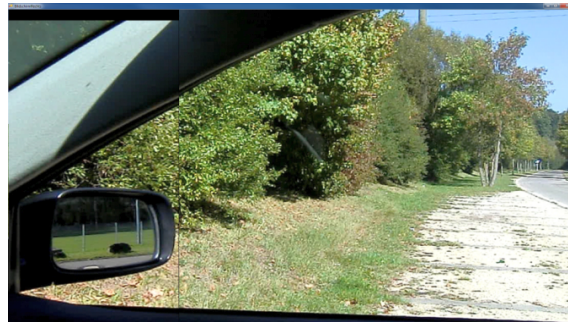
drehen und den rechten Außenspiegel auf der horizontalen. Führt er eine der Gesten für die Seitenfenster an den unteren oder oberen Rand des Tablet-Computers, so setzt ein automatisches vollständiges Öffnen oder Schließen des Fensters ein, das mit einem einfachen Fingertipp auf das Display wieder abgebrochen werden kann. Dies ist der Automatik eines realen elektrischen Fensterhebers nachempfunden.

Der Benutzer erhält ein direktes Feedback auf seine Aktionen, indem er z.B. die Fahrt der Fenster in Echtzeit nachverfolgen kann. Im Sinne der Usability kann eine Aktion leicht rückgängig gemacht werden, indem ein Fenster wieder ein Stück nach unten gefahren wird, nachdem es zuvor ein Stück nach oben gefahren wurde. Dieser Aspekt gilt analog auch für sämtliche manipulierbaren Objekte des vorderen Bildschirms.

Wie in Abschnitt 5.1 auf Seite 58 bereits erwähnt, kann durch die Anzeigen auf dem linken und rechten Bildschirm auch eine Bewegung des Fahrersitzes nachgestellt werden. Dazu werden die Bilder der Fenster und Außenspiegel entweder nach vorne oder hinten verschoben um eine Änderung der Sitzlängsrichtung zu simulieren oder nach oben oder unten für eine Änderung der Sitzhöhe. Das Sprachkommando, das aufgrund der Ergebnisse der Benutzerstudie eigentlich für die Änderung von Sitzhöhe und Sitzlängsrichtung verwendet werden müsste, war „*Fahrersitz*“. Doch wäre es problematisch, eine Verstellung auf beiden Achsen gleichzeitig durchführen zu können und auch die Gesten für beide Einstellungen sind jeweils dieselben (siehe Tabelle 4.3 auf Seite 53). Deshalb wurde das Kommando „*Fahrersitz*“ durch die beiden zusätzlichen Kommandos „*Sitzposition*“ für eine Änderung der Sitzlängsrichtung und „*Sitzhöhe*“ für eine Änderung der Sitzhöhe ersetzt.



(a) Ansicht des linken Bildschirms



(b) Ansicht des rechten Bildschirms

Abbildung 5.5.: Die Ansichten der links und rechts des Benutzers angebrachten Bildschirme

5.5.3. Hinterer Bildschirm

Auf dem hinteren Bildschirm wird ein Heckscheibenwischer in verschiedenen Zuständen angezeigt. Der Stub, der die Ausgabe auf diesem Bildschirm setzt, bekommt von der Steuerungsapplikation mitgeteilt, ob er den Scheibenwischer auf Dauerwischen, Intervallwischen, Wischen mit Wischwasser oder ganz abstellen soll. Die technische Umsetzung wurde mit aufgezeichneten Videos realisiert, die die drei erstgenannten Zustände darstellen können. Für den abgestellten Scheibenwischer wird ein Standbild angezeigt. Abbildung 5.6 zeigt die Darstellung auf dem hinteren Bildschirm.



Abbildung 5.6.: Ansicht des hinteren Bildschirms beim Scheibenwischen mit Wischwasser

Einen Gesamtüberblick über den erstellten Prototypen gibt abschließend die Abbildung 5.7. Sie zeigt die reale Anordnung aller Komponenten. Die Anzeigen der Bildschirme sind jeweils den Abbildungen 5.4 bis 5.6 zu entnehmen.



Abbildung 5.7.: Gesamtaufbau des Prototypen

6. Evaluierung des Prototypen und Interaktionsstils

Dieses Kapitel behandelt die Durchführung einer Evaluierung des erstellten Prototypen aus Kapitel 5 in Verbindung mit dem konzipierten Interaktionsstil aus Kapitelabschnitt 3.2 auf Seite 35.

Dabei sollte die Annahme überprüft werden, dass dieser Interaktionsstil die Fahrerablenkung reduzieren kann oder zumindest nicht erhöht und von den Benutzern auch positiv angenommen wird. Dazu wurde eine Vergleichsevaluierung mit einem traditionellen Interaktionsstil durchgeführt, deren Ergebnisse abschließend erörtert und illustriert werden.

6.1. Vorbereitung der Studie

Unter Einbeziehung der in Kapitel 5 beschriebenen Möglichkeiten der prototypischen Umsetzung bezüglich Objektauswahl und -manipulation, wurde eine Auswahl an Objekten und Funktionen getroffen, die die Teilnehmer der Evaluierung zur Bearbeitung bekamen. Letztendlich umfasste die Auswahl sechs verschiedene Objekte mit einem für die Evaluierung angepassten Funktionsspektrum. Tabelle 6.1 gibt einen Überblick über diese Festlegung. So wurden beide Außenspiegel, der Heckscheibenwischer, die beiden vorderen Seitenfenster und der Tempomat in die Aufgabenliste aufgenommen, die Sitzheizung, die beiden Luftdüsen und das Head-Up-Display hingegen nicht.

6.2. Aufbau der Studie

Die Evaluierung wurde in einer Laborumgebung durchgeführt und es wurde mit einigen Ausnahmen der Aufbau des in Kapitel 5 vorgestellten Prototypen in dieser Studie eingesetzt. Der vordere Bildschirm wurde durch einen 55"-Fernsehbildschirm ersetzt und auf diesem kamen auch nicht das in Kapitelabschnitt 5.5.1 auf Seite 66 beschriebene virtuelle Armaturenbrett und der CARS-Fahrsimulator zum Einsatz. Stattdessen wurde der in Kapitelabschnitt 2.1.2 auf Seite 19 vorgestellte Lane-Change-Test (LCT) auf dem Frontbildschirm

6. Evaluierung des Prototypen und Interaktionsstils

Interaktionsobjekt	Interaktionsfunktion	Ausgangssituation des Interaktionsobjektes
Linker Außenspiegel	Ganz nach oben drehen	Vertikal mittig
Linker Außenspiegel	Ganz nach unten drehen	Vertikal mittig
Rechter Außenspiegel	Ganz nach links drehen	Horizontal mittig
Rechter Außenspiegel	Ganz nach rechts drehen	Horizontal mittig
Heckscheibenwischer	Auf Dauerwischen	Aus
Heckscheibenwischer	Auf Intervallwischen	Aus
Heckscheibenwischer	Mit Wischwasser wischen	Aus
Heckscheibenwischer	Abstellen	Aus
Fahrerfenster	Zur Hälfte öffnen	Vollständig geschlossen
Beifahrerfenster	Zur Hälfte schließen	Vollständig geöffnet
Tempomat	Geschwindigkeit um mind. 15 km/h erhöhen	Bei 30 km/h
Tempomat	Geschwindigkeit um mind. 15 km/h reduzieren	Bei 30 km/h

Tabelle 6.1.: Übersicht der Objekte und durchzuführenden Funktionen der Evaluierung

angezeigt und zur Interaktion bereitgestellt. Da das LCT-Programm keinen Fenster-, sondern nur einen Vollbildmodus zur Verfügung stellte, konnte das virtuelle Armaturenbrett nicht auf dem Frontbildschirm angezeigt werden und musste auf ein Netbook ausgelagert werden. Basierend auf den Objekten und Funktionen aus Tabelle 6.1 wurden die linke und rechte Luftdüse aus diesem Armaturenbrett entfernt. Die Sitzheizung wurde zwar noch angezeigt, hatte aber keine Funktion in dieser Evaluierung. Die Bildschirmanzeigen links, rechts und hinten blieben unverändert und entsprachen jeweils dem in den Kapitelabschnitten 5.5.2 auf Seite 68 und 5.5.3 auf Seite 70 spezifizierten Aufbau.

Sämtliche Audioausgaben des Prototypen wurden anders als in Abbildung 5.1 auf Seite 59 zu sehen, über den hinteren Computer ausgegeben, damit sich die Motorengeräusche des LCT, der auf dem vorderen Computer lief, nicht mit den Audioausgaben des Prototypen überlagern und deren Lautstärke jeweils separat steuerbar ist. Abbildung 6.1 zeigt den Aufbau der beschriebenen Anzeigen mit Ausnahme des hinteren Bildschirms, dieser war allerdings unverändert wie auf Abbildung 5.7 auf Seite 70 einsehbar, positioniert.



Abbildung 6.1.: Physischer Aufbau der Evaluierung

Neben diesen Ausgabegeräten kamen das in Abbildung 4.1 auf Seite 40 veranschaulichte und in Kapitelabschnitt 4.2 auf Seite 38 beschriebene Holzlenkrad zum Einsatz. Ebenso wurde wiederum der Android-Tablet-Computer verwendet, der innerhalb des Prototypen für die Aufnahme der Gesten verantwortlich war. Auch in der Evaluierung hatte der Tablet-Computer diese Aufgabe, es wurde aber neben der in Kapitelabschnitt 5.4.2 auf Seite 65 erläuterten TUIOdroid-Applikation noch eine zweite Android-Applikation für dieses Gerät erstellt, die für den in dieser Vergleichsevaluierung eingesetzten herkömmlichen Interaktionsstil verwendet wurde. Mit dieser Applikation war es möglich, durch das Drücken des jeweiligen Buttons auf dem Tablet-Computer, den Tempomaten oder die vier Scheibenwischerfunktionen direkt zu manipulieren bzw. zu steuern (siehe Abbildung 6.2a).

Als Eingabegerät für die Modalität Sprache wurde ein Headset verwendet, sodass die geäußerten Kommandos durch die Nähe zum Mikrophon möglichst gut erkannt werden konnten. Neben Fußpedalen zur Geschwindigkeitsregulierung innerhalb des Lane-Change-Tests, kamen auch noch zwei Nintendo Wii-Controller [52] zum Einsatz. Diese wurden innerhalb der Vergleichsevaluierung ebenfalls für den nachgestellten herkömmlichen Interaktionsstil genutzt. Als Halterung für die Wii-Controller wurde eine eigengefertigte Holzkonstruktion verwendet (siehe Abbildung 6.2b) in die die Geräte passgenau eingesetzt werden konnten. Positioniert wurden die Wii-Controller neben dem linken Bildschirm so, dass sie mit der linken Hand gut zu erreichen waren. Auf einem der beiden eingesetzten Computer lief eine Java-Anwendung, die von den über Bluetooth angeschlossenen Wii-Controllern Events

6. Evaluierung des Prototypen und Interaktionsstils

entgegennahm und über das EIToolkit (siehe Kapitelabschnitt 5.2.2 auf Seite 60) an den Prototypen sendete. So war es möglich, über zwei Knöpfe eines Wii-Controllers eine Fenserscheibe nach oben und nach unten zu fahren und über das Steuerkreuz auf einem Wii-Controller einen Außenspiegel auf der vertikalen und horizontalen Achse zu drehen. Ein Wii-Controller war für die linke Seitenscheibe und linken Außenspiegel, der andere für die rechte Seitenscheibe und rechten Außenspiegel verantwortlich.



(a) Ansicht der Android-Applikation auf dem Tablet-Computer (b) Wii-Controller und Halterung

Abbildung 6.2.: Nahaufnahmen der in der Evaluierung verwendeten Android-Applikation, Wii-Controller und Halterung zur Nachstellung eines herkömmlichen Interaktionsstils

6.3. Teilnehmer

An der Evaluierung nahmen 16 Personen teil, die wie bereits in der explorativen Benutzerstudie aus Kapitel 4 über E-Mail-Verteiler oder persönlich angeworben wurden. Auch hier war die einzige Voraussetzung für die Teilnahme der Besitz eines Führerscheines. Die Studienvergütung wurde im Vergleich allerdings verdoppelt, so bekam jeder Proband 10 € für ca. 60 Minuten Teilnahmedauer.

Tabelle 6.2 gibt eine Übersicht über allgemeine Informationen zu den Probanden der Evaluierung. Diese Informationen wurden aus den ausgefüllten Eingangsfragebögen entnommen. Ein Muster dieses Eingangsfragebogen findet sich im Anhang auf Seite 117. Dieser entspricht exakt dem Eingangsfragebogen aus der explorativen Benutzerstudie.

Alter	zwischen 21 und 29 Jahre (im Durchschnitt 23,81 Jahre)
Automobilnutzung	11 Teilnehmer gelegentlich, d.h. mehr als zweimal im Monat 5 Teilnehmer täglich
Studiengang	13 Studenten der Fachrichtung Informatik/Softwaretechnik 1 Student der Fachrichtung Umweltschutztechnik 1 Student der Fachrichtung Physik 1 Student der Fachrichtung Luft- und Raumfahrttechnik
Fahrzeugeigner	8 Teilnehmer besaßen ein eigenes Auto
Führerscheinbesitzdauer	zwischen 3 und 10 Jahre (im Durchschnitt 5,75 Jahre)
Geschlecht	16 männliche Teilnehmer keine weiblichen Teilnehmer
Händigkeit	14 Rechtshänder 2 Linkshänder

Tabelle 6.2.: Allgemeine Informationen zu den Teilnehmern der Evaluierung

6.4. Ablauf der Studie

Die einzelnen Schritte, die während der Evaluierung durchgeführt wurden, sind in den folgenden Abschnitten beschrieben. Die tatsächliche Reihenfolge während der Evaluierung stimmt mit der Reihenfolge der Abschnitte überein.

6.4.1. Einführung in die Evaluierung und erster Fragebogen

Zu Beginn wurden den Teilnehmern Grundinformationen über die Evaluierung mitgeteilt. Diese beinhalteten eine kurze Erläuterung des untersuchten Interaktionsstils sowie, dass es sich um eine Vergleichsevaluierung handelt und der in Kapitelabschnitt 3.2 auf Seite 35 entworfene Interaktionsstil mit Sprache und Gesten mit einem Interaktionsstil verglichen werden soll, der die herkömmliche Interaktion im Automobil nachstellt. Detailliertere Informationen wie die beiden Interaktionsstile genau umgesetzt wurden gab es an dieser Stelle noch nicht. Vielmehr noch der allgemeine Hinweis, dass die Studie jederzeit abgebrochen werden kann und alle Daten vertraulich behandelt werden. Dann wurden die Probanden gebeten, die Kenntnisnahme dieser Hinweise in einer Einverständniserklärung zu bestätigen. In dieser Einverständniserklärung konnten die Probanden auch die Freigabe zur Verwendung von Sprach- und Videoaufzeichnungen geben, diese Aufzeichnungen wurden aber nur bei

wenigen Probanden gemacht und waren auch nicht Bestandteil der späteren Auswertung. Bevor die Teilnehmer vor dem Fahrsimulator Platz nehmen sollten, wurde noch der im vorigen Abschnitt 6.3 bereits angesprochene Eingangsfragebogen ausgeteilt und von den Teilnehmern ausgefüllt. Auf diesem fanden sich allgemeine Fragen zur Person wie Alter, Studiengang oder Geschlecht sowie zu bisherigen Erfahrungen mit Spracheingabe/-steuerung und (Multi)-Touch-Displays.

6.4.2. Einweisung und Übungsfahrt mit dem Lane-Change-Test

Danach wurden die Probanden direkt vor den Fahrsimulator gesetzt und sollten sich dort zuallererst richtig positionieren. D.h., den Sitz, den linken und rechten Bildschirm und die Fußpedale so zu stellen, dass alle Anzeigen gut einsehbar und Lenkrad sowie die Wii-Controller mit den Händen gut erreichbar sind.

Als nächstes folgte eine Einweisung in den Lane-Change-Test (LCT). Die Einweisung wurde mit einem Screenshot aus dem LCT bebildert, der wie Abbildung 2.2 auf Seite 20 aufgebaut war. Anhand dieses Screenshots wurde den Probanden erläutert, dass eine Fahrt mit dem Lane-Change-Test stets vor oder in einer Kurve beginnt und spätestens bei Kurvenausgang das Gaspedal konstant durchgedrückt werden muss. Auf dem nun folgenden geraden Streckenabschnitt sollte stets mit Maximalgeschwindigkeit gefahren werden. Bereits an dieser Stelle wurden die Teilnehmer darauf hingewiesen, dass ab dem gelben Startschild einer LCT-Strecke die erste Interaktionsaufgabe vorgelesen wird, diese Aufgaben nacheinander und in zufälliger Reihenfolge vorgegeben werden und erst ausgeführt werden dürfen, wenn die jeweiligen Aufgabenstellungen vollständig verbal wiedergegeben wurden. Die Ausführung der Aufgaben sollte aber möglichst direkt im Anschluss an deren verbale Wiedergabe erfolgen. Allerdings wurden die Teilnehmer der Studie eindringlich darauf aufmerksam gemacht, dass das Hauptaugenmerk darauf liegt, die Schilder an den Straßenrändern (siehe Abbildung 2.2 auf Seite 20) genau zu beachten und die auf diesen Schildern jeweils angezeigten Fahrspurwechsel auf eine der beiden anderen möglichen Spuren vorzunehmen. Dies sollte sofort nach Wahrnehmung der Schilder und der darauf angezeigten Fahrspurwechsel geschehen und nicht erst nachdem ein Schild bereits passiert wurde. Auch sollte nicht einfach scharf eingelenkt oder auf die andere Spur geschlittert, sondern in gewohnter, normaler Art und Weise die Spur gewechselt werden. Abschließend gab es noch den Hinweis, dass die Fahrt auf einer Strecke ca. drei Minuten dauert und in einer Kurve endet. Für weitere, vertiefende Informationen zum Lane-Change-Test sei auf Kapitelabschnitt 2.1.2 auf Seite 19 verwiesen, in der diese Messmethode für Fahrerablenkung genauer erörtert wird.

Nach dieser Einführung in den Lane-Change-Test und die zu beachtenden Verhaltensweisen, bekam jeder Proband Gelegenheit eine Übungsfahrt mit dem LCT durchzuführen. Während dieser Übungsfahrt mussten keine Aufgaben durchgeführt werden, sie diente lediglich der

Eingewöhnung der Teilnehmer an den Fahrsimulator. Die zuvor angesprochenen Vorgaben für die Fahrt und Aufgabendurchführung wurden dabei nochmals wiederholt. Die für diese Übungsfahrten und alle anderen Fahrten verwendeten Strecken, variierten Probandenübergreifend in der Form, dass begonnen mit Strecke 1 des LCT, nacheinander die Strecken 2 bis 10 verwendet wurden und dann wieder von vorne mit Strecke 1 begonnen wurde. Durch diese abwechselnde Streckenwahl wurde Lerneffekten vorgebeugt.

6.4.3. Erste Baseline-Fahrt

Im Anschluss an die Übungsfahrt und nachdem die Probanden erklärten, mit dem Lane-Change-Test nun soweit vertraut zu sein, dass sie die Vorgaben umsetzen könnten, wurde das LCT-Programm neu gestartet und wie angesprochen eine andere Strecke gewählt. Der Neustart des LCT-Programms ist grundsätzlich nach jeder Fahrt notwendig gewesen, da erst mit dem Beenden des Programms eine Analyse-Datei für die gefahrene Strecke angelegt wird. Nur so konnte jede Fahrt jeweils einzeln ausgewertet werden. Die Analyse-Dateien für die Übungsfahrten waren dabei allerdings nicht von Relevanz.

Nun sollten die Teilnehmer der Evaluierung exakt dasselbe machen wie zuvor, nämlich eine Strecke mit dem Fahrsimulator fahren. Diesmal stand jedoch eine sogenannte Baseline-Fahrt an, d.h. eine Fahrt ohne Interaktionsaufgaben, für die die erbrachte Fahrleistung in der späteren Auswertung zu Vergleichszwecken herangezogen werden konnte. Wie zuvor bei der Übungsfahrt, wurde nur auf das Fahren unter Einhaltung der LCT-Regeln Wert gelegt, diesmal hingegen wurde die dadurch erzeugte Analyse-Datei für die Auswertung gesichert.

6.4.4. Fahrten mit unterschiedlichen Interaktionsstilen

Nachdem die Baseline-Fahrt beendet wurde, standen nacheinander zwei Fahrten mit Durchführung von Nebenaufgaben, d.h. Interaktionen an. In dieser Vergleichsevaluierung wurden dazu zwei unterschiedliche Interaktionsstile verwendet. Auf der einen Seite der in Kapitelabschnitt 3.2 auf Seite 35 konzipierte und im Prototyp aus Kapitel 5 integrierte Interaktionsstil mit Sprache und Gesten. Auf der anderen Seite ein Interaktionsstil, der der herkömmlichen Interaktion im Automobil nahe kommen sollte. Da ein reales Fahrzeug nicht zur Verfügung stand, sollte dieser herkömmliche Interaktionsstil eine implizite Objektauswahl und manuelle Durchführung von Funktionen wie von einer realen Automobilumgebung bekannt, nachstellen. Die dafür verwendeten Nintendo Wii-Controller, die Android-Applikation für den Tablet-Computer sowie die damit durchführbaren Einstellungen wurden bereits in Abschnitt 6.2 auf Seite 73 vorgestellt. Dieser Interaktionsstil wurde als Interaktion mit Buttons

und Knöpfen bezeichnet, da Schaltflächen auf Tablet-Computern (bzw. Touchscreens) umgangssprachlich eher mit der Bezeichnung „Buttons“ bezeichnet werden und „Knöpfe“ als gängigere Bezeichnung für die Steuerelemente der Wii-Controller angesehen wurde. Durch diese Unterscheidung sollten die Probanden auch ein Gefühl dafür bekommen, dass sie für diesen Interaktionsstil zum einen die Wii-Controller, zum anderen den Tablet-Computer nutzen konnten, obwohl beide Bezeichnungen eigentlich die gleiche Bedeutung haben.

In welcher Reihenfolge die zwei Konditionen in den beiden nach der Baseline-Fahrt anstehenden Fahrten zum Einsatz kamen, wurde wie bei den Strecken Probanden-übergreifend variiert. D.h., wenn ein Teilnehmer zuerst eine Interaktionsfahrt mit der Kondition Sprache und Gesten und anschließend eine mit der Kondition Buttons und Knöpfe durchführte, dann musste der nächste Teilnehmer die erste Interaktionsfahrt mit der Kondition Buttons und Knöpfe und die nächste mit Sprache und Gesten durchführen, usw. Im Folgenden werden die Interaktionsstile und die dafür notwendig gewordenen Erklärungen für die Probanden erläutert.

Einweisung in die Interaktion mit Sprache und Gesten

Den Teilnehmern wurde das Headset ausgehändigt und die TUIOdroid-Applikation (siehe Kapitelabschnitt 5.4.2 auf Seite 65) gestartet. Dann wurde das Konzept der Interaktion mit Sprache und Gesten, beginnend mit den Möglichkeiten der Aktivierung der Spracherkennung vorgestellt. Dies geschieht wie in Kapitelabschnitt 5.3 auf Seite 61 angeführt durch das initiierte Sprachkommando „Kommando“ oder durch einen Pull-To-Talk-Mechanismus (siehe Kapitelabschnitt 5.4.3 auf Seite 66). Mit welcher Methode die Probanden die Spracherkennung aktivierten, war ihnen freigestellt. Im Anschluss wurden sämtliche zur Verfügung stehenden Sprachkommandos erläutert. Die Sprachkommandos für diese Evaluierung sind identisch mit den in Tabelle 6.1 auf Seite 72 verwendeten Objektbezeichnungen. Ebenso wurden die Probanden mit den Gesten für die in Tabelle 6.1 aufgelisteten Funktionen vertraut gemacht. Dazu sei auf die relevanten Gesten aus Tabelle 4.3 auf Seite 53 bis Tabelle 4.7 auf Seite 56 verwiesen. Die Teilnehmer hatten nun die Gelegenheit, unter Anleitung und dann alleine, komplette Interaktionsvorgänge durchzuführen und Objekte so zu manipulieren. Jedes Sprachkommando und jede Geste musste mindestens einmal geübt werden, bis die Probanden angaben, damit umgehen zu können.

Nach Abschluss dieser Einarbeitung in den Interaktionsstil, wurden die Einstellungen aller Objekte der Evaluierung in den in Tabelle 6.1 beschriebenen, jeweiligen Ausgangszustand zurückversetzt. Damit war sichergestellt, dass die vorbereiteten Interaktionsaufgaben auf den richtigen Zuständen der Objekte ansetzen konnten und nicht z.B. eine Fensterscheibe geöffnet werden sollte, die bereits geöffnet war.

Einweisung in die Interaktion mit Buttons und Knöpfen

Auch für diesen Interaktionsstil wurde den Probanden die Bedienung sämtlicher Funktionen aus Tabelle 6.1 auf Seite 72 vorgestellt. Wie bereits in Abschnitt 6.2 auf Seite 74 angesprochen, konnte eine Fensterscheibe mit zwei Knöpfen auf einem Wii-Controller geöffnet oder geschlossen und ein Außenspiegel mit dem Steuerkreuz horizontal oder vertikal gedreht werden. Wie auf Abbildung 6.2a auf Seite 74 ersichtlich, waren die vier Scheibenwischerfunktionen Dauerwischen, Intervallwischen, Wischen mit Scheibenwischwasser und Scheibenwischer abstellen über Buttons auf dem Tablet-Computer einstellbar. Ebenso konnte über den Tablet-Computer die Geschwindigkeit des Tempomaten erhöht und reduziert werden.

Auch hier mussten die Probanden jede Funktion mindestens einmal mit diesem Interaktionsstil durchführen und sich einüben. Mit diesem Interaktionsstil kam die Mehrzahl der Probanden schnell zurecht und die Einstellungen aller Objekte konnten wieder in ihren Ausgangszustand zurückgesetzt werden.

Nachdem, abhängig von der Reihenfolge, einer der beiden Interaktionsstile den Probanden verständlich gemacht wurde, begann die dritte Fahrt (nach Übungs- und Baseline-Fahrt) mit dem Lane-Change-Test. Diesmal wurden ab dem Startschild der jeweiligen Strecke nacheinander Aufgaben vorgegeben, die die Probanden nebenher mit dem jeweiligen Interaktionsstil durchführen mussten. Diese Aufgaben beruhten auf den in Tabelle 6.1 auf Seite 72 aufgelisteten 12 Objekt-Funktions-Paaren. Diese 12 Aufgaben wurden für jeden Probanden in eine zufällige Reihenfolge gebracht und hatte ein Proband diese 12 Aufgaben vor Streckenende bereits alle abgearbeitet, begann seine individuelle Aufgabenliste nochmals von vorn.

Nach Abschluss dieser Fahrt wurde ein Fragebogen an die Probanden ausgegeben, mit dem sie die Fahrt unter dem eben eingesetzten Interaktionsstil bewerten sollten. Der Fragebogen für die Interaktion mit Sprache und Gesten findet sich im Anhang ab Seite 119, der Fragebogen für die Interaktion mit Buttons und Knöpfen ab Seite 123. In diesen Fragebögen wurden jeweils Fragen der Driving Activity Load Index-Methode (DALI) [54] zur Messung der mentalen Beanspruchung des Fahrers und der System Usability Scale-Methode (SUS) [12] zur Bewertung der System-Usability eingesetzt.

Direkt nach Ausfüllen des ersten Zwischenfragebogens begann die Einweisung in die nächste Kondition, d.h. den in der Reihenfolge nächsten Interaktionsstil. Die Vorgehensweise war dann identisch wie zuvor beschrieben: Zuerst eine vierte Lane-Change-Test-Fahrt mit Interaktionsaufgaben und danach der Fragebogen zur Bewertung des Interaktionsstils. Abbildung 6.3 zeigt einen Probanden bei einer Fahrt während der Durchführung einer Nebenaufgabe mit dem Interaktionsstil Buttons und Knöpfe.



Abbildung 6.3.: Ansicht eines Probanden bei der Fahrt mit dem Lane-Change-Test und der Durchführung von Interaktionsaufgaben

6.4.5. Zweite Baseline-Fahrt

Abschließend folgte noch die fünfte und letzte Fahrt mit dem Lane-Change-Test. Dabei handelte es sich wiederum um eine Baseline-Fahrt. D.h., eine Fahrt ohne Durchführung von Nebenaufgaben, deren Fahrleistung wie bei der ersten Baseline-Fahrt in der Auswertung zu Vergleichszwecken herangezogen wurde.

6.4.6. Abschließender Fragebogen

Die letzte Aufgabe für die Probanden bestand darin, einen abschließenden Fragebogen auszufüllen, in dem sie für die Bedienung der Objekte der Evaluierung, d.h. Außenspiegel, Scheibenwischer, Seitenscheiben und des Tempomaten einschätzen sollten, ob der Interaktionsstil Sprache und Gesten für diese Funktionen weniger ablenkend ist, als mit herkömmlichen Bedienelementen. Des Weiteren wurde noch abgefragt, welche zusätzlichen Objekte und Funktionen, die nicht Bestandteil der Evaluierung waren, sie sich für eine Bedienung mit Sprache und Gesten wünschen würden und welche Vor- und Nachteile sie bei der Interaktion mit Sprache und Gesten sehen. Nachdem auch dieser Fragebogen ausgefüllt war, bekamen die Teilnehmer der Evaluierung ihre Studienvergütung in Höhe von 10 € ausbezahlt und mussten den Empfang quittieren. Ein Muster des Abschlussfragebogens

findet sich im Anhang auf Seite 127, die Quittung auf Seite 129.

Ein Muster des Leitfadens, der für jeden Probanden individuell vervollständigt wurde, findet sich im Anhang ab Seite 130. Der gesamte Ablauf ist in Abbildung 6.4 nochmals zusammenfassend dargestellt.

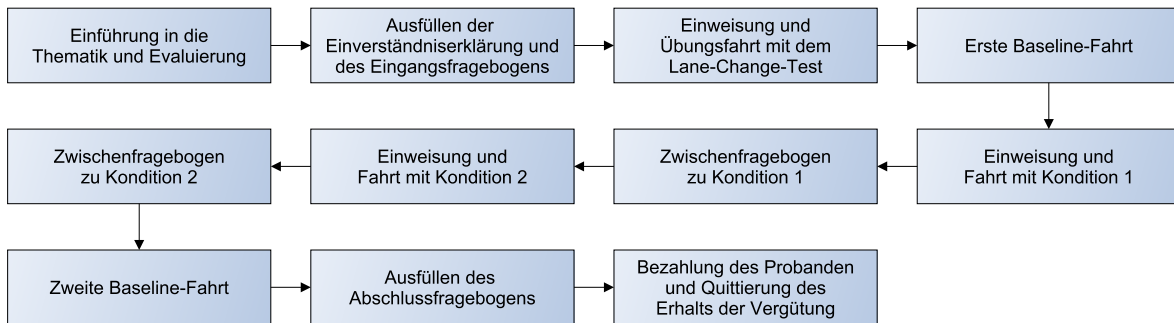


Abbildung 6.4.: Ablaufdiagramm der Evaluierung

6.5. Auswertung der Studie

Die Auswertung und Ergebnisse der Eingangs-, Zwischen- und Abschlussfragebögen werden im nachfolgenden Abschnitt dargelegt. Der sich anschließende Abschnitt schildert dann die Evaluierungsergebnisse des Lane-Change-Tests.

6.5.1. Auswertung der Fragebögen

Eingangsfragebögen

Allgemeine Informationen zu den Teilnehmern, die in den Eingangsfragebögen gesammelt wurden, sind bereits in Abschnitt 6.3 auf Seite 74 und Tabelle 6.2 auf Seite 75 angeführt worden. Ergänzend hierzu wurden in diesen Eingangsfragebögen Erfahrungen der Probanden hinsichtlich Spracheingabe und Sprachsteuerung abgefragt. 8 der 16 Probanden gaben hierbei an, keinerlei Vorerfahrungen auf diesem Gebiet zu besitzen. Wenn Vorerfahrungen angegeben wurden, dann bezogen sich diese hauptsächlich auf die Spracheingabe/-steuerung bezüglich Smartphones/Handys (6x genannt). Vereinzelt wurden auch (Tablet-)Computer (4x) und Fahrassistenzsysteme im Fahrzeug (2x) genannt.

Bezüglich Vorerfahrungen mit (Multi-)Touch-Displays dominierte ebenfalls der Bereich der Smartphones/Handys (12 der 16 Probanden). 4 Teilnehmer hatten bereits Erfahrung mit

Tablet- oder Touch-Table-Computern. Nur ein Proband verwies auf die naheliegende Interaktionserfahrung mit z.B. touch-basierten Fahrkartenautomaten. Gar keine Erfahrung konnten nur 3 Teilnehmer vorweisen.

Zwischenfragebögen beider Interaktionsstile

Die ersten Fragen, die die Probanden auf den Zwischenfragebögen für die Interaktionsstile Sprache und Gesten sowie Buttons und Knöpfe zu beantworten hatten, waren die zehn Fragen der System Usability Scale-Methode (SUS) [12]. Die SUS-Methode basiert auf einer Likert-Skala und ist als eine einfache Methode zur allgemeinen Betrachtung von subjektiven Bewertungen von Usability konzipiert worden. Die SUS-Methode hat sich als robust und zuverlässig erwiesen und wurde bereits in einer Vielzahl von Forschungsprojekten und industriellen Evaluierungen eingesetzt [12]. Für jeden Probanden wird dabei anhand eines vorgegeben Auswertungsverfahrens (siehe [12]) ein SUS-Score, d.h. eine Punktezahl berechnet, die zwischen 0 und 100 liegt und die Gesamt-Usability des untersuchten Systems kennzeichnet.

Die zehn Fragen der SUS-Methode finden sich auf den Zwischenfragebögen im Anhang ab Seite 119. Die Fragebögen für beide Interaktionsstile unterscheiden sich nur dadurch, dass die Formulierungen „Sprache und Gesten“ und „Buttons und Knöpfe“ austauschbar sind. Abbildung 6.5 stellt die jeweiligen SUS-Punktezahlen pro Proband für jede der beiden Interaktionsstile gegenüber.

Abbildung 6.5 zeigt, dass die Usability des Interaktionsstils Sprache und Gesten nur von zwei Probanden höher bewertet wurde, als die Usability des nachgestellten herkömmlichen Interaktionsstils. Bei Proband-Nr. 10 fiel dieser Bewertungsunterschied noch relativ gering aus, bei Proband-Nr. 16 hingegen, wurde die Usability der Interaktion mit Sprache und Gesten fast um das doppelte höher eingestuft. In zwei Fällen wurde die Usability beider Interaktionsformen gleich eingeschätzt, es wird jedoch deutlich, dass die Probanden insgesamt mit der Interaktion durch Nutzung der Wii-Controller und der Android-Applikation aus Abbildung 6.2 auf Seite 74 besser zurechtkamen. Dies unterstreicht den Eindruck während der Einweisung in beide Interaktionsstile, bei der die Teilnehmer die nachgestellte herkömmliche Interaktionsform schneller beherrschten. Die mittlere SUS-Punktezahl für den Interaktionsstil Sprache und Gesten lag bei 69,06 und für den Interaktionsstil Buttons und Knöpfe bei 79,38.

Die nächsten sieben Fragen verlangten eine Einschätzung des Grades der Beanspruchung, den die Probanden bezüglich sieben verschiedener Faktoren empfanden. Diese Faktoren beruhen auf der Driving Activity Load Index-Methode (DALI), die von der NASA-TLX-Methode abgeleitet wurde [54]. Die DALI-Methode erlaubt eine subjektive Evaluierung der mentalen Auslastung, speziell für den Automobilkontext. Dazu sollten von den Probanden

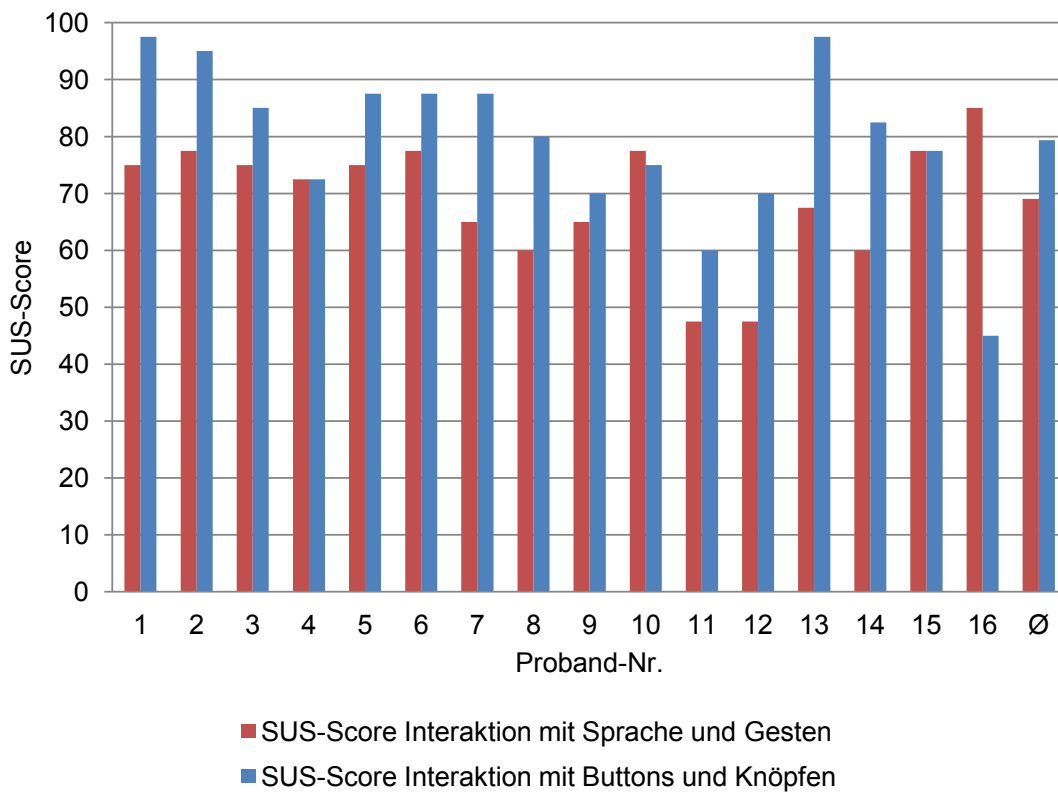


Abbildung 6.5.: SUS-Punktezahlen je Proband und Interaktionsstil

die Anforderungen an die gesamte Aufmerksamkeit, die visuelle Anforderung, auditive Anforderung, manuelle Anforderung, Stress, zeitliche Anforderung und Interferenz auf einer Skala von Niedrig(0) bis Hoch(5) eingeschätzt werden. Eine genauere Erläuterung dieser sieben Faktoren findet sich auf den Zwischenfragebögen im Anhang ab Seite 121.

Wie zuvor angegeben, wurde zur Einschätzung der Beanspruchung für die Faktoren jeweils eine Skala von 0 bis 5 vorgegeben. Die Auswertung der DALI-Faktoren erfolgte auf die Art, dass der Mittelwert und die Standardabweichung für jeden Faktor, jeweils über alle 16 Probanden hin gesehen, berechnet wurden. Abbildung 6.6 zeigt die mittleren Werte der Beanspruchung für alle sieben Faktoren, jeweils im Vergleich der Interaktionsstile Sprache und Gesten und Buttons und Knöpfe.

Die größte Diskrepanz ist bezüglich der mittleren auditiven Anforderung an die Probanden zu sehen. Dies kann damit zusammenhängen, dass bei der Interaktion mit Sprache und Gesten, jedes gültige sowie unvollständig qualifizierte Sprachkommando eine akustische Ausgabe über einen Signalton und eine Computerstimme zur Folge hatten (siehe Kapitelabschnitt 5.3 auf Seite 62). Dadurch lässt sich die deutlich höhere Anforderung an das Gehör erklären. Manuelle und visuelle Anforderung waren wiederum bei der Interaktion

6. Evaluierung des Prototypen und Interaktionsstils

mit Buttons und Knöpfen höher. Einige Teilnehmer erklärten diesbezüglich am Ende der Evaluierung, dass sie beispielsweise die Wii-Controller zuerst lokalisieren mussten und im Gegensatz dazu bei der Interaktion mit Sprache und Gesten alle Gesten zentral auf dem Tablet-Computer ausführen konnten. Die Anforderung an die gesamte Aufmerksamkeit war bei der Interaktion mit Sprache und Gesten zwar höher, dies lässt sich auch durch den ungewohnten Interaktionsstil erklären, den die Probanden in sehr kurzer Zeit zu erlernen hatten. Dazu gehörte auch, dass sie zum einen die Sprachkommandos, zum anderen die Gesten aus dem Gedächtnis abrufen mussten. Dass der zeitliche Druck bei der Interaktion mit Sprache und Gesten im Mittel höher eingestuft wurde, kann damit zusammenhängen, dass das System bei einigen Probanden das Sprachkommando nicht sofort oder falsch erkannte und sie den Vorgang daraufhin wiederholen mussten. Dies erhöhte unter Umständen den Druck auf die Probanden, eine Aufgabe aufgrund der zeitraubenden Fehlschläge in Kürze abschließen zu müssen. Ebenso kann dies eine Erklärung für das leicht erhöhte Stressempfinden sein.

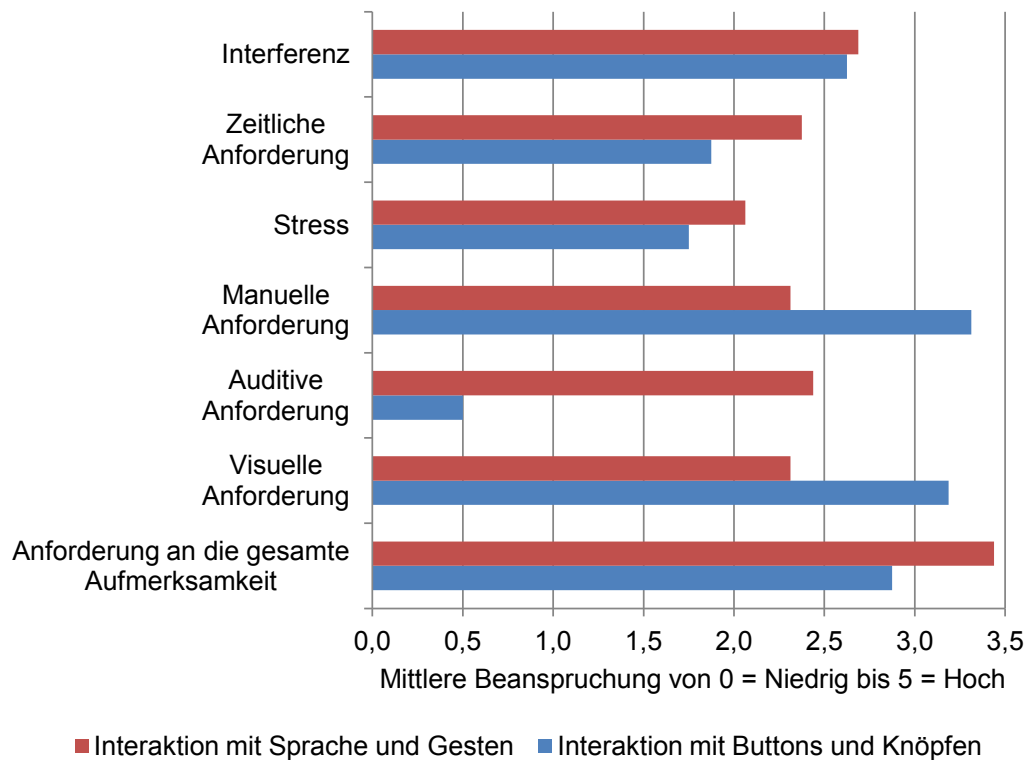


Abbildung 6.6.: Analyseergebnisse der mittleren Beanspruchung für jeden DALI-Faktor

Mit der allerletzten Frage auf den Zwischenfragebögen, sollte zu jedem Interaktionsstil angegeben werden, ob er den Probanden gefiel. 15 der 16 Teilnehmer bejahten dies für die Interaktion mit Sprache und Gesten, 14 der 16 Teilnehmer für die Interaktion mit Buttons und Knöpfen.

Abschlussfragebögen

Im abschließenden Fragebogen waren die Probanden dann nochmals aufgefordert, zu folgender Aussage Stellung zu nehmen: „*Ich halte eine Bedienung über die Kombination Sprachkommando und Touch-Geste für diese Aufgabe für weniger ablenkend als mit herkömmlichen Bedienungselementen.*“ Diese Aussage sollten sie jeweils in Bezug auf die Einstellung der Außenspiegel, die Bedienung der Scheibenwischer, der Seitenscheiben und des Tempomaten anwenden. Im Rahmen einer Likert-Skala hatten sie die Antwortmöglichkeiten „*stimme voll und ganz zu*“ (5), „*stimme zu*“ (4), „*unentschieden*“ (3), „*stimme nicht zu*“ (2) und „*stimme gar nicht zu*“ (1). Durch die Zuweisung der jeweils in Klammern angegebenen Werte, konnte für all diese Aufgaben jeweils ein Mittelwert über alle Probanden hin gesehen berechnet werden. Für die Einstellung der Außenspiegel lag dieser bei 3,13, für die Bedienung der Scheibenwischer bei 2,69, für das Öffnen/Schließen der Seitenscheiben ebenfalls bei 2,69 und für die Einstellung des Tempomaten bei 3,00. Dies kann so interpretiert werden, dass die Teilnehmer der Evaluierung hier eher unentschieden waren, ob diese neue Form der Interaktion mit Sprache und Gesten wirklich weniger ablenkend ist.

Vorstellbar war für immerhin 11 der 16 Probanden, die Interaktion mit Sprache und Gesten gegenüber der herkömmlichen Interaktion bevorzugt einzusetzen.

An zusätzlichen Funktionen zur Bedienung mit Sprache und Gesten, die nicht Bestandteil der Evaluierung waren, gaben die Probanden am häufigsten die Steuerung von Radio-, Musik- und Multimediafunktionen an. Darauf folgten die Bedienung des Navigationssystems und von Klimafunktionen.

Ein paar Teilnehmer nutzten diese Freitext-Frage, um ganz andere Wünsche mitzuteilen. So hätten zwei Probanden lieber mehr Sprachkommandos (z.B. „*Fenster halb öffnen*“) und weniger Gesten, einer würde sogar eine reine Sprachsteuerung bevorzugen. Ein Proband wollte den Tablet-Computer in den Bereich der Mittelkonsole auslagern, um so nach seinen Angaben besser lenken zu können. Auch Anzeigen direkt auf dem Tablet-Computer wurden gewünscht.

An Vorteilen bezüglich der Interaktion mit Sprache und Gesten gegenüber der herkömmlichen Interaktion wurde vor allem die geringere Ablenkung genannt. Die Auswahl der Steuerelemente kostet demnach weniger Aufmerksamkeit und das Fokussieren von Bedienelementen entfällt. Diese Aussagen gingen Hand in Hand mit Aussagen wie, dass Bedienelemente damit zentral am selben Ort liegen und nicht mehr gesucht werden müssen. Des Weiteren würde man so Platz für Steuerelemente einsparen und hätte weniger Knöpfe im Fahrzeug.

Die Nachteile lagen vor allem im Bereich der befürchteten Fehlerkennungen der Sprachkommandos und Fehlinterpretationen von Gesten. Aber auch, dass die Durchführung einer Aktion länger dauern würde, weil z.B. Spracheingabe an sich langsamer sei. Außerdem liegt bereits eine Gewöhnung an den herkömmlichen Interaktionsstil vor, sodass dieser von einem

Probanden auch als intuitiver und schneller eingeschätzt wurde. Bemängelt wurden auch Ungenauigkeiten bei den Einstellungen mit Gesten, die dann eine Korrektur erforderten. Von zwei Probanden wurde darüber hinaus auch die Geräuschkulisse als Hindernis eingestuft. Fehlendes Feedback, sodass die korrekte Ausführung einer Funktion nachträglich überprüft werden musste, wurde von drei Probanden angemerkt.

6.5.2. Analyse der aufgezeichneten Lane-Change-Test-Daten

Abbildung 6.7 zeigt die jeweilige Anzahl der von den einzelnen Probanden während der Fahrt mit den beiden unterschiedlichen Interaktionsstilen durchgeführten Aufgaben. Für die Abfolge und Aufbau dieser Aufgaben sei an Abschnitt 6.4.4 auf Seite 79 und Tabelle 6.1 auf Seite 72 erinnert. Mit dem Interaktionsstil Sprache und Gesten konnte kein einziger Proband mehr Aufgaben durchführen als mit dem Interaktionsstil Buttons und Knöpfe. Maximal konnten 15 Aufgaben (im Durchschnitt 12,06) mit der Interaktion über Sprache und Gesten während einer Fahrt erledigt werden und mit der Interaktion über Buttons und Knöpfe maximal 21 (im Durchschnitt 18,88).

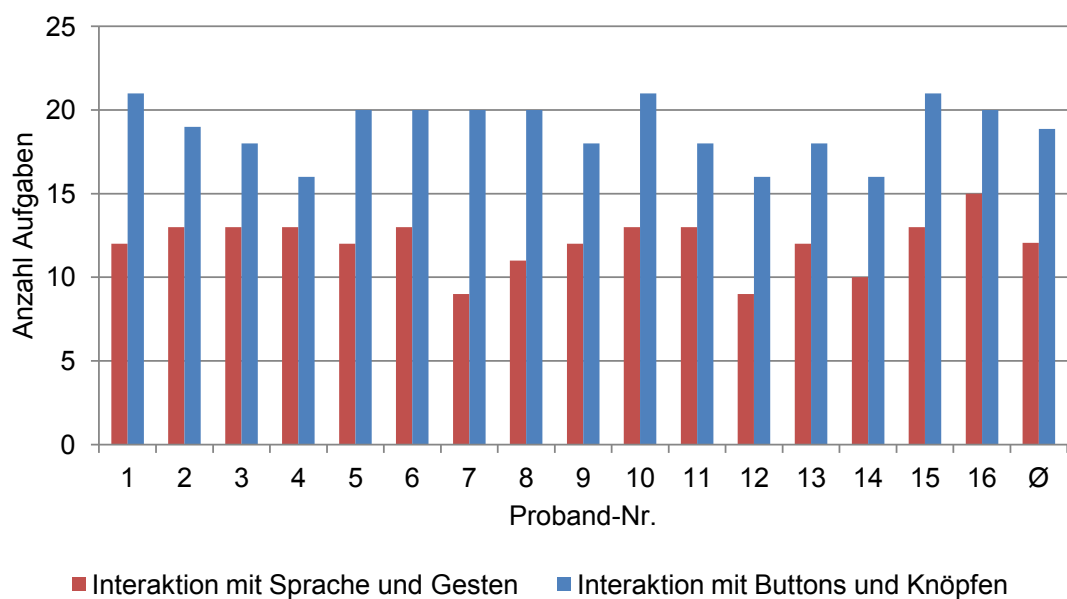


Abbildung 6.7.: Darstellung der Anzahl der durchgeführten Interaktionsaufgaben je Proband und Interaktionsstil während der Evaluierung

Wie in Abschnitt 6.4.3 auf Seite 77 bereits erläutert, musste das Lane-Change-Test-Programm nach jeder Fahrt beendet werden, damit eine Analyse-Datei erzeugt wird. Von den fünf Fahrten pro Teilnehmer war die anfängliche Übungsfahrt nicht für die Auswertung von Relevanz und somit waren es am Ende für jeden Probanden vier Analyse-Dateien, die zur Auswertung herangezogen wurden. Mit einem Analyse-Programm, das Bestandteil des Lane-Change-Tests ist, wurde die mittlere Fahrabweichung, die Standardabweichung und zu Kontrollzwecken die Länge der gefahrenen Strecke für jeden Probanden und für jede ihrer vier relevanten Fahrten (Baseline-Fahrt 1, Interaktionsfahrt 1, Interaktionsfahrt 2 und Baseline-Fahrt 2) berechnet. Wie in Abbildung 2.3 auf Seite 20 illustriert, wird die mittlere Fahrabweichung als Abweichung der tatsächlichen Fahrleistung von einem normativen Modell berechnet. Abbildung 6.8 zeigt die berechneten mittleren Fahrabweichungen für alle Fahrten über jeweils alle Probanden hin gesehen. Da das Analyse-Tool bereits die mittlere Fahrabweichung für jede Einzelfahrt berechnet, wäre die korrektere Bezeichnung eigentlich mittlere mittlere Fahrabweichung. Im Folgenden wird die kurze Formulierung verwendet, es sei jedoch darauf hingewiesen, dass damit stets die mittlere mittlere Fahrabweichung gemeint ist.

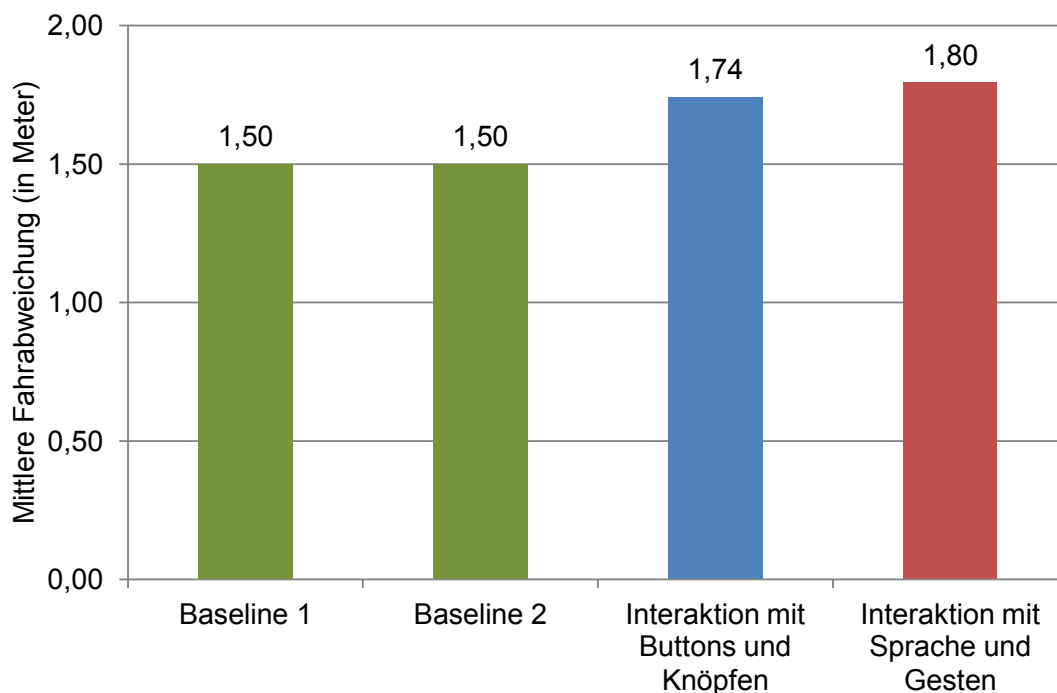


Abbildung 6.8.: Mittlere Fahrabweichung aller vier Fahrtstypen der Evaluierung

6. Evaluierung des Prototypen und Interaktionsstils

Die mittleren Fahrabweichungen der Baseline-Fahrten betragen jeweils 1,50 Meter und stimmen auf zwei Nachkommastellen gerundet überein. Selbst die Standardabweichungen der beiden Baseline-Fahrten sind auf zwei Nachkommastellen gerundet identisch. Abbildung 6.8 zeigt, dass die mittleren Fahrabweichungen der Interaktionsfahrten erwartungsgemäß über denen der Baseline-Fahrten liegen, bei denen keine Nebenaufgaben durchgeführt werden mussten.

Die Interaktionsfahrt mit Sprache und Gesten wurde gegenüber der Interaktionsfahrt mit Buttons und Knöpfen auf Signifikanz geprüft, jedoch konnte kein signifikanter Unterschied festgestellt werden, da die Wahrscheinlichkeit, dass die ermittelten Fahrabweichungen der einzelnen Fahrten durch Zufall zustande gekommen sind bei ca. 32% liegt. Aufgrund dieses Ergebnisses kann die Nullhypothese, die keinen Unterschied annimmt, auch nicht abgelehnt werden. Eine Erklärung dafür kann die geringe Anzahl der Evaluierungsteilnehmer (16) sein. Auch wenn kein signifikanter Unterschied festgestellt werden kann, scheint der Unterschied der mittleren Fahrabweichungen beider Interaktionsstile jedoch gering zu sein, sodass die Interaktion mit Sprache und Gesten keine unüberwindbaren Nachteile im Hinblick auf die Fahrerablenkung mit sich zu bringen scheint.

7. Zusammenfassung und Ausblick

In dieser Arbeit wurde ein neuartiges automobiles Interaktionskonzept aufgestellt und untersucht. Im Hinblick auf immer neue Funktionen im Fahrzeug und damit einhergehende gestiegene Anforderungen an den Fahrer, sollte dieses Konzept auf einen möglichen Beitrag zur Senkung der Fahrerablenkung hin erforscht werden. Das Ziel war es, mit einer multimodalen Eingabeform, die die Modalitäten Sprache und Gestik vereint, eine möglichst intuitive Bedienung von Funktionen im Fahrzeug zuzulassen. Dieser Interaktionsstil sollte im besten Fall dazu beitragen die Fahrerablenkung zu reduzieren.

Die Motivation für die Vereinigung speziell der Modalitäten Sprache und Gestik beruhte auf der Annahme, dass Nachteile jeder einzelnen Modalität durch die geschaffene Kombination vermindert und durch die entstehenden Vorteile verschiedene Usability-Aspekte adressiert werden können.

Zur Erstellung des Interaktionskonzeptes wurden verschiedene Vor- und Nachteile einzelner Eingabemodalitäten näher betrachtet. So ist der naheliegende zentrale Vorteil von Spracheingabe im Automobil vor allem der, dass die Hände auf dem Lenkrad verweilen können. Doch ist Spracheingabe alleine keine Lösung um gestiegene Fahrerablenkung wieder einzudämmen. So sind vor allem Feineinstellungen mit Sprache schwer zu bewerkstelligen, da natürlichsprachliche Eingabe zum einen an aktuelle technische Grenzen stößt und zum anderen die Formulierung eines solchen Kommandos wiederum kognitive Auslastung nach sich ziehen würde. Dies trifft vor allem dann zu, wenn Korrekturen oder Nachjustierungen nötig werden, die mit Sprache alleine insgesamt zu viel Zeit in Anspruch nehmen würden.

Aus den genannten Gründen war die Modalität Sprache insoweit Teil des Interaktionsstils, dass mit Sprachkommandos sichtbare Objekte bzw. deren Funktionen in der Umgebung des Fahrers ausgewählt werden können. Dabei beschränken sich diese Sprachbefehle auf kurze, prägnante und vor allem eindeutige Kommandos. Dies ist gleichzeitig ein Beitrag zur Verminderung eventueller Fehlerkennungen, wenn nur kurze Wortfolgen erkannt werden müssen. Mit diesem Vorgehen wird dem Benutzer kognitive Arbeit abgenommen, da er Interaktionsobjekte direkt vor sich sieht und sich so an Sprachkommandos schneller erinnern kann. Beispiele für solche Kommandos sind z.B. „Fahrerfenster“ oder „Tempomat“. Wie an diesen Beispielen zu sehen, genügt es bereits das Objekt an sich zu benennen, wenn es nur

eine einzige Funktion hat (z.B. Fensterfenster nach oben/unten fahren oder Tempomatgeschwindigkeit erhöhen/reduzieren). Wenn ein Objekt mehrere Funktionen anbietet, muss die jeweilige Funktion in das Sprachkommando mit aufgenommen werden. Beispielsweise also „*Head-Up-Display-Helligkeit*“. Falls ein Sprachkommando unvollständig qualifiziert war, muss das System fehlertolerant sein und Hilfe anbieten können um den Benutzer seine Eingabe wiederholen zu lassen.

Was mit Sprache nur schwer zu lösen ist, kann mit Gesten sehr effektiv durchgeführt werden: fein-granulare Einstellungen lassen sich mit Gesten leichter durchführen. Außerdem können im Sinne einer guten Usability, Änderungen leicht zurückgenommen oder korrigiert werden. Falls also z.B. eine Fensterscheibe nach oben gefahren werden soll, der Fahrer aber merkt, dass sie zu weit geöffnet wurde, kann er bei entsprechenden intuitiven Gesten eine Geste in die entgegengesetzte Richtung durchführen und so die Fensterscheibe wieder ein Stück herablassen. In dem betrachteten Fahrzeugumfeld bekommt der Fahrer des Weiteren ein sofortiges Feedback auf eine durchgeführte Gestenmanipulation.

Von all diesen Vorteilen von Gesten wurde im Rahmen der multimodalen Eingabebform in dieser Arbeit Gebrauch gemacht. Nachdem ein Objekt oder eine Funktion mit der Modalität Sprache ausgewählt war, fand ein Wechsel der Modalitäten statt, sodass mit einer Geste das zuvor ausgewählte Objekt oder die Funktion manipuliert bzw. durchgeführt werden konnte. Die Gesteneingabe fand dabei auf einem multitouch-fähigen Tablet-Computer statt, der wiederum direkt auf dem Lenkrad befestigt war, sodass dieser Interaktionsstil eine zentralisierte Bedienung ermöglicht.

Nach Fertigstellung eines tragbaren Konzeptes, war der nächste Schritt die Sammlung von Sprachkommandos und Gesten von tatsächlichen Benutzern. Dies wurde in Form einer Studie bewältigt, für die zuvor eine Auswahl an Objekten und Funktionen zusammengestellt wurde. Dazu zählten beispielsweise die Seitenfenster, die Scheibenwischer oder die Sitzheizung. Von 12 Probanden wurden im Rahmen dieser Studie Sprachbefehle zur Auswahl dieser Objekte und Funktionen gesammelt sowie Gesten für deren Manipulation. Es zeigte sich, dass der vorgeschlagene Interaktionsstil aus Sprache und Gesten von potentiellen Benutzern verstanden wird und es entstand in der Folge ein benutzerdefiniertes Set an Sprachkommandos und Gesten.

Eine große Bandbreite an verschiedensten Variationen von Sprachbefehlen wurde während der Studie aufgezeichnet, was die Notwendigkeit eines möglichst breit angelegten Befehlsspektrums einer Spracherkennungstechnologie im Automobil deutlich macht. Der Wiedererkennungsfaktor der den Studienteilnehmern präsentierten Objekte war sehr hoch, bis auf ein Head-Up-Display, das von wenigen Probanden nicht identifiziert werden konnte, wurden alle Objekte mit Sprachkommandos verbunden. Daraus kann gefolgert werden, dass die verbale Auswahl sichtbarer Objekte ein hohes Potential an intuitiver Bedienung bietet.

Bezüglich der während dieser Studie gesammelten Gesten, gab es eine hohe Übereinstimmung zwischen den Probanden. So bestand der Großteil der Gesten aus mit dem Finger nur nach links, rechts, oben oder unten geführten Bewegungen. Da die große Mehrzahl der Probanden auf Konsistenz für Gesten von jeweils konträren Funktionen achtete, ließ sich in der Folge ein ebenso konsistentes wie logisches Gestenset entwickeln. Lediglich für komplexere Funktionen wie die der Scheibenwischer, gab es eine weniger hohe Übereinstimmungsrate zwischen den Teilnehmern der Studie. Waren die zuvor angesprochenen einfachen Richtungsgesten noch intuitiv, so trifft diese Eigenschaft auf diese Gesten in weniger hohem Maße zu. Dies unterstreichen längere Denkprozesse der Probanden bis Gesten für komplexere Funktionen erdacht wurden.

Die Studie brachte in der Folge die notwendigen Grundlagen für die Konzipierung und die Entwicklung eines Prototypen, der die erstellten Sets aus Sprachkommandos und Gesten integrierte und den Interaktionsstil umsetzte. In einer Laborumgebung wurde dazu ein Aufbau mit vier Bildschirmen installiert, sodass ein Benutzer vor sich einen Fahrsimulator und ein virtuelles Armaturenbrett, links und rechts Seitenfenster und -spiegel sowie hinter sich eine Heckscheibe mit Scheibenwischer präsentiert bekam. Dieser Prototyp wurde so konstruiert, dass sich komplette Interaktionsvorgänge durchführen ließen. D.h., zuerst konnte die Spracherkennung u.a. mit einem Pull-To-Talk-Mechanismus aktiviert werden, dann ein Sprachkommando zur Objekt-/Funktionsauswahl geäußert und anschließend eine Touch-Geste auf einem in eine Lenkradoberfläche eingebetteten Multitouch-Tablet-Computer zur Manipulation eingesetzt werden. Eine unabhängige Implementierung der verschiedenen Ein- und Ausgabekomponenten lässt des Weiteren einen individuellen Austausch selbiger zu.

Der letzte Teil der Arbeit beschäftigte sich mit einer Evaluierung, in der 16 Probanden den Prototyp mit zwei unterschiedlichen Interaktionsstilen verwendeten. Dazu wurde die Interaktion mit Sprache und Gesten einer herkömmlichen, manuellen Interaktion gegenübergestellt. Diese herkömmliche Interaktion wurde so nachgestellt, dass eine Bedienung verschiedener Funktionen mit Nintendo Wii-Controllern und einer speziellen Applikation auf dem Tablet-Computer möglich war. In der als Vergleichsevaluierung aufgesetzten Studie, wurden die Teilnehmer mit einem Fahrsimulator, dem Lane-Change-Test [46] konfrontiert, mit dem sie mehrere kurze Fahrten zu absolvieren hatten. Von den insgesamt vier Fahrten je Proband waren zwei Fahrten ohne Interaktionsaufgaben durchzuführen und zwei Fahrten mit Nebenaufgaben. Bei diesen zwei Fahrten wurden mehrere Aufgaben vorgegeben, die während der einen Fahrt mit dem Interaktionsstil Sprache/Gestik durchgeführt werden sollten und während der anderen Fahrt mit der nachgestellten, herkömmlichen Interaktionstechnik. Die durchzuführenden Aufgaben bestanden beispielsweise darin, einen Heckscheibenwischer auf eine bestimmte Funktion wie Dauer- oder Intervallwischen einzustellen oder die Seitenspiegel auf eine bestimmte Einstellung zu drehen.

Die Auswertung dieser Evaluierung zeigte, dass 12 der 16 Evaluierungsteilnehmer die Usability der Bedienung mit Sprache und Gesten geringer einstufen, als die Usability der herkömmlichen Bedienung. Diese Usability-Bewertung wurde mit der System Usability Scale-Methode (SUS) [12] vorgenommen, die auch zeigte, dass der Unterschied der mittleren Usability-Bewertung zwischen beiden Interaktionsstilen allerdings nicht groß ist. Mit der Driver Activity Load Index-Methode (DALI) [54] sollten die Probanden zudem die empfundene subjektive Beanspruchung hinsichtlich mehrerer Faktoren einschätzen. Dabei wurde ersichtlich, dass die Probanden im Mittel zwar eine höhere Anforderung an die gesamte Aufmerksamkeit beim Einsatz von Sprache und Gesten sahen, jedoch ist auch hier der Unterschied zum herkömmlichen Interaktionsstil nicht gravierend. Gerade bezüglich visueller Anforderung schnitt die multimodale Interaktion mit Sprache und Gesten besser ab, als die herkömmliche. Dies lag vor allem daran, dass hier die manuellen Steuerelemente zuerst lokalisiert werden mussten. Und eben dies ist eine Situation aus dem realen Automobilkontext, sodass der untersuchten Interaktionsform hier ein gewisses Potential bescheinigt werden kann.

Grundsätzlich konnten sich zweidrittel aller Probanden vorstellen, Sprache und Gesten gegenüber der manuellen Interaktion bevorzugt einzusetzen. Dennoch schaffte es kein einziger Proband, während der Fahrt mit dem Fahrsimulator mehr oder gleich viele Aufgaben mit dem multimodalen Interaktionsstil durchzuführen, als mit dem herkömmlichen. In dieser Hinsicht war ihnen die herkömmliche Interaktion vertrauter und fiel ihnen leichter. Als weiteres Resultat gab der Einsatz des Fahrsimulators Aufschluss darüber, wie hoch bei den Fahrten mit beiden Interaktionsformen jeweils die mittlere Fahrabweichung, d.h. die Abweichung der tatsächlichen Fahrleistung von einer Referenzlinie ist. Dabei konnte zwar kein signifikanter Unterschied zwischen den mittleren Fahrabweichungen beider Interaktionsstile festgestellt werden, doch lagen die ermittelten Ergebnisse für beide Stile auch hier nicht weit auseinander, sodass die Fahrerablenkung durch die multimodale Interaktion nicht eindeutig erhöht wurde.

Insgesamt konnten alle Teilnehmer der Evaluierung sowohl Sprachkommandos wie auch Gesten bereits nach einer kurzen Einweisungsphase einsetzen und eine längere Lernphase kann somit die Handhabung dieses Interaktionsstils weiter verbessern. Auch stellte sich die Integration eines breiten Wortschatzes an Sprachkommandos als sehr positiv heraus, so nutzten manche Probanden auch Sprachbefehle, die ihnen eigentlich nicht vorgegeben waren, aber vom System dennoch erkannt wurden.

Ausblick

Die abschließende Evaluierung offenbarte ein hohes Potential, die eine weitere Erforschung der Interaktion mit Sprache und Gesten unterstreicht. Dass die Usability beim Einsatz von Sprache und Gestik im Vergleich zu einer herkömmlichen Interaktion als geringer eingestuft wurde, hängt auch mit der speziellen Realisierung des Prototypen zusammen. Hier bestehen Möglichkeiten der Erweiterung, z.B. durch Erschaffung einer realistischeren Umgebung mit mehr Feedback wie Geräuschen oder Anzeigen.

Da der aktuelle Ansatz nur auf Aktionen, die vom Fahrer ausgehen bezogen war, ließe sich der Interaktionsstil auch auf den Beifahrer ausweiten, der damit weniger fahrrelevante Funktionen wie Info- oder Entertainmentsysteme ansteuern könnte. Ein intelligentes System würde durch Mikrophone auf beiden Seiten zudem erkennen können, welcher Fahrzeuginsasse ein Kommando geäußert hat und dementsprechend den Kontext der zulässigen Funktionen einschränken.

Durch den Einsatz eines Eye-Trackers könnten Fehlerkennungen von Sprachkommandos vermieden werden, indem das geäußerte Sprachkommando und das aktuell fokussierte Objekt verglichen werden und das System so beispielsweise bei lauten Nebengeräuschen trotzdem die Auswahl des richtigen Objektes durchführen kann.

Weitere Untersuchungen in Form von Studien können in mehreren Belangen zielführend sein. Es ließen sich so die bereits erfassten Sprachkommandos nochmals evaluieren und auf fehlende Sprachbefehle testen oder bestehende Kommandos im Sinne eines kleineren Wortschatzes verwerfen. Selbiges gilt für das bestehende Gestenset: zwar waren die meisten Gesten einfache Richtungsgesten, doch trifft dies speziell auf einfache Funktionen zu. Für komplexere Funktionen müssen Gesten demnach besonders sorgfältig ausgewählt werden, um annähernd intuitiv zu sein.

Eine Langzeitevaluierung mit mehr Probanden würde auch stichhaltigere Aussagen darüber zulassen, inwiefern die Fahrerablenkung durch den Interaktionsstil mit Sprache und Gesten tatsächlich beeinflusst wird. So ist vor allem auf Lerneffekte unter den Teilnehmern zu achten und inwieweit sich diese auf das Fahrverhalten niederschlagen. Auch kann so die Akzeptanz nach einer längeren Nutzung besser eingeschätzt werden.

Eine Übertragung des Aufbaus des Prototypen in ein reales Fahrzeug ist nicht eins zu eins möglich und zieht weitere Überlegungen nach sich. So kann z.B. ein Tablet-Computer nicht über einem Airbag auf dem Lenkrad befestigt werden. Hier muss nach Lösungsmöglichkeiten gesucht werden, die beispielsweise Technologien wie flexible OLED-Touch-Displays bieten.

A. Anhang

A.1. Einverständniserklärung - Explorative Benutzerstudie

Einverständniserklärung zur Studie „User-defined Voice Commands and Touch Gestures“

Teilnehmer-Nr.: _____ Name: _____
Vorname: _____

Mir wurde der Zweck der Studie erläutert und ich wurde darüber informiert, dass ich die Studie jederzeit abbrechen kann.

Ich wurde des Weiteren darüber in Kenntnis gesetzt, dass die während der Studie gesammelten Video- und Audiodaten zur Auswertung der Studie herangezogen werden. Sämtliche Daten werden für die wissenschaftliche Nutzung gesammelt und hierbei vertraulich behandelt.

Der Verwendung von Bildmaterial mit meiner Person in Präsentationen, Ausarbeitungen und Veröffentlichungen (auch online)

- stimme ich zu.
 stimme ich nicht zu.

Stuttgart, den

Unterschrift:

Teil B – Erfahrungen des Probanden

B.1 Haben Sie bereits Erfahrungen mit Spracheingabe oder Sprachsteuerung sammeln können?

- ja
 nein

Falls ja, in welchen Bereichen?: _____

B. 2 Haben Sie bereits Erfahrungen mit (Multi-)Touch-Displays?

- ja
 nein

Falls ja, in welchen Bereichen?: _____

A.3. Zwischenfragebogen - Explorative Benutzerstudie

Zwischenfragebogen zur Studie „User-defined Voice Commands and Touch Gestures“

Teilnehmer-Nr.: _____ Name: _____

Vorname: _____

Aussage: „Die Kombination aus dem eben genannten Sprachkommando und der Touch-Geste ist während einer realen Autofahrt [.....] auszuführen.“

Aufgaben-ID	sehr einfach	einfach	mittelschwer	schwer	sehr schwer

A.4. Abschlussfragebogen - Explorative Benutzerstudie

Abschlussfragebogen zur Studie „User-defined Voice Commands and Touch Gestures“

Teilnehmer-Nr.: _____ Name: _____

Vorname: _____

Teil C – Allgemeine Bewertung der multimodalen Eingabe

C.1 Wie schätzen Sie die allgemeine Akzeptanz hinsichtlich der Kombination Sprachkommandos und Touch-Gesten im Automobilbereich ein?

- gar keine Akzeptanz
- geringe Akzeptanz
- durchschnittliche Akzeptanz
- hohe Akzeptanz
- sehr hohe Akzeptanz

C.2 Haben Sie eine Hand für die Eingabe der Gesten vom Lenkrad genommen?

- ja
- nein

C.3 Haben Sie bei der Touchgesten-Eingabe bestimmte Finger bevorzugt eingesetzt?

- ja und zwar folgende Finger: _____
- nein

C.4 Haben Sie bei Ihren Gesten auf Konsistenz bezüglich vorher eingegebener Gesten geachtet?

- ja
- nein

C.5 Wieviele Finger würden Sie für eine Touch-Geste maximal einsetzen wollen? _____

Teil D – Beurteilung der Funktionsspektrums

D.1 Nehmen Sie jeweils zu den nachstehenden Aufgaben im Automobil Stellung bezüglich der Aussage:

„Ich halte eine Bedienung über die Kombination Sprachkommando und Touch-Geste für diese Aufgabe für sinnvoll.“

(Kreuz im zutreffenden Feld setzen)

	stimme voll und ganz zu	stimme zu	unentschieden	stimme nicht zu	stimme gar nicht zu
Sitzlängsrichtung ändern					
Sitzlehne neigen					
Sitzhöhe ändern					
Kopfstützenhöhe ändern					
Sitzheizung regulieren					
Außenspiegel einstellen					
Scheibenwischer bedienen					
Seitenscheiben öffnen/schließen					
Tempomat einstellen					
Luftdüse einstellen					
Head-Up-Display einstellen					

D.2 Welche zusätzlichen Funktionen, die nicht Teil der Studie waren, würden Sie sich bezüglich einer Bedienung mit der Kombination Sprachkommandos und Touch-Gesten noch wünschen?:

Teil E – Beurteilung der Interaktionspotentiale

E.1 Welche Vorteile sehen Sie bei der Interaktion mit Sprachkommandos und Touch-Gesten?

E.2 Welche Nachteile sehen Sie bei der Interaktion mit Sprachkommandos und Touch-Gesten?

Vielen Dank für Ihre Teilnahme!

A.5. Quittung - Explorative Benutzerstudie

Studie „User-defined Voice Commands and Touch Gestures“

Bestätigung über den Erhalt der Studienvergütung

Teilnehmer-Nr.: _____

Name: _____

Vorname: _____

Die Vergütung für die Studie in Höhe von 5,- Euro habe ich erhalten.

Stuttgart, den

Unterschrift:

A.6. Leitfaden - Explorative Benutzerstudie

Leitfaden

Einführung in die Studie

- „Die Studie ist Teil meiner Diplomarbeit, in der eine neue Form der Interaktion im Automobil untersucht werden soll: Nämlich eine Kombination von Sprachkommandos und Touch-Gesten.“
- „Mit Sprachkommandos sollen Objekte oder Objektfunktionen eindeutig ausgewählt werden. In einem weiteren Schritt wird mittels Touch-Gesten dann eine Einstellung auf den ausgewählten Objekten oder Objektfunktionen durchgeführt. Eine Interaktion besteht somit aus zwei Teilen.“
- „In der Studie sollen nun von den Studienteilnehmern Sprachkommandos zur Objektauswahl/Objektfunktionsauswahl und Touch-Gesten zu deren Manipulation gesammelt werden.“
- „Während der Studie werden Sprach- und Videoaufzeichnungen gemacht, welche zur Auswertung der Studie herangezogen werden. Die gesammelten Daten werden vertraulich behandelt.“
- „Die Studie kann selbstverständlich jederzeit abgebrochen werden.“
- „Die Studie beginnt mit einem Eingangsfragebogen und einer Einverständniserklärung, dass die gesammelten Daten verwendet werden dürfen. Nach Ende des praktischen Teils der Studie gibt es noch einen Abschlussfragebogen.“

Eingangsfragebogen und Einverständniserklärung an den Probanden aushändigen und ausfüllen lassen

- Einverständniserklärung vorlesen

Detaillierte Erläuterung der Vorgehensweise

- „Stell dir vor, du sitzt in einem Oberklasse-Fahrzeug, das viele Extras und Einstellmöglichkeiten hat und auf einem bestimmten Objekt auch mehrere verschiedenartige Funktionen ausgeführt werden können.“
- „Mehrere Aufgabenpaare werden nun in zufälliger Reihenfolge präsentiert.“
- „Für den ersten Teil eines Aufgabenpaares wird ein Bild eingeblendet und die Aufgabe besteht jeweils darin das dort markierte Objekt oder wenn du es erkennst die Objekteigenschaft/-funktion mit einem Sprachkommando eindeutig auszuwählen.“

- „Das Objekt/die Objektfunktion soll ja durch das Sprachkommando eindeutig vom System identifiziert werden können, d.h. darauf achten, dass Objekte auch mehrfach im Auto vorkommen können. Ist das Sprachkommando nicht eindeutig, werde ich nochmals nachfragen bis es eindeutig ist.“
- „Der zweite Teil eines Aufgabenpaares besteht aus der Ausführung einer Geste zur Manipulation des zuvor benannten Objekts oder Objekteigenschaft/-funktion. Welche Einstellung genau ausgeführt werden soll, wird dir das Tablet per Audiowiedergabe mitteilen.“

Beispielaufgabe

- CARS-Simulator dazu starten und einüben lassen.
- „Ein Beispiel ist die Auswahl des Lichts im Fahrzeuginnenraum. Durch ein Sprachkommando wie 'Innenraumlicht', 'Leselampe', 'Licht im Innenraum', 'Deckenlicht', 'Innenbeleuchtung' oder 'Leselicht' wird hier das Objekt ausgewählt. Auch möglich wäre direkt die Objekteigenschaft/-funktion zu verwenden, wie z.B. Deckenlichthelligkeit. Die genannten Sprachkommandos sind aber noch nicht ausreichend eindeutig, in diesem Falle würde ich nachfragen um welches Licht es sich handelt, denn das Licht kann ja mehrmals im Fahrzeug vorkommen.“
- „Ist dann das Objekt ausgewählt kann man die vorgegebene Aufgabe mit einer Geste ausführen. Die Aufgabenstellung kann ich bei akustischen Schwierigkeiten auch nochmals wiederholen.“
- „Die Gesten sollen auf dem Tablet (weißer Untergrund) ausgeführt werden.“
- „Es ist ein Multi-Touch-Display, d.h. mehrere Finger können eingesetzt werden.“
- „Eine wichtige Vorgabe ist, dass eine Hand stets auf dem Lenkrad verbleiben muss.“
- „Ein Fahr Simulator läuft nämlich die ganze Zeit mit und du sollst dabei die Spur halten und Hindernissen ausweichen.“
- „Beim ersten Aufgabenteil, den Sprachkommandos kannst du einfach nur gerade aus fahren, damit du dir das auf dem Tablet erscheinende Bild genau betrachten kannst.“
- „Beim zweiten Teil, die Gesten aber nur dann ausführen, wenn die Verkehrslage im Simulator es zulässt.“
- „Lautes Denken ist gewollt. Bitte immer erklären, warum du etwas wie tust.“
- „Nachdem ein Aufgabenpaar abgeschlossen wurde, werde ich dir immer dieselbe Frage stellen, die dann jeweils durch Auswahl eines Skalenwertes beantwortet werden muss. Dann geht's weiter zum nächsten Aufgabenpaar.“
- „Es besteht während der ganzen Studie keinerlei Zeitdruck.“
- „Gibt es Fragen? Bei Verständnisschwierigkeiten werde ich dir soweit möglich helfen.“

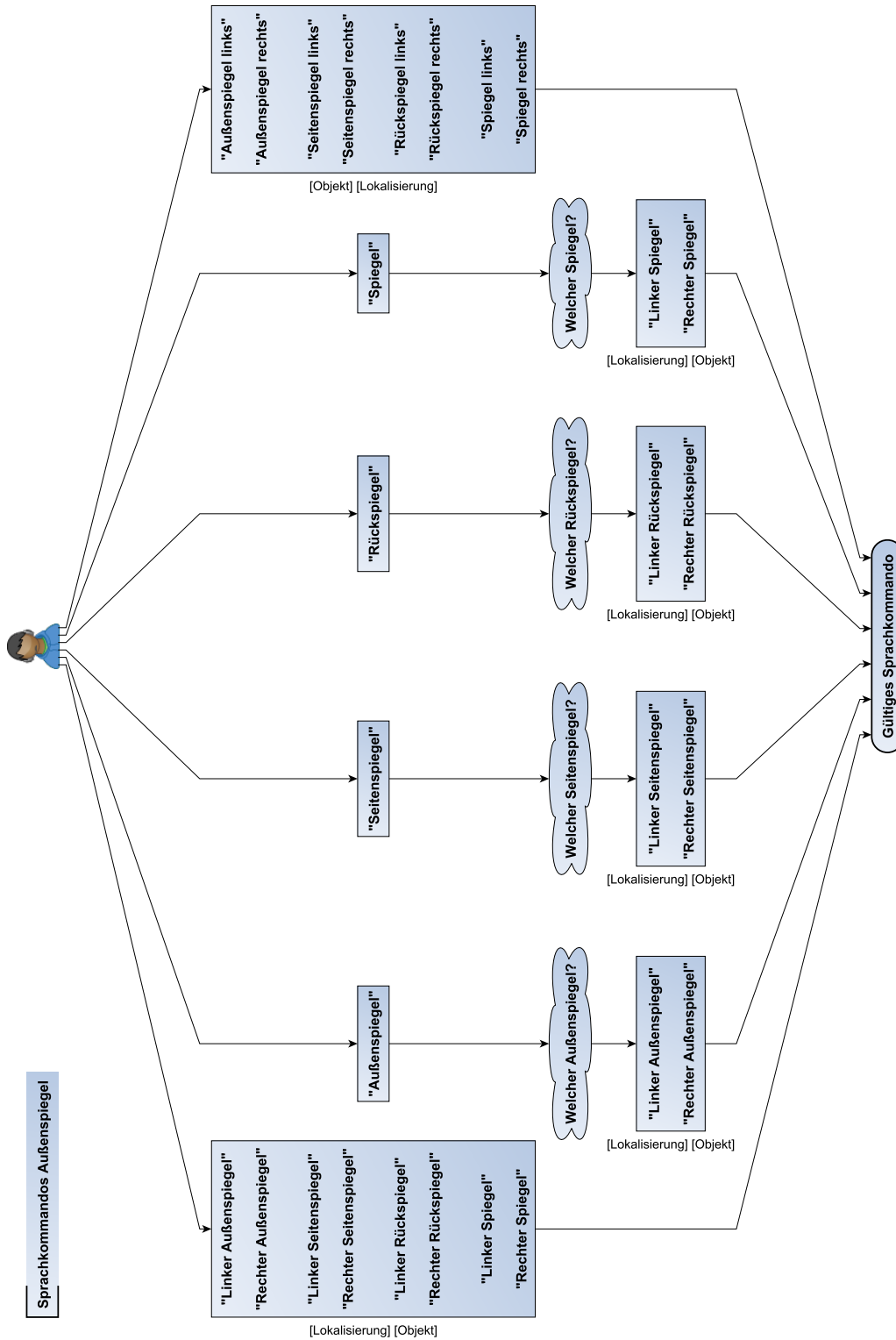
Beginn der Aufgabenreihe

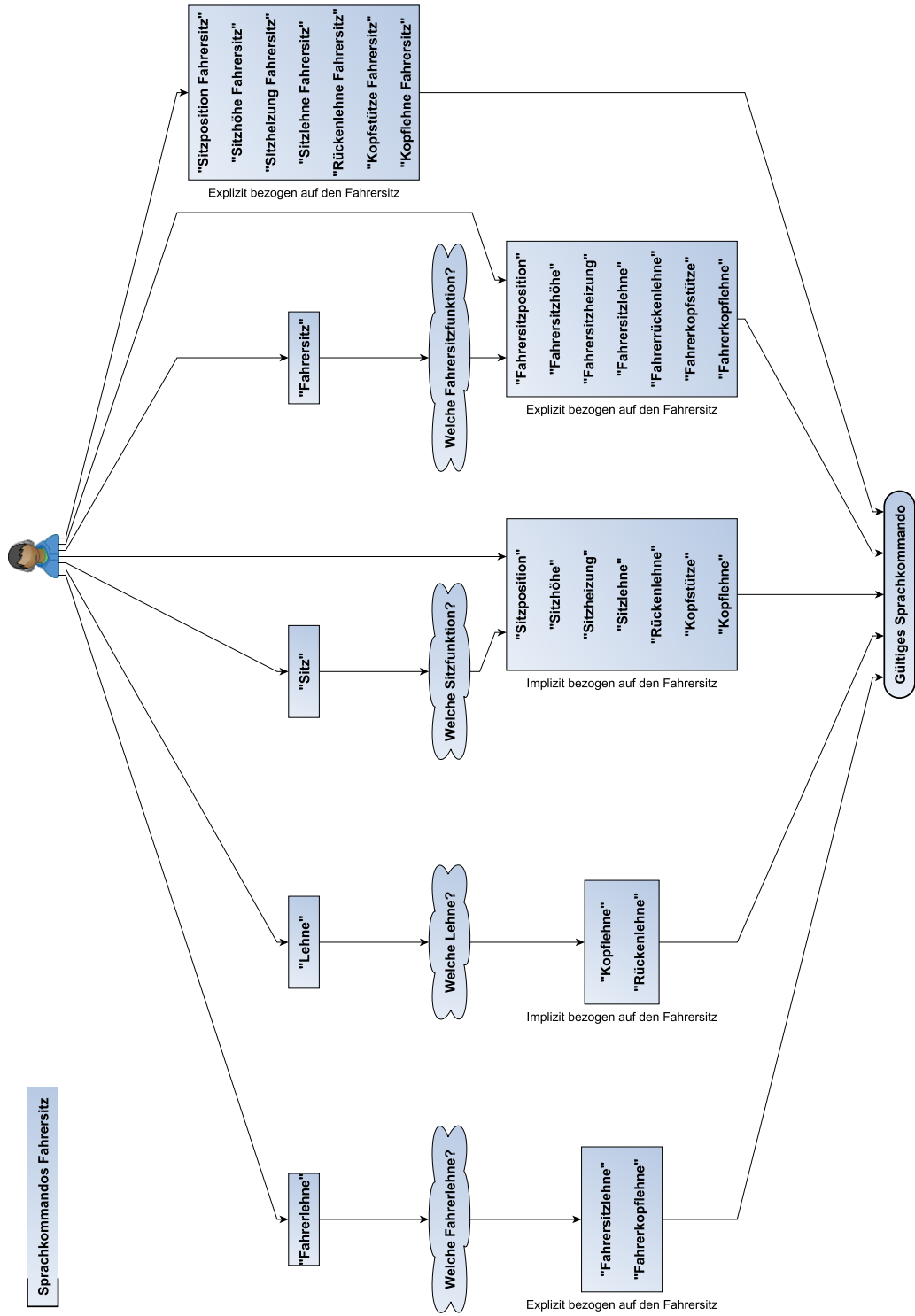
- Videoaufzeichnung starten.
- An lautes Denken wenn nötig erinnern.
- Bei nicht eindeutigen oder falschen Sprachkommandos: Nachsetzen, bis eine Eindeutigkeit erzielt wurde. Gelingt dies nach mehreren Versuchen nicht, dann weiter zur Gesteneingabe.
- Wenn Proband aus seiner Sicht einen groben Fehler während der Gesteneingabe macht, dann Eingabe zurücksetzen.

Nach Abschluss der Aufgabenreihe

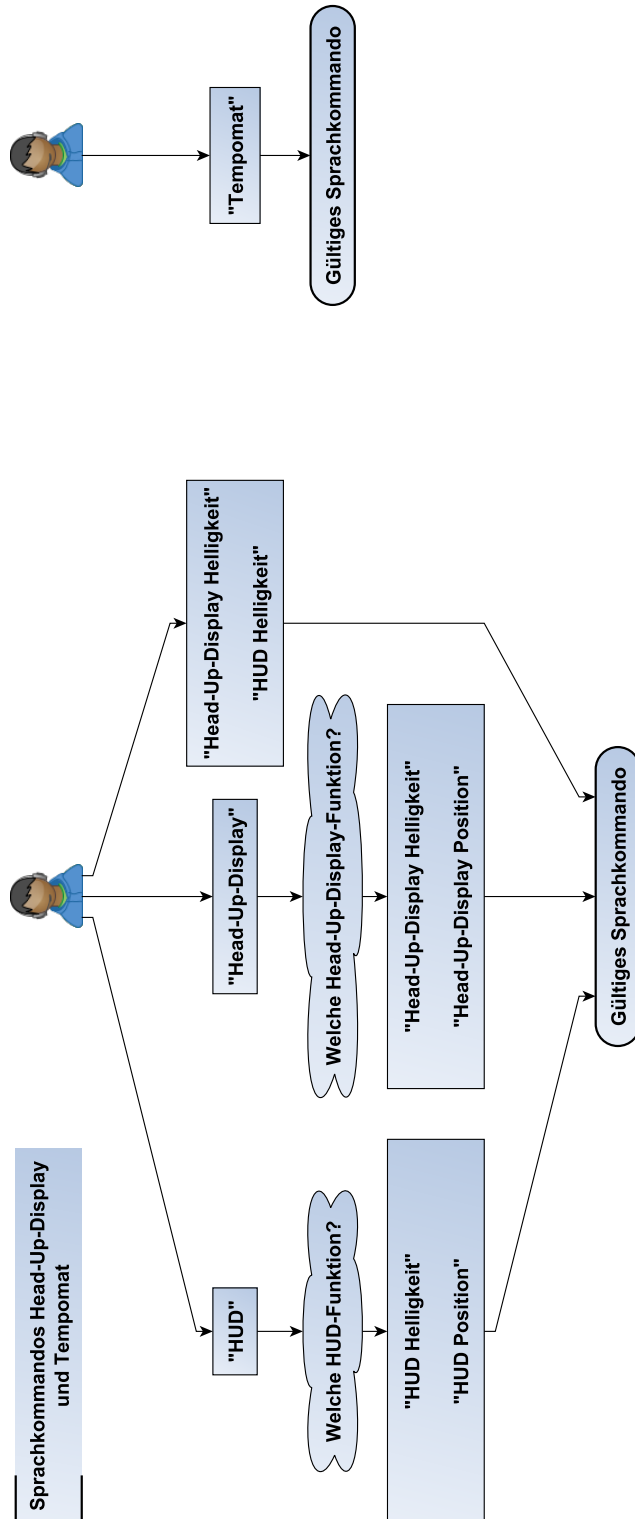
- Videoaufzeichnung stoppen.
- Abschlussfragebogen ausfüllen lassen.
- Studienvergütung ausbezahlen und Empfang quittieren lassen

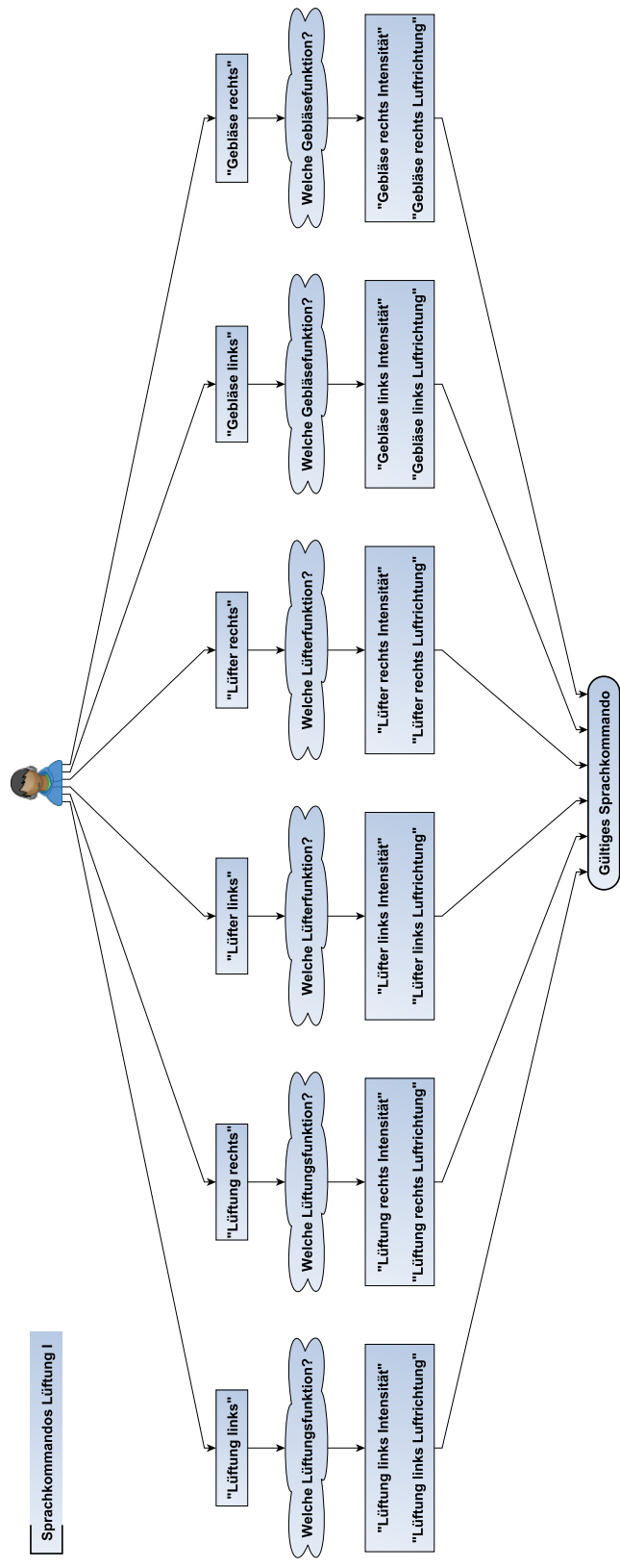
A.7. Sprachkommando-Set als Ergebnis der explorativen Benutzerstudie



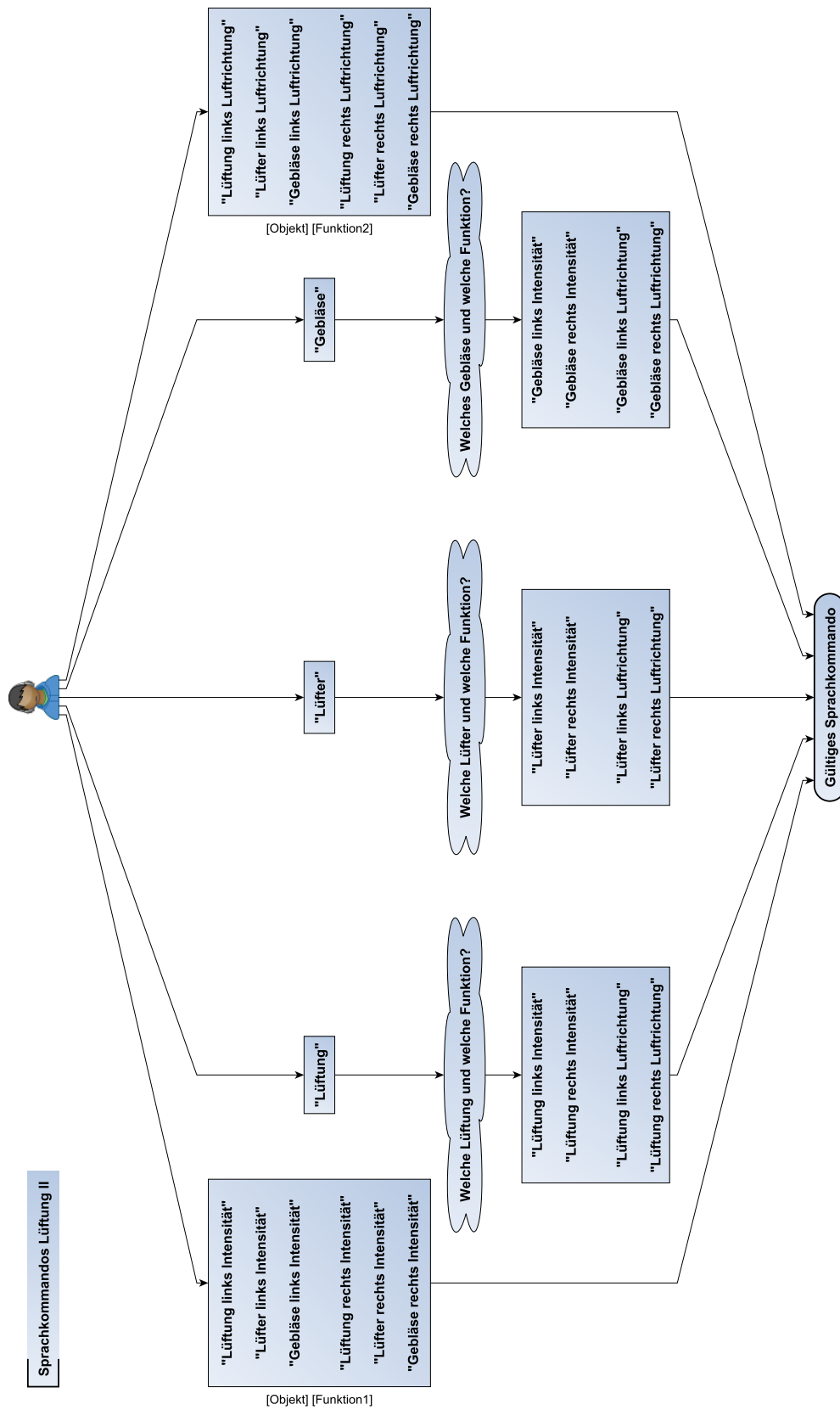


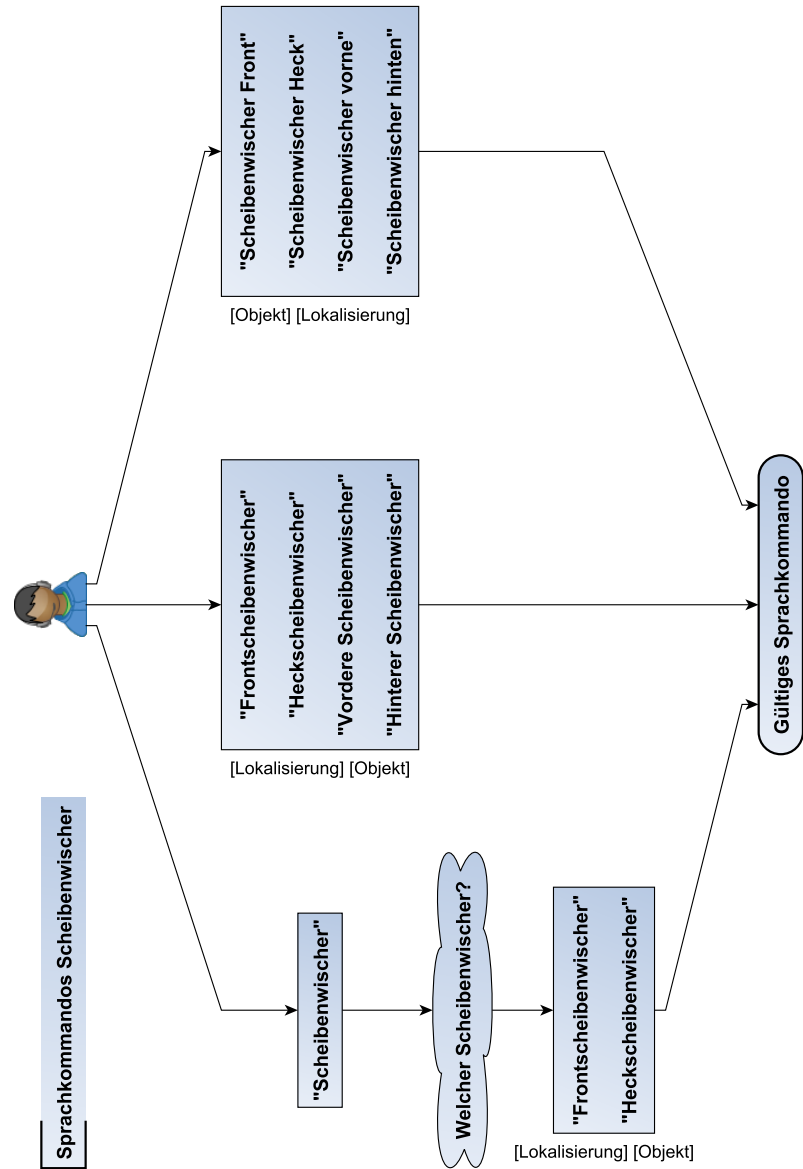
A.7. Sprachkommando-Set als Ergebnis der explorativen Benutzerstudie



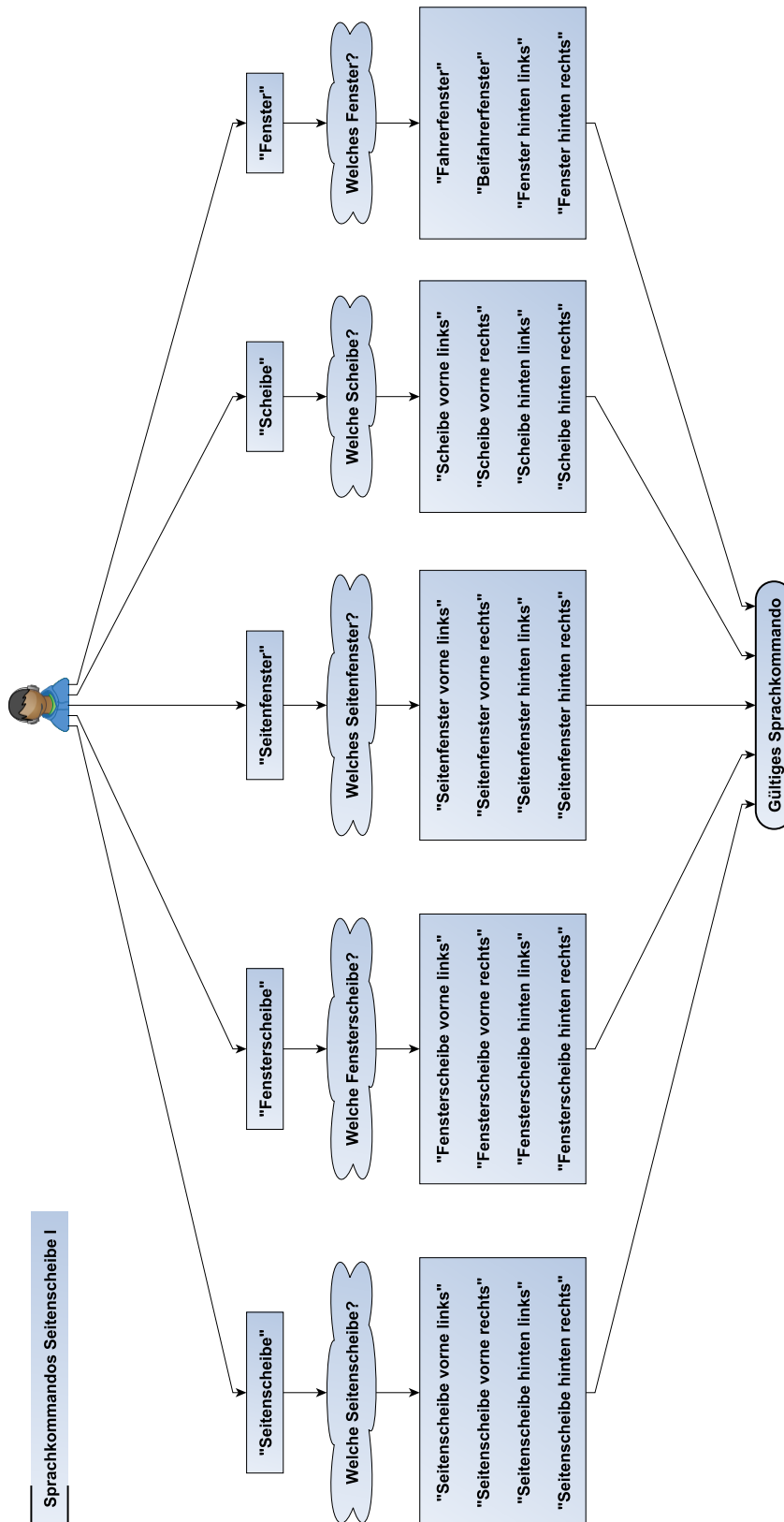


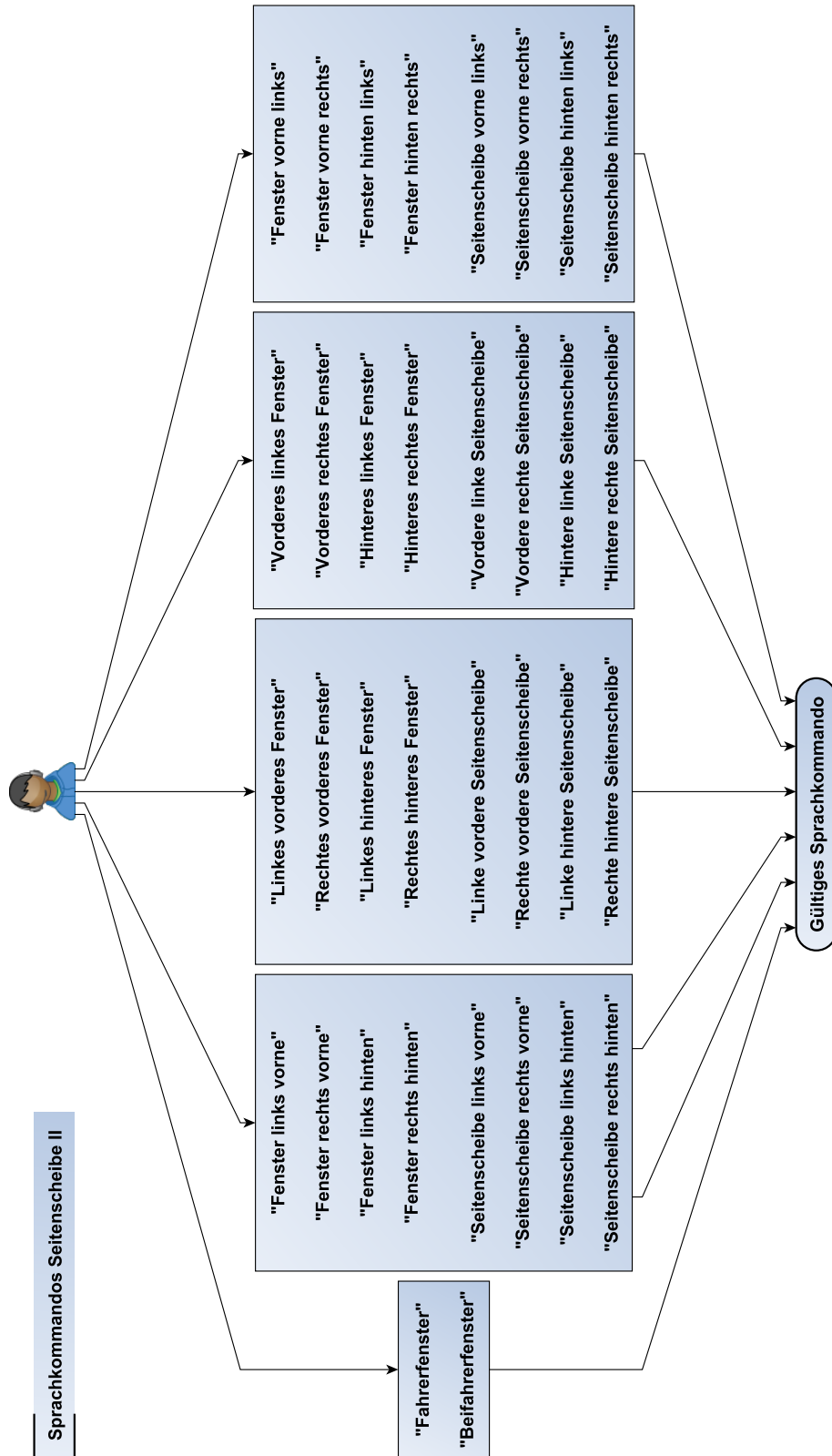
A.7. Sprachkommando-Set als Ergebnis der explorativen Benutzerstudie

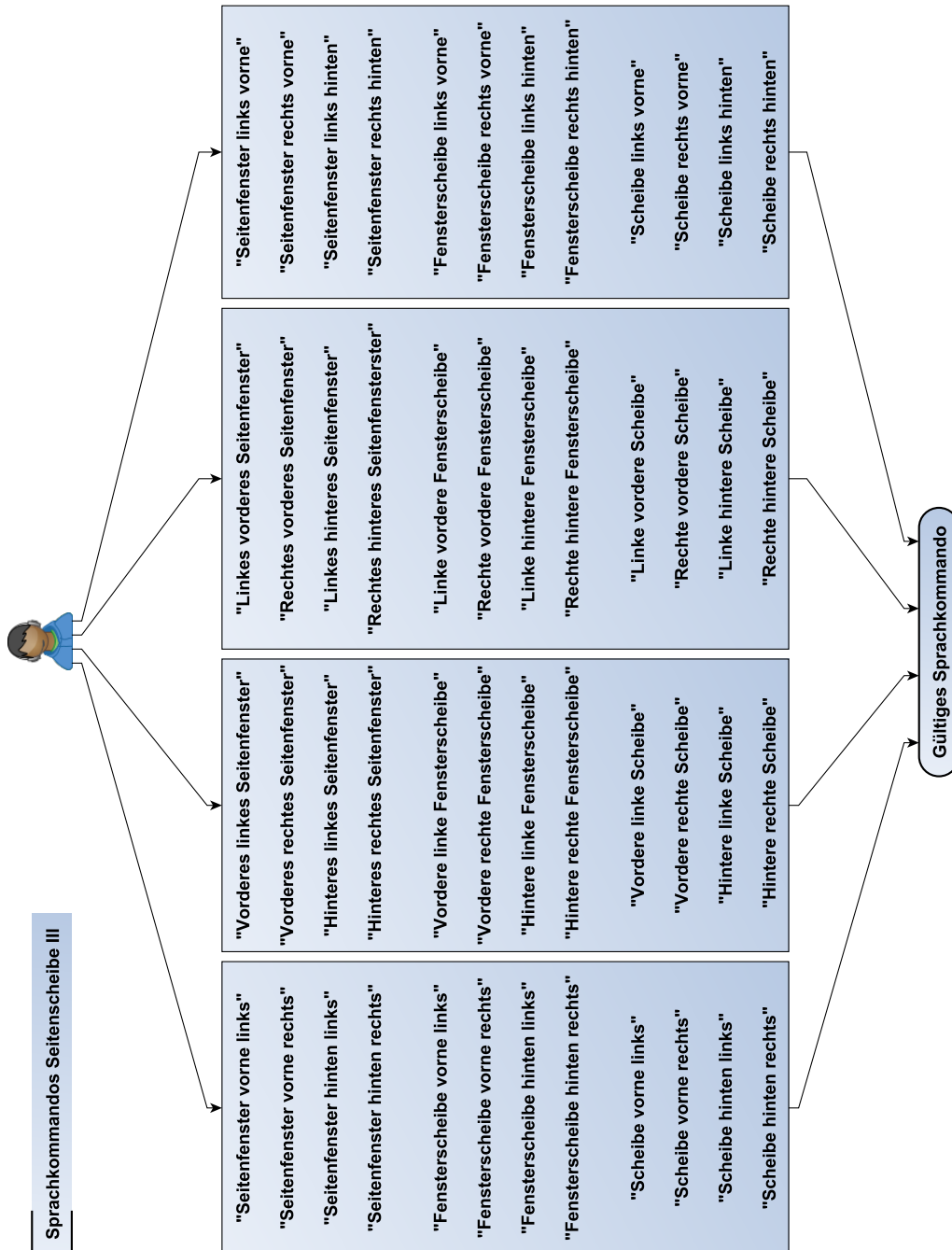




A.7. Sprachkommando-Set als Ergebnis der explorativen Benutzerstudie







A.8. Einverständniserklärung - Evaluierung

Einverständniserklärung zur Studie „Multimodale vs. Unimodale Interaktion“

Teilnehmer-Nr.: _____ Name: _____

Vorname: _____

Mir wurde der Zweck der Studie erläutert und ich wurde darüber informiert, dass ich die Studie jederzeit abbrechen kann.

Ich wurde des Weiteren darüber in Kenntnis gesetzt, dass die während der Studie gesammelten Video- und Audiodaten zur Auswertung der Studie herangezogen werden. Sämtliche Daten werden für die wissenschaftliche Nutzung gesammelt und hierbei vertraulich behandelt.

Der Verwendung von Bildmaterial mit meiner Person in Präsentationen, Ausarbeitungen und Veröffentlichungen (auch online)

stimme ich zu.

stimme ich nicht zu.

Stuttgart, den

Unterschrift:

Teil B – Erfahrungen des Probanden

B.1 Haben Sie bereits Erfahrungen mit Spracheingabe oder Sprachsteuerung sammeln können?

- ja
 nein

Falls ja, in welchen Bereichen?: _____

B.2 Haben Sie bereits Erfahrungen mit (Multi-)Touch-Displays?

- ja
 nein

Falls ja, in welchen Bereichen?: _____

A.10. Zwischenfragebogen Interaktion mit Sprache und Gesten - Evaluierung

Zwischenfragebogen zur Studie „Multimodale vs. Unimodale Interaktion“ - Interaktion mit Sprache und Gesten -

Teilnehmer-Nr.: _____ Name: _____

Vorname: _____

Teil C – Bewertung der multimodalen Eingabe

C.1 Nehmen Sie jeweils zu den nachstehenden Aussagen Stellung.

(Kreuz im zutreffenden Feld setzen)

	stimme gar nicht zu	stimme nicht zu	unent- schieden	stimme zu	stimme voll und ganz zu
1. Ich könnte mir vorstellen, Interaktion mit Sprache und Gesten im Automobil häufig einzusetzen.					
2. Die Interaktion mit Sprache und Gesten war unnötig komplex.					
3. Die Interaktion mit Sprache und Gesten war einfach durchzuführen.					
4. Ich würde Unterstützung durch das System benötigen (z.B. bei Fehlern) um die Interaktion mit Sprache und Gesten durchführen zu können.					
5. Die verschiedenen Funktionen des Systems waren gut integriert.					

A. Anhang

	stimme gar nicht zu	stimme nicht zu	unent- schieden	stimme zu	stimme voll und ganz zu
6. Die Interaktion mit Sprache und Gesten hatte zu viele Inkonsistenzen.					
7. Die meisten Autofahrer(innen) würden die Interaktion mit Sprache und Gesten schnell erlernen können.					
8. Die Interaktion mit Sprache und Gesten war umständlich durchzuführen.					
9. Ich konnte mit der Interaktion mit Sprache und Gesten sicher umgehen.					
10. Ich musste eine Menge lernen, bevor ich mit der Interaktion mit Sprache und Gesten zurechtkam.					

C.2 Geben Sie jeweils den Grad der Beanspruchung, den Sie während des Versuches empfunden haben, auf einer Skala von Niedrig(0) bis Hoch(5) an.

1.) Anforderung an die gesamte Aufmerksamkeit

Aufmerksamkeit, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Insgesamt alle mentalen (denken, entscheiden...), visuellen, auditiven und manuellen Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen.)

Niedrig Hoch

2.) Visuelle Anforderung

Visuelle Anforderung, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Visuelle Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen (alles, was mit Sehen zu tun hat).)

Niedrig Hoch

3.) Auditive Anforderung

Auditive Anforderung, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Auditive Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen (alles, was mit Hören zu tun hat).)

Niedrig Hoch

4.) Manuelle Anforderung

Manuelle Anforderung, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Manuelle Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen (alles, was mit der Handhabung zu tun hat).)

Niedrig Hoch

5.) Stress

Stressempfinden während der Fahrt, wie Unsicherheit, Irritation, Entmutigung, usw..

(Stressniveau während der Fahrt wie Irritation, Müdigkeit, Unsicherheit, Entmutigung, etc.)

Niedrig *Hoch*

6.) Zeitliche Anforderung

Zeitlicher Druck während der Fahrt.

(Gefühlte Belastung und spezifische Beeinträchtigung durch die schnelle Abfolge der Aufgaben.)

Niedrig *Hoch*

7.) Interferenz

Störung durch die Zweitaufgabe beim Fahren und deren Auswirkungen auf das Fahren.

(Beeinträchtigung des Fahrerzustandes und Auswirkungen auf die Fahrleistung durch gleichzeitige Zweitaufgabe während des Fahrens.)

Niedrig *Hoch*

C.3 Hat ihnen die Interaktion mit Sprache und Gesten gefallen?

[] ja

[] nein

A.11. Zwischenfragebogen Interaktion mit Buttons und Knöpfen - Evaluierung

Zwischenfragebogen zur Studie „Multimodale vs. Unimodale Interaktion“ - Interaktion mit Buttons und Knöpfen -

Teilnehmer-Nr.: _____ Name: _____

Vorname: _____

Teil D – Bewertung der multimodalen Eingabe

D.1 Nehmen Sie jeweils zu den nachstehenden Aussagen Stellung.

(Kreuz im zutreffenden Feld setzen)

	stimme gar nicht zu	stimme nicht zu	unent- schieden	stimme zu	stimme voll und ganz zu
1. Ich könnte mir vorstellen, Interaktion mit Buttons und Knöpfen im Automobil häufig einzusetzen.					
2. Die Interaktion mit Buttons und Knöpfen war unnötig komplex.					
3. Die Interaktion mit Buttons und Knöpfen war einfach durchzuführen.					
4. Ich würde Unterstützung durch das System benötigen (z.B. bei Fehlern) um die Interaktion mit Buttons und Knöpfen durchführen zu können.					
5. Die verschiedenen Funktionen des Systems waren gut integriert.					

A. Anhang

	stimme gar nicht zu	stimme nicht zu	unent- schieden	stimme zu	stimme voll und ganz zu
6. Die Interaktion mit Buttons und Knöpfen hatte zu viele Inkonsistenzen.					
7. Die meisten Autofahrer(innen) würden die Interaktion mit Buttons und Knöpfen schnell erlernen können.					
8. Die Interaktion mit Buttons und Knöpfen war umständlich durchzuführen.					
9. Ich konnte mit der Interaktion mit Buttons und Knöpfen sicher umgehen.					
10. Ich musste eine Menge lernen, bevor ich mit der Interaktion mit Buttons und Knöpfen zurechtkam.					

D.2 Geben Sie jeweils den Grad der Beanspruchung, den Sie während des Versuches empfunden haben, auf einer Skala von Niedrig(0) bis Hoch(5) an.

1.) Anforderung an die gesamte Aufmerksamkeit

Aufmerksamkeit, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Insgesamt alle mentalen (denken, entscheiden...), visuellen, auditiven und manuellen Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen.)

Niedrig Hoch

2.) Visuelle Anforderung

Visuelle Anforderung, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Visuelle Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen (alles, was mit Sehen zu tun hat).)

Niedrig Hoch

3.) Auditive Anforderung

Auditive Anforderung, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Auditive Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen (alles, was mit Hören zu tun hat).)

Niedrig Hoch

4.) Manuelle Anforderung

Manuelle Anforderung, die während der Fahrt nötig war, um die Aufgabe zu erfüllen.

(Manuelle Faktoren, die während der Fahrt erforderlich waren, um die Gesamtleistung zu erzielen (alles, was mit der Handhabung zu tun hat).)

Niedrig Hoch

5.) Stress

Stressempfinden während der Fahrt, wie Unsicherheit, Irritation, Entmutigung, usw..

(Stressniveau während der Fahrt wie Irritation, Müdigkeit, Unsicherheit, Entmutigung, etc.)

Niedrig *Hoch*

6.) Zeitliche Anforderung

Zeitlicher Druck während der Fahrt.

(Gefühlte Belastung und spezifische Beeinträchtigung durch die schnelle Abfolge der Aufgaben.)

Niedrig *Hoch*

7.) Interferenz

Störung durch die Zweitaufgabe beim Fahren und deren Auswirkungen auf das Fahren.

(Beeinträchtigung des Fahrzustandes und Auswirkungen auf die Fahrleistung durch gleichzeitige Zweitaufgabe während des Fahrens.)

Niedrig *Hoch*

D.3 Hat ihnen die Interaktion mit Buttons und Knöpfen gefallen?

[] ja

[] nein

A.12. Abschlussfragebogen - Evaluierung

Abschlussfragebogen zur Studie „Multimodale vs. Unimodale Interaktion“

Teilnehmer-Nr.: _____ Name: _____

Vorname: _____

Teil E – Vergleich der Modalitäten

E.1 Nehmen Sie jeweils zu den nachstehenden Aufgaben im Automobil Stellung bezüglich der Aussage:

„Ich halte eine Bedienung über die Kombination Sprachkommando und Touch-Geste für diese Aufgabe für weniger ablenkend als mit herkömmlichen Bedienungselementen.“

(Kreuz im zutreffenden Feld setzen)

	stimme voll und ganz zu	stimme zu	unentschieden	stimme nicht zu	stimme gar nicht zu
Außenspiegel einstellen					
Scheibenwischer bedienen					
Seitenscheiben öffnen/schließen					
Tempomat einstellen					

E.2 Können Sie sich vorstellen, die Interaktion mit Sprache und Gesten bevorzugt gegenüber der herkömmlichen Interaktion mit Buttons und Knöpfen einzusetzen?

[] ja

[] nein

A. Anhang

E.3 Welche zusätzlichen Funktionen, die nicht Teil der Studie waren, würden Sie sich bezüglich einer Bedienung mit der Kombination Sprachkommandos und Touch-Gesten noch wünschen?:

Teil F – Beurteilung der Interaktionspotentiale

F.1 Welche Vorteile sehen Sie bei der Interaktion mit Sprachkommandos und Touch-Gesten im Vergleich mit der Interaktion mit Buttons und Knöpfen?

F.2 Welche Nachteile sehen Sie bei der Interaktion mit Sprachkommandos und Touch-Gesten im Vergleich mit der Interaktion mit Buttons und Knöpfen?

Vielen Dank für Ihre Teilnahme!

A.13. Quittung - Evaluierung

Studie „Multimodale vs. Unimodale Interaktion“

Bestätigung über den Erhalt der Studienvergütung

Teilnehmer-Nr.: _____

Name: _____

Vorname: _____

Die Vergütung für die Studie in Höhe von 10,- Euro habe ich erhalten.

Stuttgart, den

Unterschrift:

A.14. Leitfaden - Evaluierung

Leitfaden – Proband []

Einführung in die Studie

- „Die Studie ist Teil meiner Diplomarbeit, in der eine neue Form der Interaktion im Automobil untersucht werden soll: Nämlich eine Kombination von Sprachkommandos und Touch-Gesten.“
- „Mit Sprachkommandos sollen Objekte oder Objektfunktionen eindeutig ausgewählt werden. In einem weiteren Schritt wird mittels Touch-Gesten dann eine Einstellung auf den ausgewählten Objekten oder Objektfunktionen durchgeführt. Eine Interaktion besteht somit aus zwei Teilen.“
- „In der Studie soll nun mit einem Fahrsimulator und einem erstellten Prototypen diese Interaktion mit Testkandidaten untersucht und mit der manuellen Interaktion über Buttons und Knöpfe verglichen werden.“
- „(Während der Studie werden Sprach- und Videoaufzeichnungen gemacht, welche zur Auswertung der Studie herangezogen werden.) Die gesammelten Daten werden vertraulich behandelt.“
- „Die Studie kann selbstverständlich jederzeit abgebrochen werden.“
- „Die Studie beginnt mit einem Eingangsfragebogen und einer Einverständniserklärung, dass die gesammelten Daten verwendet werden dürfen.“

Eingangsfragebogen und Einverständniserklärung an den Probanden aushändigen und ausfüllen lassen

- Einverständniserklärung selber vorlesen

Vorbereitung

- Sitz richtig einstellen
- Fenster (d.h. Bildschirme) positionieren
- Pedale positionieren
- LCT-Beispielbild öffnen

Einweisung in den Lane Change Test

- Erklärung anhand des Screenshots
- „Eine Fahrt beginnt vor einer Kurve, die passiert werden muss. Nach dem Kurvenausgang muss das Gaspedal konstant durchgedrückt und die Maximalgeschwindigkeit gehalten werden.“
- „Ab dem gelben Startschild beginnen jeweils die Interaktionsaufgaben.“
- „Die Aufgaben werden dann von mir in zufälliger Reihenfolge hintereinander gestellt und müssen direkt nach Ende der verbalen Anweisung durchgeführt werden, aber nicht bereits während ich noch spreche!“
- „Das Hauptaugenmerk liegt allerdings darauf, während der Fahrt mit dem Fahrsimulator die Schilder an den Straßenrändern genau zu beachten. Diese Schilder zeigen an, auf welche der drei Spuren ein Fahrspurwechsel vorgenommen werden muss. Dieser Fahrspurwechsel muss nachdem er erkannt wurde auch direkt durchgeführt werden. Dabei aber nicht einfach scharf einlenken oder schlittern, sondern in normaler Weise die Spur wechseln.“
- „Eine Strecke endet, sobald die nächste Kurve (nach ca. drei Minuten) erreicht wird.“

Übungsfahrt mit dem LCT

- LCT starten (Strecke-Nr. [])
- „Du hast nun Gelegenheit dich mit dem Fahrsimulator vertraut zu machen. Dazu machst du nun eine Testfahrt ohne Interaktion.“
- „Nochmals: Beachte und reagiere dabei auf die Schilder am Straßenrand!“
- Testfahrt ohne Interaktion durchführen und die LCT-Regeln auf Einhaltung überwachen
- LCT beenden

Detailliertere Erläuterung der Vorgehensweise

- „Mehrere Fahrten mit dem Fahrsimulator stehen jetzt an. Eine erste Fahrt ohne jegliche Interaktion, eine Fahrt mit „Interaktion mit Sprache und Gesten“, eine Fahrt mit „Interaktion mit Buttons und Knöpfen“ und abschließend nochmals eine Fahrt ohne jede Interaktion. Nach den Interaktionsfahrten sind jeweils Fragebögen auszufüllen.“

Durchführung der Aufgabenreihen

Baseline 1

- LCT starten (Strecke-Nr. [])
- Hinweis: „Fahraufgabe ist die zentrale Aufgabe.“
- Baseline-Fahrt durchführen und LCT-Regeln auf Einhaltung überwachen.
- LCT beenden
- LCT-File umbenennen

Condition 1 []

- Kondition 1 erklären und einüben lassen (siehe X1 oder X2)
- Richtige Tablet-App starten
- LCT starten (Strecke-Nr. [])
- Hinweis: „Fahraufgabe ist die zentrale Aufgabe.“
- Fahrt starten und Aufgaben dabei ansagen und durchführen lassen. LCT-Regeln auf Einhaltung überwachen.
- LCT beenden
- Zwischenfragebogen ausfüllen lassen
- LCT-File umbenennen
- Prototyp in Ausgangszustand zurücksetzen (siehe Aufgabenbogen)

Condition 2 []

- Kondition 2 erklären und einüben lassen (siehe X1 oder X2)
- Richtige Tablet-App starten
- LCT starten (Strecke-Nr. [])
- Hinweis: „Fahraufgabe ist die zentrale Aufgabe.“
- Fahrt starten und Aufgaben dabei ansagen und durchführen lassen. LCT-Regeln auf Einhaltung überwachen.
- LCT beenden
- Zwischenfragebogen ausfüllen lassen
- LCT-File umbenennen
- Prototyp in Ausgangszustand zurücksetzen (siehe Aufgabenbogen)

Baseline 2

- LCT starten (Strecke-Nr. [])
- Hinweis: „Fahraufgabe ist die zentrale Aufgabe.“
- Baseline-Fahrt durchführen und LCT-Regeln auf Einhaltung überwachen.
- LCT beenden
- LCT-File umbenennen

Nach Abschluss der Aufgabenreihe

- Abschlussfragebogen ausfüllen lassen
- Studienvergütung ausbezahlen und Empfang (2x) quittieren lassen
- LCT-Aufzeichnungen umbenennen falls noch nicht vollständig geschehen und ins SVN sichern

X1: Einweisung in die multimodale Interaktion

- „Ich gebe dir nun einen Überblick über die Bedienung der einzelnen Funktionen. Die Objekte und deren Funktionen können mit Sprachkommandos ausgewählt und dann mit Gesten manipuliert werden.“
- Überblick über die Sprachkommandos und Gesten zu den einzelnen Objekten und Funktionen der Evaluierung geben und die Bedienung vorführen (sowohl mit Sprachaktivierung als auch PTT)
- Einüben lassen in den Gebrauch der einzelnen Kommandos und Gesten
- Prototyp in Ausgangszustand zurücksetzen (siehe Aufgabenbogen)

X2: Einweisung in die unimodale Interaktion

- „Ich gebe dir nun einen Überblick über die Bedienung der einzelnen Funktionen durch den Einsatz von Buttons und Knöpfen. Die Funktionen der Objekte können damit direkt durchgeführt werden.“
- Überblick über die Buttons und Knöpfe zu den einzelnen Objekten und Funktionen der Evaluierung geben und die Bedienung vorführen
- Einüben lassen in den Gebrauch der einzelnen Buttons und Knöpfe
- Prototyp in Ausgangszustand zurücksetzen (siehe Aufgabenbogen)

Literaturverzeichnis

- [1] ADOLPH, M.: Decreasing Driver Distraction / ITU Telecommunication Standardization Bureau. 2010. – ITU-T Technology Watch Report (Zitiert auf Seite 16)
- [2] ALTHOFF, F. ; LINDL, R. ; WALCHSHÄUSL, L.: Robust Multimodal Hand- and Head Gesture Recognition for Controlling Automotive Infotainment Systems. In: *VDI-Tagung: Der Fahrer im 21. Jahrhundert*. Braunschweig, Deutschland, November 2005 (Zitiert auf den Seiten 11 und 29)
- [3] ALTHOFF, F. ; MCGLAUN, G. ; LANG, M. K. ; RIGOLL, G.: Evaluating multimodal interaction patterns in various application scenarios. In: *Proceedings of the 5th International Gesture Workshop (GW 2003)*. Genua, Italien, 2003, S. 421–435 (Zitiert auf Seite 11)
- [4] ANGUITA, J. ; HERNANDO, J. ; PEILLON, S. ; BRAMOULLE, A.: Detection of Confusable Words in Automatic Speech Recognition. In: *IEEE Signal Processing Letters* 12 (2005), August, Nr. 8, S. 585 – 588 (Zitiert auf Seite 34)
- [5] BACH, K. M. ; JAEGER, M. G. ; SKOV, M. B. ; THOMASSEN, N. G.: You Can Touch, but You Can't Look: Interacting with In-Vehicle Systems. In: *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems (CHI '08)*. New York, NY, USA : ACM, 2008, S. 1139–1148 (Zitiert auf Seite 29)
- [6] BARÓN, A. ; GREEN, P.: Safety and Usability of Speech Interfaces for In-Vehicle Tasks while Driving: A Brief Literature Review / The University of Michigan Transportation Research Institute (UMTRI). 2006. – Forschungsbericht (Zitiert auf den Seiten 27 und 33)
- [7] BENCINA, R. ; KALTENBRUNNER, M. ; JORDA, S.: Improved Topological Fiducial Tracking in the reacTIVision System. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops - Volume 03*. Washington, DC, USA : IEEE Computer Society, 2005, S. 99–106 (Zitiert auf Seite 65)
- [8] BERNSEN, N. O. ; DYBKJAER, L.: *Multimodal Usability*. 2. London : Springer-Verlag, 2009. – S. 72 (Zitiert auf Seite 31)
- [9] BIBLIOGRAPHISCHES INSTITUT GMBH, Dudenverlag: *Definition "Geste"*. <http://www.duden.de/rechtschreibung/Geste>. – letzter Aufruf am 13.11.2011 (Zitiert auf Seite 29)

- [10] BMW GROUP: *PressClub*. <http://www.press.bmwgroup.com>. – letzter Aufruf am 18.11.2011 (Zitiert auf den Seiten 18, 25, 26 und 29)
- [11] BOLT, R. A.: Put-That-There: Voice and Gesture at the Graphics Interface. In: *Proceedings of the 7th annual conference on Computer graphics and interactive techniques (SIGGRAPH '80)*. New York, NY, USA : ACM, 1980, S. 262–270 (Zitiert auf Seite 32)
- [12] BROOKE, J.: SUS - A quick and dirty usability scale. In: JORDAN, P. W. (Hrsg.) ; THOMAS, B. (Hrsg.) ; WEERDMEESTER, B. A. (Hrsg.) ; McCLELLAND, A. L. (Hrsg.): *Usability Evaluation in Industry*. 1. London : Taylor and Francis Ltd, Juni 1996, Kapitel 21, S. 189–194 (Zitiert auf den Seiten 79, 82 und 92)
- [13] BUXTON, W.: Taxonomies of Input. Version: Januar 2009. <http://www.billbuxton.com/input04.Taxonomies.pdf>. In: *Human Input to Computer Systems: Theories, Techniques and Technology*. William Buxton <http://www.billbuxton.com>, Januar 2009, Kapitel 4, 4.1-4.16. – letzter Aufruf am 12.11.2011 (Zitiert auf Seite 24)
- [14] CHANG, J. ; LIEN, A. ; LATHROP, B. ; HEES, H.: Usability Evaluation of a Volkswagen Group In-Vehicle Speech System. In: *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '09)*. New York, NY, USA : ACM, 2009, S. 137–144 (Zitiert auf den Seiten 11, 28 und 34)
- [15] CITROEN: *Lane Departure Warning System - LDWS*. <http://www.piecescitroensport.citroen.com/CWW/en-US/TECHNOLOGIES/SECURITY/AFIL/>. – letzter Aufruf am 13.11.2011 (Zitiert auf Seite 27)
- [16] COMMISSION OF THE EUROPEAN COMMUNITIES: Commission Recommendation of 22 December 2006 on safe and efficient in-vehicle information and communication systems: Update of the European Statement of Principles on human machine interface. In: *Official Journal of the European Union* 50 (2007), Februar, Nr. L32, S. 200–241 (Zitiert auf den Seiten 15, 16 und 26)
- [17] DIN DEUTSCHES INSTITUT FÜR NORMUNG: *Road vehicles - Ergonomic aspects of transport information and control systems - Dialogue management principles and compliance procedures*. Oktober 2003. – Ref.No.: DIN EN ISO 15005:2003 (Zitiert auf Seite 15)
- [18] DIN DEUTSCHES INSTITUT FÜR NORMUNG: *Ergonomics of human-system interaction - Part 110: Dialogue principles*. September 2008. – Ref.No.: DIN EN ISO 9241-110:2008 (Zitiert auf Seite 35)
- [19] DIN DEUTSCHES INSTITUT FÜR NORMUNG: *Road vehicles - Ergonomic aspects of transport information and control systems - Specifications and test procedures for in-vehicle visual presentation*. Juni 2011. – Ref.No.: DIN EN ISO 15008:2011 (Zitiert auf Seite 15)

-
- [20] DÖRING, T. ; KERN, D. ; MARSHALL, P. ; PFEIFFER, M. ; SCHÖNING, J. ; GRUHN, V. ; SCHMIDT, A.: Gestural Interaction on the Steering Wheel: Reducing the Visual Demand. In: *Proceedings of the 2011 annual conference on Human factors in computing systems (CHI '11)*. New York, NY, USA : ACM, 2011, S. 483–492 (Zitiert auf den Seiten 30 und 34)
- [21] ECKER, R. ; BROY, V. ; HERTZSCHUCH, K. ; BUTZ, A.: Visual Cues supporting Direct Touch Gesture Interaction with In-Vehicle Information Systems. In: *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '10)*. New York, NY, USA : ACM, 2010, S. 80–87 (Zitiert auf den Seiten 11 und 29)
- [22] ENDRES, C. ; SCHWARTZ, T. ; MÜLLER, C. A.: Geremin: 2D Microgestures for Drivers Based on Electric Field Sensing. In: *Proceedings of the 16th international conference on Intelligent user interfaces (IUI '11)*. New York, NY, USA : ACM, 2011, S. 327–330 (Zitiert auf Seite 32)
- [23] GEISER, G.: Man Machine Interaction in Vehicles. In: *ATZ* 87 (1985), S. 74–77 (Zitiert auf Seite 22)
- [24] GELAU, C. ; HENNING, M.J. ; KREMS, J.F.: On the reliability of the occlusion technique as a tool for the assessment of the HMI of in-vehicle information and communication systems. In: *Applied Ergonomics* 40 (2009), Nr. 2, S. 181–184 (Zitiert auf Seite 18)
- [25] GONZÁLEZ, I. ; WOBROCK, J. ; CHAU, D. ; FAULRING, A. ; MYERS, B.: Eyes on the Road, Hands on the Wheel: Thumb-based Interaction Techniques for Input on Steering Wheels. In: *Proceedings of Graphics Interface 2007 (GI '07)*. New York, NY, USA : ACM, 2007, S. 95–102 (Zitiert auf den Seiten 11 und 30)
- [26] GOOGLE CODE: *TUIOdroid - Open Source TUIO App for Android*. <http://code.google.com/p/tuiodroid>. – letzter Aufruf am 14.11.2011 (Zitiert auf Seite 65)
- [27] GRAHAM, R. ; CARTER, C.: Comparison of Speech Input and Manual Control of In-Car Devices while on the Move. In: *Personal and Ubiquitous Computing* 4 (2000), Nr. 2/3, S. 155–164 (Zitiert auf den Seiten 33 und 34)
- [28] GÄRTNER, U. ; KÖNIG, W. ; WITTIG, T.: Evaluation of Manual vs. Speech Input When Using a Driver Information System in Real Traffic. In: *Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*. Iowa City, USA : University of Iowa, 2001, S. 7–13 (Zitiert auf Seite 27)
- [29] GRUENSTEIN, A. ; ORSZULAK, J. ; LIU, S. ; ROBERTS, S. ; ZABEL, J. ; REIMER, B. ; MEHLER, B. ; SENEFF, S. ; GLASS, J. ; COUGHLIN, J.: City Browser: Developing a Conversational Automotive HMI. In: *Proceedings of the 27th international conference extended abstracts on*

- Human factors in computing systems (CHI EA '09)*. New York, NY, USA : ACM, 2009, S. 4291–4296 (Zitiert auf Seite 34)
- [30] HAMMER, M.R.F. ; DOUGLAS, F.C. ; TOBIN, D.: Technology and driver distraction—the need for industry guidelines. In: FAULKS, I.J. (Hrsg.) ; REGAN, M. (Hrsg.) ; STEVENSON, M. (Hrsg.) ; BROWN, J. (Hrsg.) ; PORTER, A. (Hrsg.) ; IRWIN, J.D. (Hrsg.): *Distracted driving*. Sydney, NSW : Australasian College of Road Safety, 2007, S. 525–576 (Zitiert auf den Seiten 16 und 25)
- [31] HASSEL, L. ; HAGEN, E.: Evaluation of a Dialogue System in an Automotive Environment. In: *Proceedings of the 6th SIGdial Workshop on Discourse and Dialogue*. Lissabon, Portugal, September 2005, S. 155–165 (Zitiert auf Seite 28)
- [32] HATAOKA, N. ; ARAKI, M. ; MATSUDA, T. ; TAKAHASHI, M. ; OHTAKI, R. ; OBUCHI, Y.: Evaluation of Interface and In-Car Speech - Many Undesirable Utterances and Sever Noisy Speech on Car Navigation Application -. In: *Proceedings of IEEE 10th Workshop on Multimedia Signal Processing*. Cairns, Australien, Oktober 2008, S. 956 –959 (Zitiert auf den Seiten 28 und 34)
- [33] HOLLEIS, P.: *EIToolkit*. <http://www.eitoolkit.de>. – letzter Aufruf am 14.11.2011 (Zitiert auf den Seiten 60 und 61)
- [34] HUA, Z. ; NG, W. L.: Speech Recognition Interface Design for In-Vehicle System. In: *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '10)*. New York, NY, USA : ACM, 2010, S. 29–33 (Zitiert auf den Seiten 11, 28 und 34)
- [35] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION: *Road vehicles - Ergonomic aspects of transport information and control systems - Occlusion method to assess visual distraction due to the use of in-vehicle systems*. April 2007. – Ref.No.: ISO 16673:2007 (Zitiert auf Seite 18)
- [36] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION: *Road vehicles - Ergonomic aspects of transport information and control systems - Simulated lane change test to assess in-vehicle secondary task demand*. September 2010. – Ref.No.: ISO 26022:2010 (Zitiert auf Seite 19)
- [37] KALTENBRUNNER, M. ; BOVERMANN, T. ; BENCINA, R. ; COSTANZA, E.: TUIO - A Protocol for Table-Top Tangible User Interfaces. In: *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW 2005)*. Vannes, Frankreich, 2005 (Zitiert auf Seite 65)
- [38] KERN, D. ; MAHR, A. ; CASTRONOVO, S. ; SCHMIDT, A. ; MÜLLER, C.: Making Use of Drivers' Glances onto the Screen for Explicit Gaze-Based Interaction. In: *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2010)*. New York, NY, USA : ACM, 2010 (Zitiert auf Seite 11)

- [39] KERN, D. ; MÜLLER, M. ; SCHNEEGASS, S. ; WOLEJKO-WOLEJSZO, L. ; SCHMIDT, A.: CARS – Configurable Automotive Research Simulator. In: *Automotive User Interfaces and Interactive Applications (AUIIA 08)*. Lübeck : Workshop at Mensch und Computer 2008, 2008 (Zitiert auf den Seiten 21, 22, 38 und 66)
- [40] KERN, D. ; SCHMIDT, A.: Design Space for Driver-based Automotive User Interfaces. In: *Proceedings of the First International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI 2009)*. New York, NY, USA : ACM, 2009, S. 3–10 (Zitiert auf den Seiten 24, 25, 27 und 33)
- [41] KERN, D. ; SCHMIDT, A. ; ARNSMANN, J. ; APPELMANN, T. ; PARARASASEGARAN, N. ; PIEPIERA, B.: Writing to Your Car: Handwritten Text Input While Driving. In: *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems (CHI EA '09)*. New York, NY, USA : ACM, 2009, S. 4705–4710 (Zitiert auf den Seiten 30 und 31)
- [42] KLAUER, S. G. ; DINGUS, T. A. ; NEALE, V. L. ; SUDWEEKS, J. D. ; RAMSEY, D. J.: The Impact of Driver Inattention On Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data / U.S. National Highway Traffic Safety Administration. 2006. – Forschungsbericht (Zitiert auf den Seiten 16 und 17)
- [43] KRAY, C. ; NESBITT, D. ; DAWSON, J. ; ROHS, M.: User-Defined Gestures for Connecting Mobile Phones, Public Displays, and Tabletops. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10)*. New York, NY, USA : ACM, 2010, S. 239–248 (Zitiert auf Seite 31)
- [44] LONG, A. C. Jr. ; LANDAY, J. A. ; ROWE, L. A.: Implications For a Gesture Design Tool. In: *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit (CHI '99)*. New York, NY, USA : ACM, 1999, S. 40–47 (Zitiert auf Seite 34)
- [45] MARTENS, M. H. ; VAN WINSUM, W.: *Measuring distraction: the peripheral detection task*. Driver Distraction Internet Forum. <http://www-nrd.nhtsa.dot.gov/departments/HumanFactors/driver-distraction/PDF/34.PDF>. Version: Juni 2000. – letzter Aufruf am 10.11.2011 (Zitiert auf Seite 18)
- [46] MATTES, S.: The Lane-Change-Task as a Tool for Driver Distraction Evaluation. In: STRASSER, H. (Hrsg.) ; KLUTH, K. (Hrsg.) ; RAUSCH, H. (Hrsg.) ; BUBB, H. (Hrsg.): *Quality of Work and Products in Enterprises of the Future*. Stuttgart : Ergonomia Verlag, 2003, S. 57–60 (Zitiert auf den Seiten 19 und 91)
- [47] MICROSOFT: *Microsoft Text-to-Speech*. <https://www.microsoft.com/reader/de/downloads/tts.mspx>. – letzter Aufruf am 14.11.2011 (Zitiert auf Seite 63)

- [48] MICROSOFT MSDN: *SpeechRecognizer-API*. <http://msdn.microsoft.com/en-us/library/ms587649.aspx>. – letzter Aufruf am 14.11.2011 (Zitiert auf Seite 64)
- [49] MICROSOFT MSDN: *SpeechSynthesizer-API*. <http://msdn.microsoft.com/en-us/library/ms596245.aspx>. – letzter Aufruf am 14.11.2011 (Zitiert auf Seite 63)
- [50] MÜLLER, C. ; WEINBERG, G.: Multimodal Input in the Car, Today and Tomorrow. In: *IEEE Multimedia* 18 (2011), Januar, Nr. 1, S. 98–103 (Zitiert auf den Seiten 32, 33 und 34)
- [51] NIELSEN, J.: Heuristic Evaluation. In: NIELSEN, J. (Hrsg.) ; MACK, R.L. (Hrsg.): *Usability Inspection Methods*. 1. New York, NY, USA : John Wiley & Sons Inc., April 1994, Kapitel 2, S. 25–62 (Zitiert auf Seite 35)
- [52] NINTENDO: *Wii*. <http://wii.nintendo.de>. – letzter Aufruf am 18.11.2011 (Zitiert auf Seite 73)
- [53] OVIATT, S.: Multimodal Interfaces. In: SEARS, A. (Hrsg.) ; JACKO, J. A. (Hrsg.): *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. 2. Mahwah, New Jersey : Lawrence Erlbaum Associates Inc, September 2007, S. 413–432 (Zitiert auf den Seiten 31 und 32)
- [54] PAUZÍÉ, A.: A Method to assess the driver mental workload : The driving activity load index (DALI). In: *IET Intelligent Transport Systems Journal* 2 (2008), Nr. 4, S. 315–322 (Zitiert auf den Seiten 79, 82 und 92)
- [55] REISSNER, U.: Gestures and Speech in Cars. Version: März 2007. http://www14.informatik.tu-muenchen.de/konferenzen/Jass07/courses/3/3_Gestures-and-Speech-in-Cars_Report.pdf. In: *Electronic Proceedings of Joint Advanced Student School 2007 (JASS 2007)*. St. Petersburg, März 2007. – letzter Aufruf am 14.11.2011 (Zitiert auf Seite 34)
- [56] RIENER, A.: *Sensor-Actuator Supported Implicit Interaction in Driver Assistance Systems*, Universität Linz, Diss., 2009. – S. 80 (Zitiert auf Seite 26)
- [57] SHNEIDERMAN, B. ; PLAISANT, C. ; COHEN, M. ; JACOBS, S.: *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. 5. Boston, MA, USA : Pearson, 2009 (Zitiert auf Seite 35)
- [58] STEVENS, A. ; QUIMBY, A. ; BOARD, A. ; KERSLOOT, T. ; BURNS, P.: Design guidelines for safety of in-vehicle information systems / TRL Limited. 2002. – Forschungsbericht (Zitiert auf den Seiten 15 und 16)
- [59] TAYLOR, P.: *Text-to-Speech Synthesis*. 1. Cambridge : Cambridge University Press, 2009 (Zitiert auf Seite 27)

- [60] TÖNNIS, M. ; BROY, V. ; KLINKER, G.: A Survey of Challenges Related to the Design of 3D User Interfaces for Car Drivers. In: *Proceedings of the 1st IEEE Symposium on 3D User Interfaces (3DUI '06)*. Washington, DC, USA : IEEE Computer Society, 2006, S. 127–134 (Zitiert auf den Seiten 23, 25, 26 und 32)
- [61] TOMA, F. E.: Fahrer-Interaktionssysteme: Integration von Entertainment, Information, Komfort-Elektronik und Fahrer-Assistenz. In: TIEMANN, R.C. (Hrsg.): *Cockpits im Automobil*. 1. Renningen : Expert Verlag, Mai 2005, S. 8–18 (Zitiert auf Seite 27)
- [62] TUIO.ORG: *Software Implementing TUIO*. <http://www.tuio.org/?software>. – letzter Aufruf am 14.11.2011 (Zitiert auf Seite 65)
- [63] UNIVERSITÄT DUISBURG-ESSEN, Pervasive Computing and User Interface Engineering: CARS. <http://cars.pcuie.uni-due.de>. – letzter Aufruf am 17.11.2011 (Zitiert auf Seite 21)
- [64] URBAS, L. ; LEUCHTER, S. ; PAPE, N. ; TRÖSTERER, S.: Prospektive Gestaltung der Ablenkungswirkung von In Vehicle Information Systems. In: JÜRGENSOHN, Th. (Hrsg.) ; KOHLREP-ROMETSCH, H. (Hrsg.): *Tagungsband der 2. Berliner Fachtagung Fahrermodellierung, 19./20. Juli 2008*. Düsseldorf : VDI-Verlag, 2009, S. 74–83 (Zitiert auf den Seiten 17, 18 und 19)
- [65] U.S. DEPARTMENT OF TRANSPORTATION: National Household Travel Survey. Version: Januar 2010. <http://www.fhwa.dot.gov/policy/2008cpr/pdfs/cp2008.pdf>. In: *2008 Status of the Nation's Highways, Bridges, and Transit: Conditions and Performance*. U.S. Department of Transportation, Januar 2010, Kapitel 15, 15.1.-15.24. – letzter Aufruf am 22.11.2011 (Zitiert auf Seite 11)
- [66] U.S. NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION: Distracted Driving 2009 / U.S. National Highway Traffic Safety Administration. 2010. – Forschungsbericht (Zitiert auf den Seiten 16 und 17)
- [67] WEINBERG, G. ; HARSHAM, B. ; FORLINES, C. ; MEDENICA, Z.: Contextual Push-to-Talk: Shortening Voice Dialogs To Improve Driving Performance. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. New York : ACM, September 2010, S. 113–122 (Zitiert auf den Seiten 11 und 28)
- [68] WINTER, U. ; GROST, T. J. ; TSIMHONI, O.: Language Pattern Analysis for Automotive Natural Language Speech Applications. In: *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '10)*. New York, NY, USA : ACM, 2010, S. 34–41 (Zitiert auf Seite 27)

- [69] WOBROCK, J. O. ; MORRIS, M. R. ; WILSON, A. D.: User-Defined Gestures for Surface Computing. In: *Proceedings of the 27th international conference on Human factors in computing systems (CHI '09)*. New York, NY, USA : ACM, 2009, S. 1083–1092 (Zitiert auf den Seiten 30, 31, 42 und 52)
- [70] WOBROCK, J. O. ; MYERS, B. A. ; AUNG, H. H. ; LOPRESTI, E. F.: Text Entry from Power Wheelchairs: EdgeWrite for Joysticks and Touchpads. In: *Proceedings of the 6th international ACM SIGACCESS conference on Computers and accessibility (Assets '04)*. New York, NY, USA : ACM, 2004, S. 110–117 (Zitiert auf Seite 30)
- [71] WRIGHT, M.: Open Sound Control: an enabling technology for musical networking. In: *Organised Sound* 10 (2005), Dezember, Nr. 3, S. 193–200 (Zitiert auf Seite 65)
- [72] WU, Q.: An Overview of Driving Distraction Measure Methods. In: *IEEE 10th International Conference on Computer-aided Industrial Design & Conceptual Design (CAID&CD 2009)*. Wenzhou, China, November 2009, S. 2391 –2394 (Zitiert auf Seite 18)
- [73] ZELLER, A. ; WAGNER, A. ; SPRENG, M.: iDrive - Zentrale Bedienung im neuen 7er von BMW. In: *Elektronik im Kraftfahrzeug, VDI-Berichte Nr. 1646*. Düsseldorf : VDI-Verlag, 2001, S. 997–1009 (Zitiert auf Seite 25)

Inhalt der CD-ROM

Abstract

- Abstract der Arbeit in Deutsch und Englisch

Diplomarbeit

- DA_KienastMichael_MultimodaleInteraktion.pdf

Literatur

- Sämtliche in digitaler Form vorliegende Literatur

Software der explorativen Benutzerstudie

- Android-Applikation für eine Beispielaufgabe
- Android-Applikation mit 26 Aufgaben zur Sammlung von Sprachkommandos und Gesten
- Android-Applikation zur Analyse der aufgezeichneten Daten

Auswertung der explorativen Benutzerstudie

- Auswertungen der Eingangs-, Zwischen- und Abschlussfragebögen
- Auswertung der aufgezeichneten Sprachkommandos
- Auswertung der aufgezeichneten Gesten

Software und Dokumentation/Handbuch des Prototypen

- C#-Projektmappe sowie sämtliche benötigten Bibliotheken, Videos und Grafiken
- Für den Prototypen angepasster CARS-Fahrsimulator
- Android-Applikation TUIOdroid zur Gestenaufnahme
- Dokumentation und Handbuch für den Prototypen

Software der Evaluierung

- Android-Applikation zur Bedienung von Funktionen über Buttons auf dem Lenkrad
- WiiConnectorStub zum Versand von Wii-Controller-Events über das EIToolkit
- Lane-Change-Test, Analyse-Tool und aufgezeichnete Fahrtdaten

Auswertung der Evaluierung

- Auswertungen der Eingangs-, Zwischen- und Abschlussfragebögen
- Auswertung der aufgezeichneten LCT-Daten

Erklärung

Hiermit versichere ich, diese Arbeit selbständig verfasst und nur die angegebenen Quellen benutzt zu haben.

(Michael Kienast)