

0	Einleitung	1
1	Einordnung der Themenstellung, Grundlagen und Zielsetzung	3
1.1	Endbenutzersysteme und Benutzerfreundlichkeit	3
1.2	Argumente für und gegen die natürliche Sprache	5
1.2.1	Benutzergruppen	6
1.2.2	Argumente für die Verwendung natürlichsprachiger FAS	7
1.2.3	Argumente gegen die Verwendung natürlichsprachiger FAS	10
1.2.4	Diskussion der Argumente gegen natürlichsprachige FAS	13
1.3	Komplexitätsbarriere und 'habitability'	15
1.3.1	Die Komplexitätsbarriere in der Computerwissenschaft	15
1.3.2	Ausgrenzung der künstlichen Intelligenz durch Krägeloh 1978	17
1.3.2.1	Zum Selbstverständnis der Künstlichen Intelligenz	
1.3.2.2	Natürlichsprachige FAS der siebziger Jahre	
1.3.2.3	Fazit	
1.3.3	'Habitability'	21
1.3.3.1	Die Vorschläge von Watt 1968	
1.3.3.2	Linguistische und konzeptuelle Vollständigkeit	
1.3.4	Subsets bestehender FAS und ihre Begründung	27

VIII		
1.4	Klärungsdialog	30
1.5	Zusammenfassung	33
2	Argumente, die den Aufbau natürlichsprachiger FAS bestimmen	35
2.1	Zur Argumentation von PLIDIS	38
2.1.1	Systemaufbau	38
2.1.2	Praxisorientierung und Linguistik	39
2.1.3	Fazit zu PLIDIS	41
2.2	'User Speciality Languages' (USL): Beschreibung, Testtext und Argumentationsstrukturen	43
2.2.1	Benutzergruppen, Aufgabenspezifikation und ihr Einfluß auf die Grundkonzeption	44
2.2.1.1	Benutzergruppen	
2.2.1.2	Aufgabenspezifikation	
2.2.1.3	Konsequenzen	
2.2.1.4	Nebenziel: USL als Werkzeug für die Linguistik	
2.2.2	Gesamtüberblick über USL	47
2.2.2.1	Hauptkomponenten	
2.2.2.2	Dreistufige Analyse	
2.2.2.3	Zugrundeliegende Prinzipien und ihre Konsequenzen	
2.2.3	ISBL von PRTV	50
2.2.3.1	Strukturen	
2.2.3.2	Operationen zur Bildung von ISBL-Ausdrücken	
2.2.3.3	Anwendung der ISBL-Ausdrücke in Anfragen und weitere Befehle	
2.2.3.4	Neu- und Umdefinition von Relationen	
2.2.4	Anwendungsabhängige Strukturen und ihre Umsetzung	57
2.2.5	Anwendungsunabhängige Strukturen	59
2.2.5.1	Zeitausdrücke	
2.2.5.2	Das Fragepronomen <u>wieviel</u>	
2.2.5.3	Koordination	
2.2.5.4	Pronomen	

2.2.6	Formale Erweiterung der natürlichsprachigen Anfragemöglichkeiten	71
2.2.7	Hilfsprogramme	72
2.2.8	Zusammenfassung zu USL	73
2.3	Fazit	74
3	Evaluierungsstudien und Evaluierungsmethoden	78
3.1	Evaluierungsstudien zu natürlichsprachigen FAS	78
3.1.1	Pauschalhinweise auf Anwendungstests	79
3.1.2	Informelle Tests	79
3.1.3	Kontrollierte Tests	81
3.1.4	Tests in realen Anwendungssituationen	82
3.1.5	Simulierte natürlichsprachige MMI	86
3.1.6	Fazit	87
3.2	Evaluierung formaler Anfragesprachen und Vergleiche mit natürlichsprachigen Abfragemöglichkeiten	89
3.2.1	Formale Anfragesprache	90
3.2.2	Formale versus natürlichsprachige Anfragesprache	91
3.3	Evaluierungsmethoden und Systementwicklung	93
3.3.1	Intuition, Fragebögen, kontrollierte Tests	94
3.3.2	Zwischenmenschlicher Dialog versus MMI	95
3.3.3	Reale Anwendungssituation	99
3.3.4	Einbettung der Einzelevaluierung in ein größeres Evaluierungskonzept	99
4	Die KFG-Studie	106
4.1	Themenstellung, Verlauf der Studie und Hilfsmittel der Evaluierung	106
4.1.1	Themenstellung und Benutzermotivation	106
4.1.2	USL-Version	107

X		
4.1.3	Datengrundlage und Einrichten der Studie	108
4.1.4	Sitzungen zur Problemlösung	109
4.1.5	Primärdaten der Auswertung und ihre Gewinnung	110
4.1.5.1	Sitzungsprotokolle und ihre Aufbereitung	
4.1.5.2	Die Notizen zu den Sitzungsbeobachtungen und die Eigenkommentierungen des Benutzers	
4.1.5.3	Fragebögen	
4.2	Fehlerdefinition und Gesamtfehlerquoten	113
4.2.1	Fehlerdefinition und Restriktionen	114
4.2.1.1	Fehlerdefinition	
4.2.1.2	Restriktionen	
4.2.1.3	Anfrage und Datenbankinhalt als Konstanten	
4.2.1.4	Das Ermitteln der Fehler	
4.2.2	Die Gesamtfehlerquoten in der KFG-Studie	117
4.2.2.1	Gesamtfehlerquoten pro Benutzer und Sitzung	
4.2.2.2	Untergliederungen	
4.2.2.3	Einfluß des Dialogtyps	
4.3	Fehlertypen und ihre Entwicklung	127
4.3.1	Softwareevaluierung versus Evaluierung experimenteller Systeme mit natürlichsprachiger Abfragekomponente	127
4.3.1.1	Softwareevaluierung	
4.3.1.2	Evaluierung experimenteller Systeme	
4.3.1.3	Natürlichsprachige Abfrage als Gegenstand des Interesses	
4.3.1.4	Fazit	
4.3.2	Richtlinien der Klassifizierung und der Ermittlung der Fehlerklasse	129
4.3.2.1	Abgrenzung der Fehlerklassen untereinander	
4.3.2.2	Zuordnung zu einer Fehlerklasse	
4.3.2.3	Fehlermeldungen des Systems	
4.3.2.4	Ermittlung der Fehlerklassen in einer Fehlersituation	
4.3.3	Die Fehlerklassen der KFG-Studie	133
4.3.3.1	Beschreibung der einzelnen Fehlerklassen	

4.3.3.2	Überblick	
4.3.3.3	Probleme der Zählung	
4.3.3.3.1	Doppelfehler	
4.3.3.3.2	Folgefehler	
4.3.4	Verteilung der Fehler auf die einzelnen Fehlerklassen	152
4.3.4.1	Verteilung der Fehlerklassen auf die Benutzer	
4.3.4.2	Fehlerklassen mit hohen Fehleranzahlen	
4.3.4.3	Entwicklung der Hauptfehlergruppen des Benutzers S (Teil: Natürlichsprachiger Subset im weiteren Sinn)	
4.3.4.3.1	Entwicklung von <u>Präp+W</u>	
4.3.4.3.2	Entwicklung von <u>Rel+Rund</u>	
4.3.4.3.3	Entwicklung von <u>SP+VB+Dur</u>	
4.3.4.3.4	Entwicklung von <u>Vund+Vansch</u>	
4.3.4.3.5	Interpretation	
4.3.4.4	Entwicklung der Fehlerklassen, die eine funktionale Lücke anzeigen	
4.3.4.5	Entwicklung der Fehlergruppe <u>Tipp</u>	
4.4	Fehlerfolgen und Fehlerstrategien	165
4.4.1	Einzelfehler, isolierte Fehler und Fehlerfolgen	165
4.4.1.1	Abgrenzung	
4.4.1.2	Definition der Fehlerfolge	
4.4.1.3	Einschränkungen	
4.4.1.4	Doppelfolgen	
4.4.1.5	Pseudofehlerfolge bei korrekter Antwort "Es gibt kein Objekt mit der gewünschten Eigenschaft"	
4.4.1.6	Abschluß der Fehlerfolge	
4.4.2	Anteil der Fehler in Fehlerfolgen an der Gesamtfehlerquote	173
4.4.3	Fehlerstrategien	174
4.4.3.1	Hauptgruppen	
4.4.3.2	Die Strategien der KFG-Studie	
4.4.3.3	Strategienverteilung und -häufigkeit	
4.4.4	Fehlerklassen und Fehlerfolgen	179

## XII

4.4.4.1	Häufigkeiten der Fehlerfolgen (Hauptfehlerklassen)	
4.4.4.2	Einzelanalysen der Hauptfehlerklassen	
4.4.4.3	Vergleich der erfolgreich gehandhabten Fehlerfolgen von <u>Sp</u> und <u>Präp</u>	
4.4.4.4	Fehlerfolgen zu <u>Tipp</u>	
4.4.4.5	Fazit	
4.5	Zur Definition und Verwendung der Relationen in der KFG-Studie	189
4.5.1	Undefinierte Relationen und die Verwendung der Relationen durch die Benutzer	191
4.5.2	Von Rel oder Präp betroffene Relationen	192
4.5.3	Definierte Relationen versus benutzte Relationen	196
4.5.4	Fazit	198
4.6	Strukturuntersuchungen zur KFG-Studie	199
4.6.1	Strukturzählungen	199
4.6.2	Anfragestrukturen	200
4.7	Erweiterung der natürlichsprachigen Anfragen durch formale Elemente	201
4.7.1	Mengenoperationen und Variablendefinitionen statt Koordination	202
4.7.2	Pronominalisierung und Mengenbildungsoperationen	203
4.7.3	Fazit	205
4.8	Zusammenfassung der Ergebnisse der KFG-Studie	206
4.8.1	Inhaltliche Ergebnisse	206
4.8.2	Methodik der Fehleranalyse	208
5	Die TA-Studie	209
5.1	Bisherige Informationsgewinnung des TAB	210
5.1.1	Die TA-Berichtsdatei	210
5.1.2	Benutzung der Berichtsdatei	211

5.1.3	Die formale Abfragesprache IQRP	211
5.2	Aufbau oder USL-Anwendung für den TA	212
5.2.1	Generelles Vorgehen	212
5.2.2	Anpassung der TA-Berichtsdatei an die USL-Anwendung	213
5.2.3	Das TA-Vokabular und die TA-Relationen	214
5.2.3.1	Vortests mit dem TAB und Definition des TA-Vokabulars	
5.2.3.2	Probleme beim Aufbau der ersten Relationenliste	
5.2.3.2.1	Elliptische Ausdrucksweise	
5.2.3.2.2	Eindeutige Präpositionen	
5.2.3.2.3	Fazit	
5.2.3.3	In USL-T2 noch nicht berücksichtigte Probleme	
5.3	Erste Realisierung der TA-Anwendung mit USL-T1	221
5.3.1	Tests mit reduzierter Datenmenge	221
5.3.1.1	Abhängigkeit der Frageformulierung vom Datentyp	
5.3.1.2	Formaler versus natürlichsprachiger Ausdruck der Zeit	
5.3.2	Erste Tests mit der vollständigen Datenbank (USL-T1)	222
5.3.3	USL-T2 als Konsequenz aus den Vortests	224
5.3.3.1	Leistungserweiterungen	
5.3.3.2	Die Performanz von USL-T2	
5.4	Erste Terminalsitzung des TAB mit USL-T2	226
5.4.1	Fehlerursachen	226
5.4.2	Konsequenzen aus der ersten Terminalsitzung des TAB	229
5.5	Fazit	230
5.5.1	Die Verwendung nichteindeutiger Präpositionen	230
5.5.2	Syntaxfehler bei komplexen Ausdrücken	231
5.5.3	Elliptische Ausdrucksweise	231
5.5.4	Umgebung der Anfrage	231

6	Natürliche Sprache versus formale Abfragesprache	233
6.1	Testgruppen, Testanordnung und Testinhalte	233
6.1.1	Tests auf der Grundlage der KFG-Studie	234
6.1.1.1	Die Gruppe mit ISBL-Erfahrung	
6.1.1.2	Studentengruppen	
6.1.1.2.1	Erste Studentengruppe	
6.1.1.2.2	Zweite Studentengruppe	
6.1.2	Tests auf der Grundlage der TA-Studie	242
6.2	Vergleich der CPU-Zeiten	243
6.2.1	Festlegung der Vergleichsgrundlage	243
6.2.1.1	Fehlende Funktionen und Fehlersätze	
6.2.1.2	Ersatzfragen	
6.2.2	CPU-Mehrbedarf bei natürlichsprachigen Anfragen	245
6.3	Kurze formale Anfragen versus lange natürlichsprachige	246
6.4	Fehleranalyse und Lerneffekt	248
6.4.1	Fehleranalyse der Gruppe mit ISBL-Erfahrung	248
6.4.2	Die Tests mit den Studentengruppen	251
6.4.2.1	Leistungstest	
6.4.2.2	Haupttest	
6.4.3	Zusammenfassung der inhaltlichen Ergebnisse	260
6.4.3.1	Gesamtvergleich	
6.4.3.1.1	Einfluß des Vergessens	
6.4.3.1.2	Qualitative Unterschiede	
6.4.3.1.3	Formale Abfragesprachen nach 'human factors'-Gesichtspunkten	
6.4.3.2	Einfluß der EDV-, ISBL- und Datenbankerfahrung	
6.4.4	Methodische Überlegungen	265
6.4.4.1	Terminalsitzung versus Schreibtischtest	
6.4.4.2	Leistungstest versus Haupttest	
6.4.4.3	Reale Anwendung versus Übersetzungstest	

6.5	Zusammenfassung	267
6.5.1	Inhaltliche Ergebnisse	267
6.5.2	Testmethodik	268
7	Zusammenfassung der Ergebnisse und Schlußfolgerungen	270
7.1	Ergebnisse zu USL	272
7.1.1	Einzelprobleme im Zusammenhang mit der speziellen USL-Übertragbarkeitsanforderung	272
7.1.2	Erweiterung der semantischen Basis von USL	273
7.2	Aussagen, die über USL hinausgehen	275
7.2.1	Einzelfragen	275
7.2.2	Ergebnisse zur Subset-Problematik	277
7.3	Zur Methodik der Evaluierungsstudien	279
7.4	Zum Stellenwert natürlichsprachiger Abfragemöglichkeiten bei Informationssystemen (Ausblick)	280
8	Literatur	282

#### ANHÄNGE auf MICROFICHE

- 1: Listen und Tabellen zur KFG-Studie
- 2: Variablendefinitionen der KFG-Studie
- 3: Listen und Tabellen zur TA-Studie
- 4: Listen und Tabellen zum Vergleich formale versus natürlichsprachige Anfragen