

Ausarbeitung des Referates
**„Kommunikation mit anderen Spezies -
Syntaxfreie Interaktionen in der
Mensch-Maschine-Kommunikation“**
zum Seminar
„Kommunikation: Das Miteinander von
Mensch und Maschine“

Andreas Bunge ¹

13. August 2000

¹abunge@techfak.uni-bielefeld.de

Inhaltsverzeichnis

1	Einführung	4
2	WIMP	4
2.1	Was bedeutet WIMP?	4
2.2	Kritikpunkte	6
3	Einflüsse auf die Mensch-Computer-Interaktion	7
4	Syntaxfreie Interaktionen in Benutzerschnittstellen	8
4.1	Syntaxfreie Interaktionen - Definition	8
4.2	Unterschied bisherige - zukünftige Benutzerschnittstellen	8
4.3	Gestaltungsaufgaben	9
5	Interaktion mit dem Computer	10
5.1	Interaktionsorganisation	10
5.2	Interaktionstechniken	11
5.2.1	Blickgesteuerte Interaktionen	11
5.2.2	Manuelle Interaktionen	11
5.2.3	Gesten- und nonverbale Interaktionen	12
5.2.4	Sprachbasierte Interaktionen	12
5.2.5	Ausgabetechnologien	12
6	Zusammenfassung/Ausblick	13

1 Einführung

Das Problem heutiger Computersysteme liegt darin, daß sie dem Menschen nicht so dienen, wie ihr Leistungspotential es vermuten ließe. So bieten sie dem Benutzer zwar sehr viele Funktionen, es dauert aber sehr lange, bis all diese Funktionen erlernt werden können. Computerarbeit bedarf also bislang einer hohen Einarbeitungszeit.

Genau an diesem Punkt setzt diese Ausarbeitung an: Nicht der Mensch soll lernen, mit dem Computer umzugehen, sondern der Computer soll sich so auf seinen Benutzer einstellen, daß dieser sich schnell zurechtfindet.

Um dieses Ziel zu erreichen, ist es zunächst einmal wichtig, die bisherigen Ansätze zur Mensch-Computer-Interaktion zu analysieren, um ihre Schwachpunkte zu erkennen. Da auf WIMP basierende Benutzerschnittstellen am weitesten verbreitet ist, stelle ich dieses Schnittstellenprinzip in Kapitel 2 kurz vor, um dann einige Kritikpunkte aufzuzeigen, die in zukünftigen Ansätzen verbessert werden sollten. Natürlich ist es ebenso wichtig zu wissen, welche Faktoren überhaupt Einfluß auf die Kommunikation zwischen Mensch und Computer haben, nachzulesen in Kapitel 3. In Kapitel 4 wird dann der neue Ansatz der „Syntaxfreien Interaktionen“ vorgestellt, auf denen auch das im Rahmen der zugrundeliegenden Diplomarbeit ([J. H. Israel]) entstandene Betriebssystem VOS basiert.

Kapitel 5 faßt einige Interaktionstechniken zusammen, die in neuartigen Systemen verwendet werden könnten.

Wenn nicht anders angegeben, sind die in dieser Ausarbeitung verwendeten Abbildungen aus [J.H. Israel] entnommen.

2 WIMP

2.1 Was bedeutet WIMP?

WIMP ist eine Abkürzung und steht für „**W**indow **I**con **M**enu **P**ointing devices“. Aufgrund des Namens läßt sich schon vermuten, wie in etwa Programme aussehen, die auf dieser Schnittstelle basieren: Die auf dem Bildschirm angezeigten Daten werden in Fenstern organisiert, die entweder überlappend, gekachelt oder maximiert dargestellt werden können (Abb. 1) . Mit Hilfe dieser Fenstertechnik ist es für den Benutzer leicht möglich, mehrere Aufgaben parallel zu bearbeiten („Multitasking“). Er kann die gewünschten Kommandos mit Hilfe von Zeigeeinstrumenten auswählen, also i.d.R. mit der Maus. Diese Kommandos stehen dem Benutzer in Menüs zur Verfügung oder werden als Icons dargestellt (Abb. 2).

WIMP-Schnittstellen sind seit 1983 immer mehr verbreitet und folgen dem Desktop-Prinzip, d.h. die Anordnung der Daten soll sich an einem Schreibtisch orientieren, daher auch der Begriff „Ordner“ für Verzeichnis.

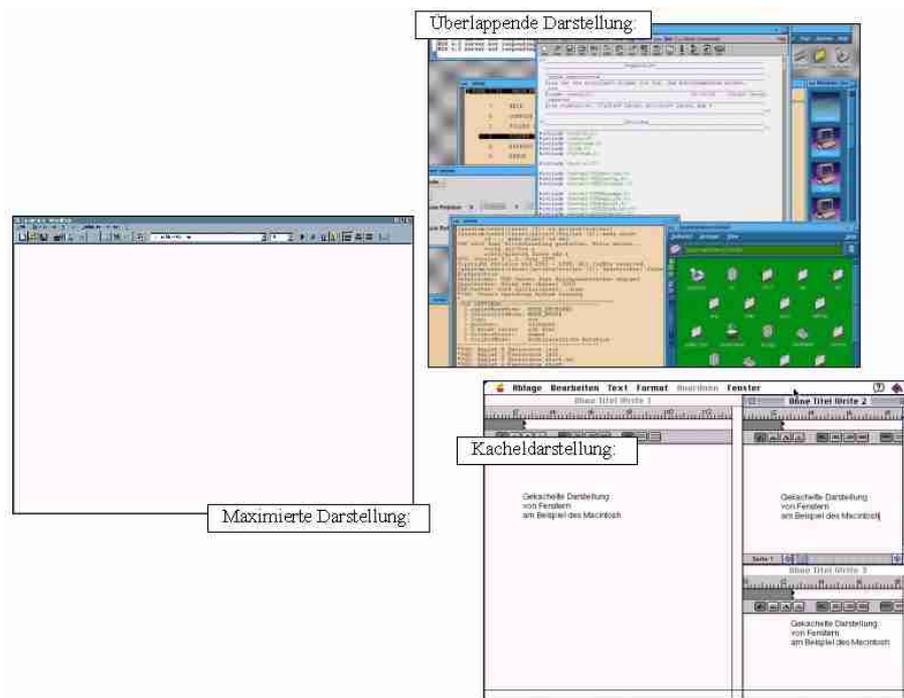


Abbildung 1: Verschiedene Fenster-Anordnungen

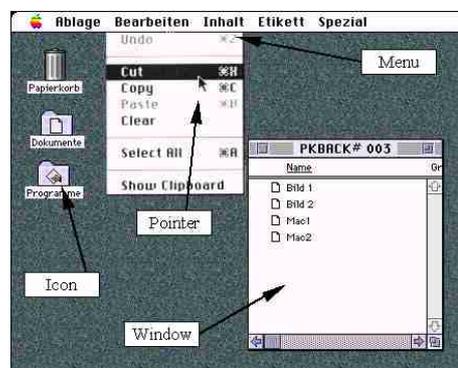


Abbildung 2: Die Elemente einer WIMP-Schnittstelle

2.2 Kritikpunkte

Der WIMP-Ansatz basiert nicht auf der systematischen Analyse der Bedürfnisse des Menschen. Es wurde also nicht analysiert, welche Bedürfnisse der Mensch bzgl. der Kommunikation mit dem Computer hat, sondern es wurde zunächst ein Ansatz entwickelt, der dann dem Benutzer „vorgesetzt“ wurde. Das eigentliche Ziel, nämlich der benutzerangemessene Umgang mit dem Rechner, ist aufgrund folgender Kritikpunkte nicht ganz erfüllt worden.

So muß der Benutzer auf jeden Fall die Kommandos erlernen, die der Computer ihm zur Verfügung stellt. Auch wenn das Lernen durch Icons und Menüs vereinfacht wird, ist das sicherlich nicht optimal.

Ein weiterer Kritikpunkt ist die Verwendung von generischen Dialogen (z.B. der „Speichern unter“-Dialog), die zwar dem Programmierer einer Anwendung viel Arbeit abnehmen, aber keine automatische Anpassung an die Bedürfnisse des Benutzers ermöglichen. Denkbar wären z.B. Menüs, die sich dermaßen anpassen, daß sich oft verwendete Einträge im oberen Teil des Menüs anfinden, was allerdings auch gewöhnungsbedürftig ist, v.a. wenn man sich schon an die bisherigen Menüs und Dialoge gewöhnt hat.

Weiterhin ist die strikt sequentielle Kommunikation zwischen Mensch und Computer zu bemängeln. Sequentielle Kommunikation heißt in diesem Fall, daß immer nur einer der beiden Kommunikationspartner - Mensch oder Computer - agieren kann, während der andere auf seinen Einsatz wartet. Auch sind verschiedene Eingabegeräte nicht sinnvoll parallel nutzbar, wie es zum Beispiel in Computerspielen der Fall ist (Maus und Tastatur). Die Folge daraus ist eine sehr niedrige Bandbreite vom Menschen zum Computer, während die Bandbreite in umgekehrter Richtung aufgrund der hohen Leistungsfähigkeit heutiger Rechner sehr hoch sein kann. Der letzte Punkt kann zur Informationsüberflutung führen.

Die Frage ist natürlich nun: Wie kann man eine bessere Mensch-Computer-Interaktion zustandebringen? Um diese Frage zu beantworten, ist es zunächst einmal wichtig zu wissen, welche Faktoren überhaupt Einfluß auf die Mensch-Computer-Kommunikation haben.

3 Einflüsse auf die Mensch-Computer-Interaktion

Folgende Punkte sind auf jeden Fall bei der Planung einer neuartigen Benutzerschnittstelle zu berücksichtigen:

- **Welche Aufgaben will der Benutzer mit dem Rechner erledigen?**
Einfache Aufgaben können z.B. das Kreieren, Auswählen, Positionieren oder das Löschen von Objekten sein. I.d.R. sind aber v.a. komplexere Aufgaben wie das Schreiben von Texten oder Programmieren von Anwendungen zu erledigen.
- **Eigenschaften/Fähigkeiten des Menschen**
Hier sind v.a. die sensorischen, kognitiven und motorischen Fähigkeiten des Menschen einzubeziehen.
- **Nachrichtenaustausch**
Hierbei ist zu beachten, daß bei jeder Art von Kommunikation auch Störungen auftreten oder auch Nachrichten gar nicht ankommen können. Entsprechende Vorkehrungen zur Fehlererkennung und -minimierung/-beseitigung müssen eingeplant werden.
- **Verfügbare Technik**
Der beste Ansatz nützt nichts, wenn die entsprechende Technik nicht vorhanden ist, um diesen Ansatz in die Realität umzusetzen. Also muß man sich immer auch vor Augen halten, was technisch realisierbar ist und was nicht.

Wird der Computer nur ungenügend an die Fähigkeiten des Menschen angepaßt, so kann das z.T. gefährliche Folgen haben. So kann der Benutzer durch Informationsüberflutung und einen zu hohen Abstraktionsgrad der dargestellten Informationen überfordert werden.

Oder aber er wird durch hochautomatisierte Prozesse und monotone Arbeitsabläufe unterfordert. In beiden Fällen erhöht sich das Risiko von Unfällen, wenn man an den Einsatz von Benutzerschnittstellen in Fabrikmaschinen denkt.

Mithilfe der im nächsten Kapitel eingeführten „Syntaxfreien Interaktionen“ soll diese Gefahr vermieden werden.

4 Syntaxfreie Interaktionen in Benutzerschnittstellen

4.1 Syntaxfreie Interaktionen - Definition

„Syntaktische Handlungsabläufe sind zusammengefügte Handlungen, bei denen jede Handlung ihren Wert erst durch die Fügung erhält. Das Interaktionsverhalten einer Benutzerschnittstelle gilt als syntaxfrei, wenn sich der Mensch während der Mensch-Computer-Interaktion die Syntax seiner Handlungen nicht verdeutlicht. Dazu kann er sich der Vielzahl seiner natürlichen und abstrakter Interaktionsmöglichkeiten bedienen.“

([Nielsen], Übersetzung aus [J. H. Israel])

Ziel der syntaxfreien Interaktionen ist die Orientierung an natürlichen Arbeitsabläufen und somit eine Verringerung der Einarbeitungszeit, indem nicht spezielle Funktionen und Kommandos des Computers erlernt werden müssen, sondern der Computer die Eingaben des Benutzers interpretiert.

4.2 Unterschied bisherige - zukünftige Benutzerschnittstellen

Alle bisherigen Ansätze und somit auch der WIMP-Ansatz sind nicht syntaxfrei, denn der Umgang mit dem Computer muß erlernt werden. Und beim Lernen muß sich der Mensch der Syntax seiner Handlungen sehr wohl bewußt sein. In zukünftigen Systemen soll es so sein, daß der Computer den Menschen „versteht“, ohne daß dieser den Umgang mit dem Computer erlernen muß.

Die Rolle des Computers wird sich dahingehend verändern, daß er selbständiger agiert, d.h. im Gegensatz zu bisherigen Systemen ist die Vorgabe nicht mehr, genau das zu tun, was der Benutzer ihm explizit „sagt“ - was zugegebenermaßen häufig nicht funktioniert -, sondern er soll das tun, was der Benutzer *will*. Das soll durch Beobachtung des Benutzers, z.B. mit Hilfe von Tracking-Systemen, realisiert werden. So kann z.B. der aktuelle Zustand des Benutzers mit Biosensoren (s. Kap. 5.2.3) o.ä. erkannt werden. Dadurch kann festgestellt werden, ob der Benutzer mit dem Ergebnis zufrieden ist, das der Computer ihm präsentiert und evtl. sofort ein neuer Vorschlag gemacht werden, falls der Benutzer unzufrieden erscheint.

Informationsüberflutung soll verhindert werden, indem nicht benötigte Objekte automatisch versteckt werden.

Weiterhin soll der Interaktionsstrom verändert werden, insbesondere sollen verschiedene Eingabegeräte parallel benutzt werden können und somit die weiter oben bemängelte Bandbreite vom Menschen zum Computer erhöht werden.

Ein wichtiger Punkt für zukünftige Schnittstellen sind auch kontinuierliche Benutzerdialoge, mit deren Hilfe die Wartezeiten Mensch \Leftrightarrow Computer verringert

werden sollen. Es gibt ja keine diskreten Kommandos mehr, die der Computer ausführen soll, sondern der Benutzer soll sich in einem ständigen Dialog mit dem Computer befinden.

Anhand dieser Punkte wird in [J. H. Israel] eine Systematik zur Gestaltung zukünftiger Benutzerschnittstellen beschrieben, auf die ich im nächsten Abschnitt näher eingehen möchte.

4.3 Gestaltungsaufgaben

Gestaltungsaufgaben	Gestaltungsbereiche		
	Technisches System		
	Eingabelemente	Dialog	Anzeigeelemente
Anpassung an Motorik und Sensorik	Abstimmung technischer Eingabegeräte an motorische Fähigkeiten	Überprüfen der Verträglichkeit von motorischen und sensorischen Anforderungen des Systems an Benutzer (in Ort und Zeit)	Anpassung der Darstellungsparameter der Anzeigen an Eigenschaften der Sinnesorgane
Codierung der Information	Zuordnung von motorischen Aktivitäten zu Nachrichten an das System	Abstimmung von Ein- und Ausgabecodes	Zuordnung von Nachrichten zu Sinnesreizen an den Benutzer
Organisation der Information	Struktur der Eingabelemente	Verknüpfung von Ein- und Ausgabeinformationen im Dialog; zeitliche, örtliche und inhaltliche Strukturierung von Informationen; Gestaltung der Benutzerführung	Struktur der Anzeigeelemente

Abbildung 3: Systematik zur Gestaltung zukünftiger Benutzerschnittstellen

In Abb. 3 lassen sich drei Gestaltungsbereiche erkennen, nämlich die Eingabe- und Ausgabeelemente und der Dialog als Verknüpfung dieser Elemente. In jedem dieser Gestaltungsbereiche sollten nun bestimmte Aufgaben erfüllt werden, um eine benutzerfreundliche Mensch-Computer-Kommunikation zu ermöglichen:

- Anpassung an die Motorik und Sensorik
- Codierung der Information
- Organisation der Information

Aus der Motorik und Sensorik des Menschen folgen Regeln für Abmessungen sowie für Betätigungseinrichtungen und -kräfte von mechanischen Eingabelementen. Für die Ausgabeelemente ergeben sich Darstellungsparameter für optische und akustische Anzeigen.

Mit Codierung von Informationen ist das Abbilden von Nachrichten auf motorische Aktivitäten und Sinnesreize gemeint. Dabei kann in Analog-/Digitalwerten,

alphanumerischen Zeichen, Grafik und in gesprochener/geschriebener Sprache codiert werden.

Bei der Organisation der Informationen werden mehrere zusammengehörige Daten zeitlich, örtlich und inhaltlich strukturiert. Als Beispiel für örtliche Strukturierung kann man den Ziffernblock auf der Tastatur sehen.

5 Interaktion mit dem Computer

5.1 Interaktionsorganisation

Wie bereits erwähnt, sollte ein Benutzer ohne lange Einarbeitungszeit mit dem Computer umgehen können. Ideal hierfür wäre natürlich, wenn keine Vorgaben für das Interaktionsverhalten seitens des Menschen gemacht würden, doch das ist natürlich sehr schwierig. So ist es beispielsweise selbst für den Menschen in manchen Situationen nahezu unmöglich, Gesten richtig zu interpretieren. Ziel sollte es dennoch sein, die Mensch-Computer-Interaktion so natürlich wie möglich zu gestalten.

Dabei ist es auch wichtig, daß man die Interaktion an das menschliche Vermögen zur Informationsverarbeitung anpasst, z.B. ist es sinnvoll, mehrere Informationskanäle des Menschen gleichzeitig zu nutzen wie beispielsweise den auditiven und den visuellen Kanal. Da es hierbei durchaus Unterschiede zwischen den Menschen gibt - auf manchen mag es störend wirken, wenn z.B. ein angezeigter Text zusätzlich vorgelesen wird -, ist es notwendig, daß der Computer sich im Laufe der Zeit eine Wissensbasis zum Interaktionsverhalten des Benutzers anlegt, damit er „weiß“, wie der aktuelle Anwender am besten Informationen aufnehmen kann. Gleiches gilt natürlich auch für Informationen, die der Benutzer an den Computer sendet. So können zum Beispiel bestimmte Gesten eines Menschen immer dasselbe meinen, während ein anderer Mensch etwas anderes ausdrücken möchte. Diese Gesten können dann gespeichert werden. Das nächste mal, wenn eine solche Geste von diesem Benutzer vorkommt, kann der Computer sofort eine Bedeutung zuordnen. Ob diese Interpretation richtig war, muß allerdings sicherheitshalber in jedem Fall noch einmal überprüft werden.

5.2 Interaktionstechniken

In zukünftigen Benutzerschnittstellen soll es möglich sein, Daten mit den verschiedensten Eingabegeräten-/techniken einzugeben und auszugebende Daten entsprechend zu präsentieren. Um dieses Ziel zu realisieren, werden verschiedene Interaktionstechniken angewandt, die nach [J.H. Israel] in die folgenden Kategorien eingeteilt werden können:

- Blickgesteuerte Interaktionen
- Manuelle Interaktionen
- Gesten- und weitere nonverbale Interaktionen
- Sprachbasierte Interaktionen
- Ausgabetechnologien

Kritisch sei an dieser Stelle erwähnt, daß diese Einteilung insofern nicht ganz glücklich ist, da verschiedene Kategorien wie Modus (die ersten vier Kategorien) und Medium (die Ausgabetechnologien) vermischt werden.

5.2.1 Blickgesteuerte Interaktionen

Das Auge ist aufgrund seiner hohen Geschwindigkeit eigentlich prädestiniert für schnelle und einfache Benutzereingaben, jedoch ist die Auswertung der Augenbewegungen sehr kompliziert, da das Auge nie wirklich stillsteht. Das Auge bewegt sich sehr plötzlich und in sogenannten Sakkaden. Wenn das Auge ein Objekt fixiert, so führt es trotzdem noch kleine zitternde Bewegungen aus. Es ist nicht leicht, diese von den gewollten Bewegungen zu unterscheiden und herauszufiltern.

Mit Hilfe von sogenannten Gazetracking-Systemen kann die Blickrichtung der Augen und die Position des Augapfels bestimmt werden. Als Musersatz ist das Auge aber nicht empfehlenswert, da es zu leicht überanstrengt wird. Wenn beispielsweise ein Objekt ausgewählt werden soll, so muß dieses Objekt förmlich angestarrt werden, da Latenzzeiten eingeführt werden müssen, um nicht mit jedem Blick eine Aktion auszuführen.

Dennoch lassen sich aus den Augenbewegungen wichtige Informationen über Wünsche und Intentionen des Benutzers herleiten.

5.2.2 Manuelle Interaktionen

Als Beispiel für manuelle Eingabegeräte ist der Datenhandschuh zu sehen. Hier findet - zumindest in voll immersiven Systemen mit Datenhelm (s. Kap. 5.2.3) - eine optische Rückkopplung statt, indem die Hand auch in der virtuellen Umgebung dargestellt wird.

Da aber meist keine haptische Rückkopplung stattfindet und durch virtuelle Objekte hindurchgegriffen werden kann, ermüdet der Arm sehr schnell.

Ein weiteres manuelles Interaktionsgerät ist der Touchscreen, mit dessen Hilfe Objekte direkt manipuliert werden können.

5.2.3 Gesten- und nonverbale Interaktionen

Gesten- und nonverbale Interaktionen können z.B. mit Datenhelmen, Gesichtstrackern und Biosensoren realisiert werden.

Datenhelme sind Geräte, die wie ein „normaler“ Helm auf den Kopf gesetzt werden und mit deren Hilfe es möglich ist, in eine virtuelle Umgebung „einzutauchen“.

Mit Gesichtstrackern können mimische Bewegungen und Bewegungen der Gesichtsmuskeln erfaßt werden und somit der emotionale Ausdruck des Gesichts bestimmt werden. Existierende Systeme bieten bereits Erkennung von Augenöffnung, Anhebung und Disanz der Augenbrauen, Kieferbewegungen, Mundöffnung, etc., sind aber noch nicht genau genug, um effektiv eingesetzt werden zu können.

Mit Biosensoren werden elektrische Nervenimpulse gemessen. Damit lassen sich wichtige Rückschlüsse über Muskelbewegungen oder Hirnaktivitäten ziehen und somit auch über die Emotionen des Benutzers.

5.2.4 Sprachbasierte Interaktionen

Die Spracherkennung ist bereits weiter entwickelt als die Gestenerkennung und bietet inzwischen eine relativ hohe Erkennungsrate- und geschwindigkeit. So werden sie wohl schon bald Bestandteil verbreiteter Benutzerschnittstellen sein. Mit der „TextToSpeech“-Technik wird zudem die parallele Nutzung mehrerer Informationskanäle möglich. So kann, wie an anderer Stelle erwähnt, ein angezeigter Text auch vorgelesen werden. In diesem Fall wird sowohl der auditive Kanal (das Ohr) als auch der visuelle Kanal (das Auge) genutzt.

5.2.5 Ausgabetechnologien

Das räumliche Sehen basiert auf der Auswertung monokularer und binokularer Merkmale. Monokulare Merkmale sind dabei Merkmale, die auch nur mit einem Auge wahrgenommen werden könnten wie z.B. Objektverdeckungen, Licht oder Schatten. Binokulare Merkmale sind Merkmale, die nur mit zwei Augen erfaßt werden können wie z.B. Tiefeninformationen.

Während monokulare Merkmale mit herkömmlichen Bildschirmen darstellbar sind, erfordert die Wahrnehmung binokularer Merkmale spezielle Geräte, wobei zwischen brillengebundenen und autostereoskopischen Geräten unterschieden wird. Autostereoskopische Geräte sind in [J. H. Israel] so beschrieben:

“Autostereoskopische Ausgabegeräte ermöglichen dem Betrachter ein stereoskopisches Sehen ohne Sehhilfe [...]. Dabei wird die Bildinformation entweder tat-

sächlich in den Raum projiziert [...], oder die für linkes und rechtes Auge erzeugten Bilder werden durch Veränderung der optischen Wege nur dem jeweiligen Auge sichtbar gemacht.“

Haptische Ausgabegeräte stimulieren den Tast-/Kraftsinn des Benutzers, der somit Oberflächeneigenschaften von Objekten wahrnehmen kann. Z.B. kann in Datenhandschuhen der Druck-/Tastsinn („Touch feedback“) stimuliert werden, indem kleine Luftkissen oder stumpfe Nadeln im Handschuh angebracht werden.

Akustische Ausgabegeräte sind z.B. Kopfhörer oder Lautsprecher, die so angeordnet sind, daß man Objekte im künstlichen Raum orten kann.

6 Zusammenfassung/Ausblick

In dieser Ausarbeitung wurden verschiedene Überlegungen gemacht, die unbedingt in die Entwicklung neuartiger Benutzerschnittstellen einfließen sollten, um dem Menschen den Umgang mit dem Computer zu erleichtern. Eine Implementation einer solchen Benutzerschnittstelle wird in der Ausarbeitung des Referates „Kommunikation mit anderen Spezies: Das visuelle 3D-Betriebssystem VOS“ von Markus Wanske vorgestellt. Hier zeigt sich auch, wie schwierig teilweise die Umsetzung von durchaus sinnvollen Vorgaben ist.

Literatur

- [J. H. Israel] Diplomarbeit Johann Habakuk Israel: Syntaxfreie Interaktionen im Rahmen des visuellen 3D-Betriebssystems VOS (1998), Technische Universität Berlin - Fachbereich Informatik
(<http://wwwwbs.cs.tu-berlin.de/user/kusche/diplom/>)
- [J. Nielsen] J.Nielsen: „Noncommand User Interfaces“, Communications of the ACM 36, 4, überarbeitete Ausgabe, 1996, S. 83-99