

Talking Heads

Jan Paller

1.2.2001

Seminar
„Multimodale Mensch-Maschine-Kommunikation“
WS 00/01
Universität Bielefeld

Zu beantwortende Fragen:

Was ist ein Talking-Head?

Wie funktioniert ein Talking Head?

Welche TH's gibt es bereits und welche Erkenntnisse hat man gewonnen?

Wozu braucht man ein solches System?

Was ist ein Talking Head?

bisher:

User -> Maschine

jetzt:

Maschine -> User

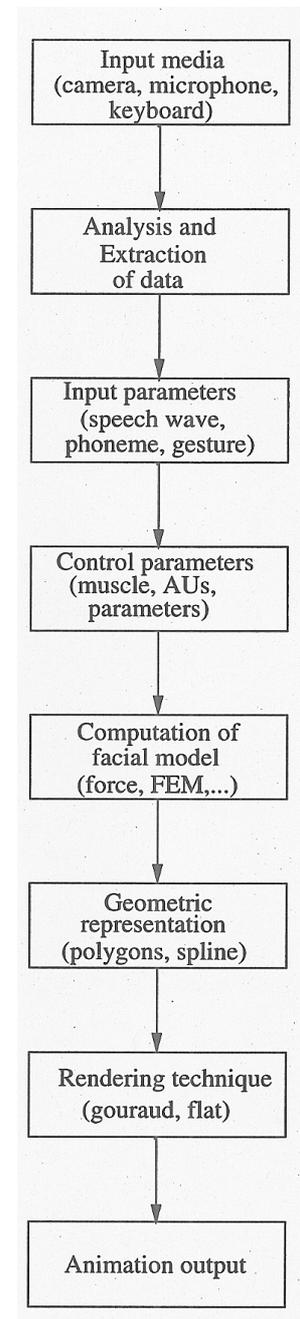
- TH-Systeme modellieren einen synthetischen Agenten, mit dem der Benutzer kommunizieren kann.
- möglichst natürliche Kommunikation gefordert

-> Sprache + Gesichtsmimik

Wie funktioniert ein Talking Head?

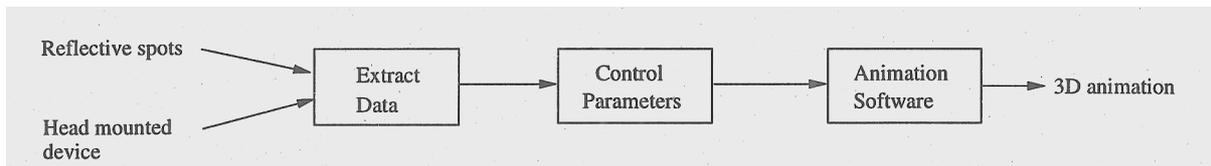
nötig:

- Modellierung des Gesichts
- Modellierung der Lippenbewegungen
- Synchronisation von Audio und Video

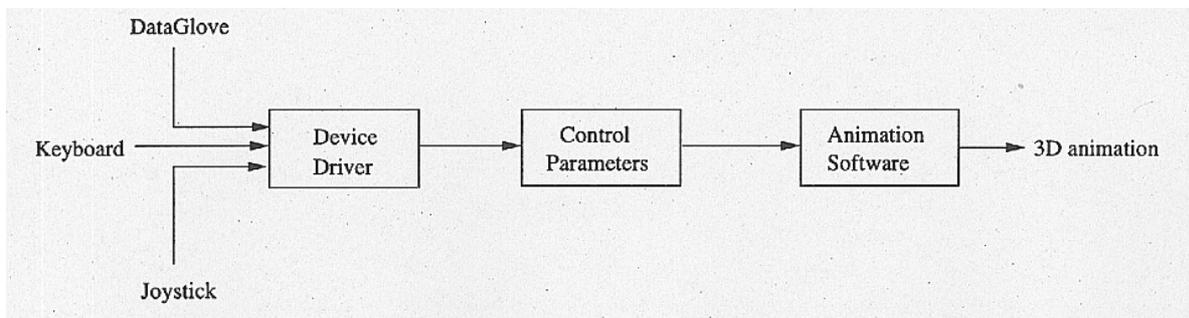


verschiedene Möglichkeiten zur Generierung des Gesichtes:

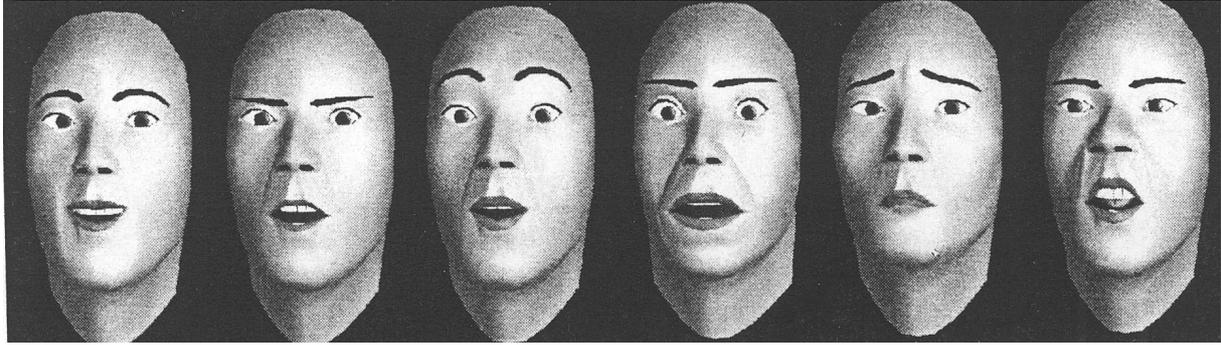
1. Spezifizieren jeder einzelnen Bewegung
2. Motion Capturing



3. mit „Puppenspieler“

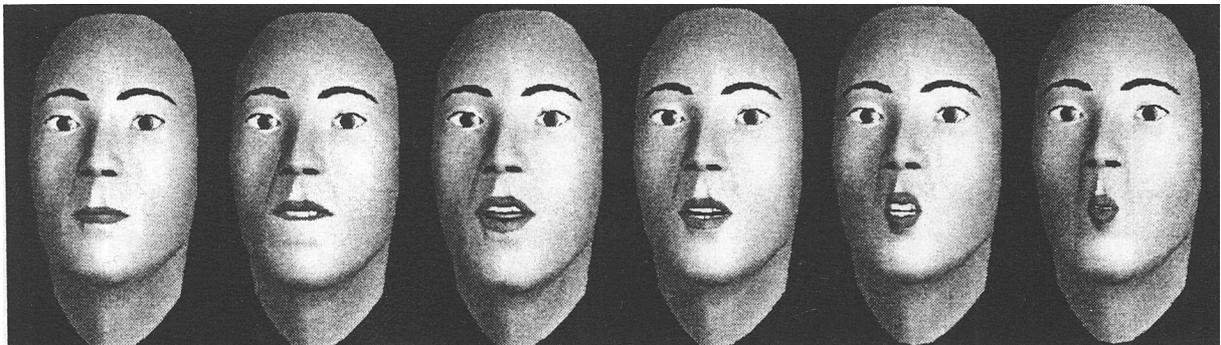


4. Text-to-Speech
5. Conversational Agent



Lippenbewegungen

- cartoonartig
- parameterbasiert
- kontextbasiert
- bildbasiert
- EMG-basiert



Synchronisation von Audio und Video

Menschen sind diesbezüglich sehr empfindlich

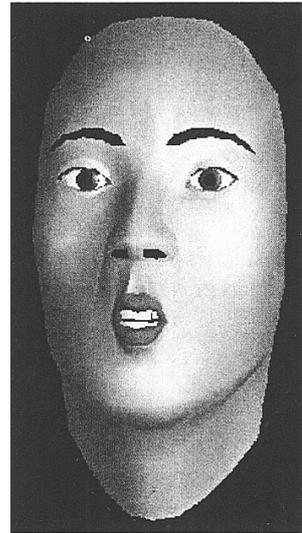
das Auge ist empfindlicher gegenüber Verzögerungen als das Ohr, je nach Schärfe des Tons

Audiokanal wird als Zeitgeber verwendet

Welche Talking Heads gibt es bereits?



August



Baldi



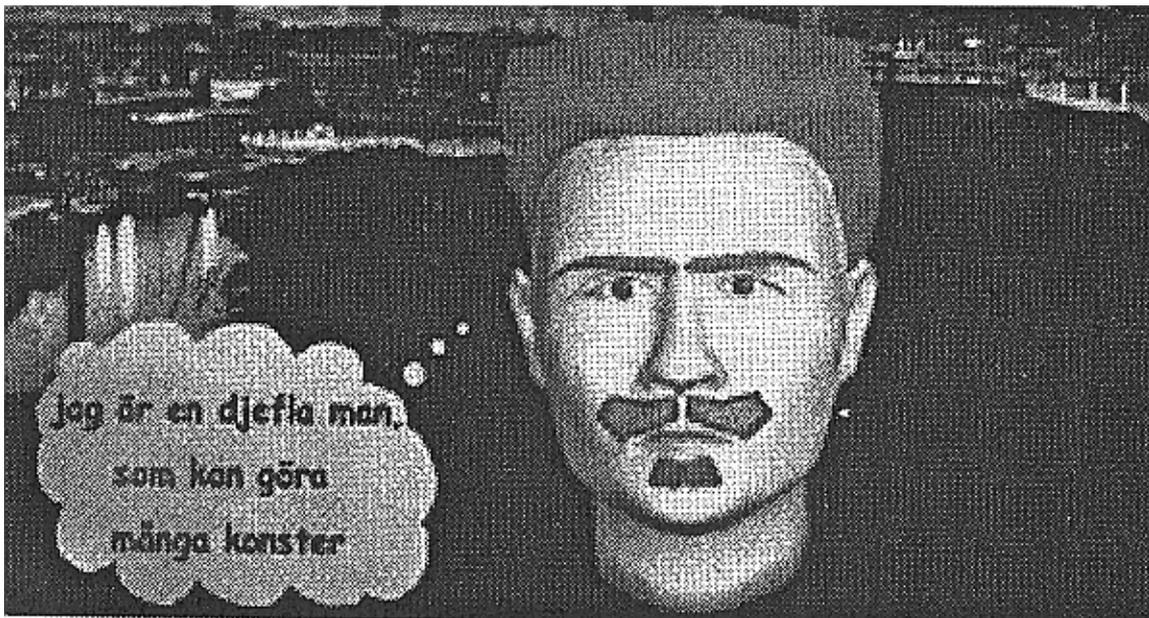
Olga

August

schwedisches multimodales Dialogsystem

kommuniziert mit:

- synthetischer Sprache
- Gesichtsausdrücken
- geschriebenem Text



war für sechs Monate im Einsatz mit grösstenteils unerfahrenen Benutzern

Wie lief die Kommunikation mit August ab?

Es wurde viel Wert auf natürliche Kommunikation gelegt

| User | System |
|---|--|
| What is your name? | → I call myself Strindberg, but I don't really have a surname |
| When were you born? | → Strindberg was born in 1849 |
| What do you do for a living? | → I can answer questions about Strindberg, The Royal Institute of Technology and Stockholm |
| How many people live in Stockholm? | → Over a million people live in the Stockholm area. |
| Do you like it here? | → People who live in glass houses should not throw stones. |
| Yes, that was a smart thing to say | → Hello, I come from the department of Speech, Music and Hearing |
| The Royal Institute of Technology | → The information is shown on the map |
| Thank you | → You are welcome |
| Goodbye | → Perhaps we will meet soon again |
| Yes, it might be that we will | → Strindberg was married three times |

Die Nutzer probierten verschiedene Ansätze aus:

- Verkleinern des eigenen Vokabulars, Benutzung systemeigener Sprachstrukturen
- möglichst kurze und allgemein übliche Äusserungen (circa 4 Worte)
- Wiederholungen

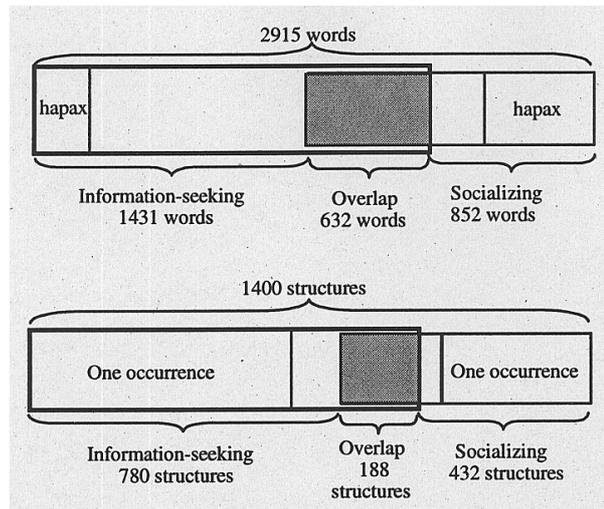
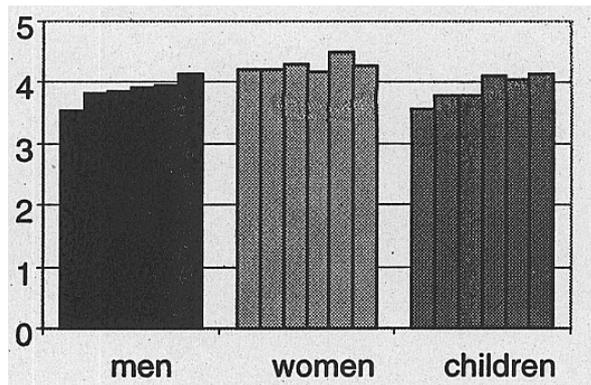
Äusserungen liessen sich aufteilen in:

| soziale | Informations- suchende |
|---------------------------------------|--|
| 15.626 Wörter | 23.604 Wörter |
| „Hallo August!“ „Wie alt bist du?“ | „Wie spät ist es?“ „Wo sind die Restaurants in Stockholm“ |

sozialer Teil wahrscheinlich durch den animierten Agenten beeinflusst

- nur 14% der Wörter nicht unter den 200 häufigsten des KTH-Corpus
- 99 Wörter (oft Schimpfworte) kamen jedoch nicht im KTH-Corpus vor

Äusserungen wurden im Verlauf des Dialogs nicht kürzer, sondern eher länger:



- offensichtlich gab es kaum komplexe Syntaxstrukturen
- Strukturen in sozialer Kategorie waren einheitlicher
- die meisten Äusserungen enthielten nur ein Verb
- komplexe Strukturen fast nur bei Referenzierungen und diese fast nur auf ganze Sätze

Wiederholungen:

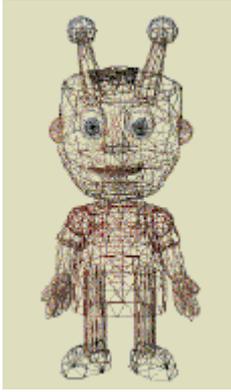
circa 80% reine Wiederholungen bzw
Umstellen der Satzstruktur

11% Reduzierung der Komplexität
9% Erhöhung der Komplexität

Bei mehr als einer Wiederholung :
erst Vereinfachung,
dann Ersetzen von Wörtern

Schlussfolgerungen:

trotz einer grossen Anzahl
verschiedener Kommunikations-
Strategien scheinen nur geringe
lexikalische und syntaktische
Kenntnisse seitens des Talking Heads
nötig zu sein



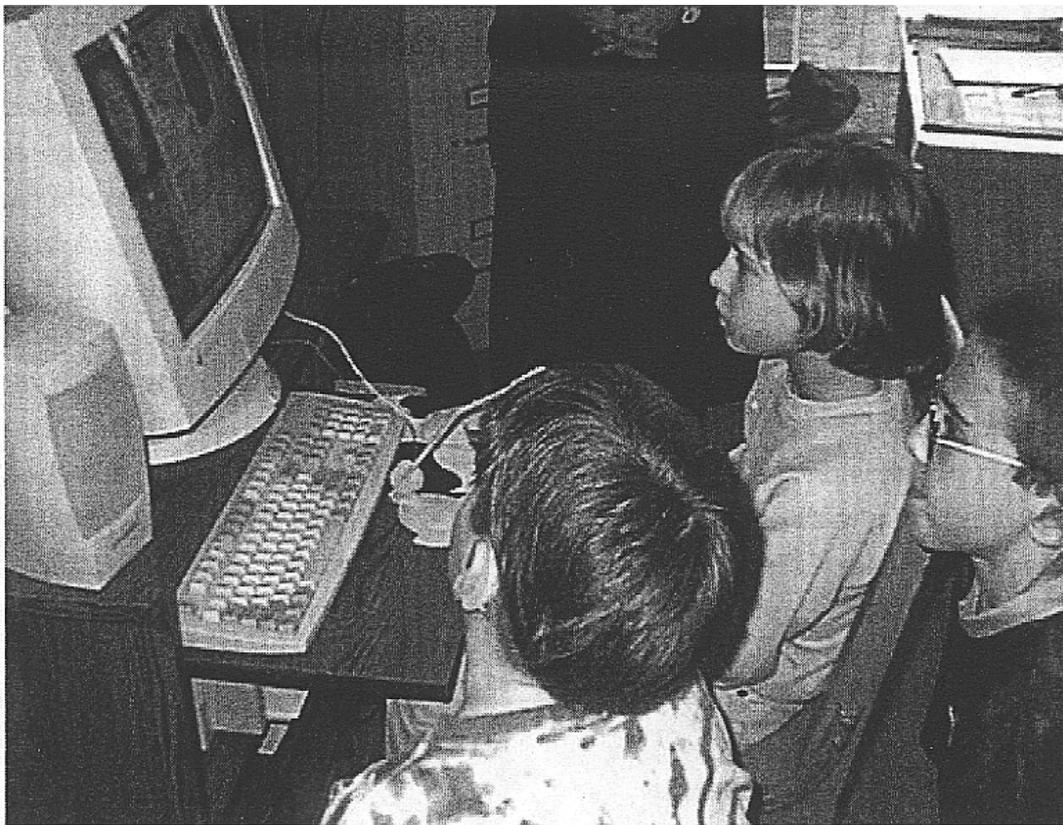
Olga



- dreidimensionale animierte Figur
- kann auf eigene Initiative hin Tips und Ratschläge geben
- hilft bei der Suche und Bearbeitung von Datenbanken

Wozu dient ein solches System?

Baldi wurde für drei Jahre zum Spracherlernen bei tauben Kindern benutzt



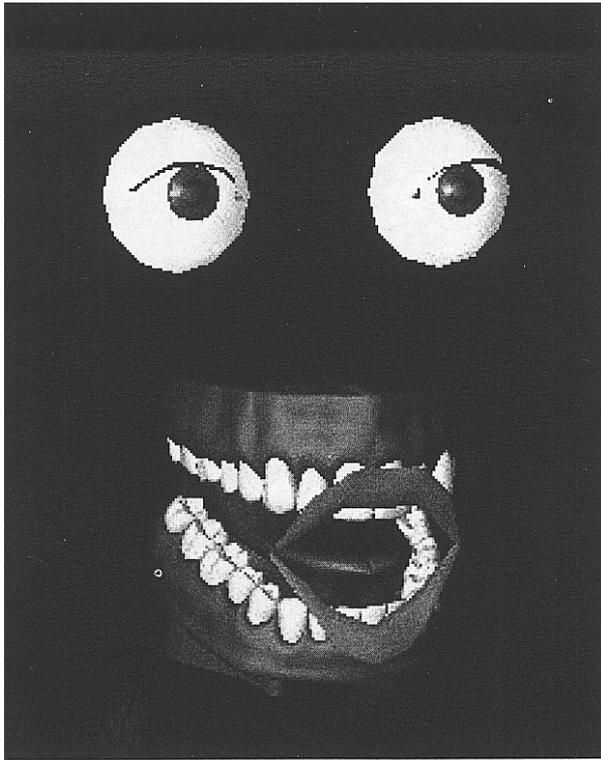
es zeigte sich ein sehr hoher Motivationsgrad

Wie geht es weiter?

August ist derzeit mit anderen Talking Heads in der Öffentlichkeit zugänglich

Baldi bekam neue Modelle für Zunge , Gaumen und Zähne

-> bessere Möglichkeiten zum Lehren von Sprache bei Hörgeschädigten



transparente Haut, ausblendbare Teile
-> z.B. besseres Erkennen der
Zungenbewegungen

Fazit

- Talking Head Systeme erleichtern die Mensch-Maschine-Kommunikation
- erhöhen die Robustheit der Kommunikation durch Multimodalität
- eignen sich zum Erlernen zwischenmenschlicher Kommunikation (auch neue Sprachen)