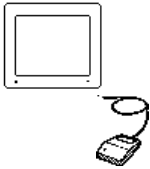
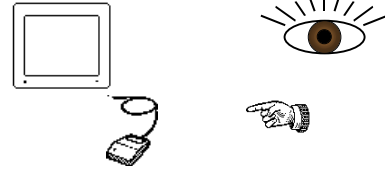


what does a graphical interface  
say about a user?



what does a graphical interface  
say about a user?

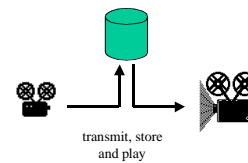


### different media

- not just a load of old text and graphics
- speech and language
- sound - beeps, buzzes, scrapes and music
- handwriting, sketches and gestures
- video, animation, web-cams
- touch, haptic and force-feedback
- bio-sensors

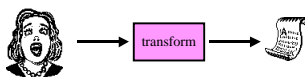
### how much does the computer understand?

- uninterpreted
  - recorded messages, web-cam, multimedia, retrieval



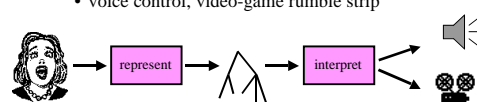
### how much does the computer understand?

- uninterpreted
- interpreted
  - transcription - conversion between media
    - OCR, dictation, 'tidying' freehand drawings. MIDI



### how much does the computer understand?

- uninterpreted
- interpreted
  - transcription - conversion between media
  - semantics - real understanding
    - voice control, video-game rumble strip



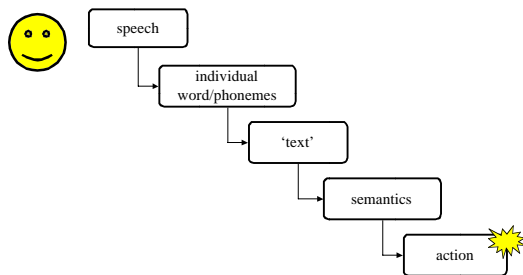
## speech - many levels

- phoneme - smallest units of speech
- allophone - combinations of phonemes
- morpheme - small words or parts of words
- words
- sentences and grammar
- semantics
- conversation

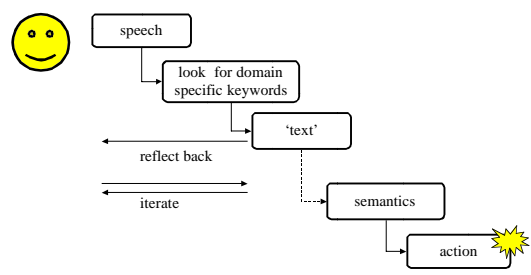
## speech recognition some issues

- phoneme co-articulation
- word separation
- speaker dependence
- small vs. large vocabulary
- different languages ... tonal etc. ...
- homophones - bough / bow
- addressing: accidental commands
- quoting: saying 'stop' in a telegram

## traditional recognition model



## interactive model



## non-speech sounds

- ecological
  - plonk in the bin
  - filling water = file copy
- non-ecological
  - earcons - short note sequences

## networked media

- QoS (quality of service)
  - bandwidth, latency, jitter, reliability, ...
- media
  - video - dropped frames OK
  - speech - missing sound - disaster
- mode
  - real-time - jitter can disrupt conversation
  - streamed - use buffering (latency) to overcome jitter

