chapter 10
universal design

universal design principles - NCSW
- equitable use
- flexibility in use
- simple and intuitive to use
- perceptible information
- tolerance for error
- low physical effort
- size and space for approach and use

Multi-Sensory Systems
- More than one sensory channel in interaction
  - e.g. sounds, text, hypertext, animation, video, gestures, vision
- Used in a range of applications:
  - particularly good for users with special needs, and virtual reality
- Will cover
  - general terminology
  - speech
  - non-speech sounds
  - handwriting
- considering applications as well as principles
Usable Senses

The 5 senses (sight, sound, touch, taste and smell) are used by us every day:
- each is important on its own
- together, they provide a fuller interaction with the natural world

Computers rarely offer such a rich interaction
Can we use all the available senses?
- ideally, yes
- practically – no

We can use • sight • sound • touch (sometimes)
We cannot (yet) use • taste • smell

Multi-modal vs. Multi-media

• Multi-modal systems
  - use more than one sense (or mode) of interaction
    e.g. visual and aural senses: a text processor may speak the words as well as echoing them to the screen
• Multi-media systems
  - use a number of different media to communicate information
    e.g. a computer-based teaching system may use video, animation, text and still images: different media all using the visual mode of interaction; may also use sounds, both speech and non-speech: two more media, now using a different mode

Speech

Human beings have a great and natural mastery of speech
- makes it difficult to appreciate the complexities
  but
- it’s an easy medium for communication
Structure of Speech

- phonemes
  - 40 of them
  - basic atomic units
  - sound slightly different depending on the context they are in; these larger units are...

- allophones
  - all the sounds in the language
  - between 120 and 130 of them
  - these are formed into...

- morphemes
  - smallest unit of language that has meaning.

Speech (cont’d)

Other terminology:

- prosody
  - alteration in tone and quality
  - variations in emphasis, stress, pauses and pitch
  - impart more meaning to sentences.

- co-articulation
  - the effect of context on the sound
  - transforms the phonemes into allophones

- syntax – structure of sentences

- semantics – meaning of sentences

Speech Recognition Problems

- Different people speak differently:
  - accent, intonation, stress, idiom, volume, etc.

- The syntax of semantically similar sentences may vary.

- Background noises can interfere.

- People often "ummm...." and "errr....."

- Words not enough - semantics needed as well

  - requires intelligence to understand a sentence

  - context of the utterance often has to be known

  - also information about the subject and speaker

  e.g. even if “Errr... I, um, don’t like this” is recognised, it is a fairly useless piece of information on its own
The Phonetic Typewriter
- Developed for Finnish (a phonetic language, written as it is said)
- Trained on one speaker, will generalise to others.
- A neural network is trained to cluster together similar sounds, which are then labelled with the corresponding character.
- When recognising speech, the sounds uttered are allocated to the closest corresponding output, and the character for that output is printed.
  - requires large dictionary of minor variations to correct general mechanism
  - noticeably poorer performance on speakers it has not been trained on

The Phonetic Typewriter (ctd)

Speech Recognition: useful?
- Single user or limited vocabulary systems
  - e.g. computer dictation
- Open use, limited vocabulary systems can work satisfactorily
  - e.g. some voice activated telephone systems
- General user, wide vocabulary systems ...
  - still a problem
    - Great potential, however
      - when users hands are already occupied
        - e.g. driving, manufacturing
      - for users with physical disabilities
      - lightweight, mobile devices
Speech Synthesis

The generation of speech

Useful
- natural and familiar way of receiving information

Problems
- similar to recognition: prosody particularly

Additional problems
- intrusive - needs headphones, or creates noise in the workplace
- transient - harder to review and browse

Speech Synthesis: useful?

Successful in certain constrained applications when the user:
- is particularly motivated to overcome problems
- has few alternatives

Examples:
- screen readers
  - read the textual display to the user utilised by visually impaired people
- warning signals
  - spoken information sometimes presented to pilots whose visual and haptic skills are already fully occupied

Non-Speech Sounds

boings, bangs, squeaks, clicks etc.

- commonly used for warnings and alarms
- Evidence to show they are useful
  - fewer typing mistakes with key clicks
  - video games harder without sound
- Language/culture independent, unlike speech
Non-Speech Sounds: useful?

- Dual mode displays:
  - information presented along two different sensory channels
  - redundant presentation of information
  - resolution of ambiguity in one mode through information in another
- Sound good for
  - transient information
  - background status information

E.g. Sound can be used as a redundant mode in the Apple Macintosh; almost any user action (file selection, window active, disk insert, search error, copy complete, etc.) can have a different sound associated with it.

Auditory Icons

- Use natural sounds to represent different types of object or action
- Natural sounds have associated semantics which can be mapped onto similar meanings in the interaction
  - e.g. throwing something away
  - the sound of smashing glass
- Problem: not all things have associated meanings
- Additional information can also be presented:
  - muffled sounds if object is obscured or action is in the background
  - use of stereo allows positional information to be added

SonicFinder for the Macintosh

- Items and actions on the desktop have associated sounds
- Folders have a papery noise
- Moving files – dragging sound
- Copying – a problem ...
  - sound of a liquid being poured into a receptacle
  - rising pitch indicates the progress of the copy
- Big files have louder sound than smaller ones
Earcons

- Synthetic sounds used to convey information
- Structured combinations of notes (motives) represent actions and objects
- Motives combined to provide rich information
  - compound earcons
  - multiple motives combined to make one more complicated earcon

Earcons (ctd)

- family earcons
  similar types of earcons represent similar classes of action or similar objects; the family of "errors" would contain syntax and operating system errors

Smiley face: Earcons easily grouped and refined due to compositional and hierarchical nature

Smiley face: Harder to associate with the interface task since there is no natural mapping

touch

- haptic interaction
  - cutaneous perception
    - tactile sensation; vibrations on the skin
  - kinesthetics
    - movement and position; force feedback
- information on shape, texture, resistance, temperature, comparative spatial factors
- example technologies
  - electronic braille displays
  - force feedback devices e.g. Phantom
  - resistance, texture
Handwriting recognition

Handwriting is another communication mechanism which we are used to in day-to-day life

• Technology
  - Handwriting consists of complex strokes and spaces
  - Captured by digitising tablet
    • strokes transformed to sequence of dots
  - Large tablets available
    • suitable for digitising maps and technical drawings
  - Smaller devices, some incorporating thin screens to display the information
    • PDAs such as Palm Pilot
    • Tablet PCs

Handwriting recognition (ctd)

• Problems
  - Personal differences in letter formation
  - Co-articulation effects

• Breakthroughs:
  - Stroke not just bitmap
  - Special ‘alphabet’ - Graffeti on PalmOS

• Current state:
  - Usable - even without training
  - But many prefer keyboards!

gesture

• Applications
  - Gestural input - e.g. “put that there”
  - Sign language

• Technology
  - Data glove
  - Position sensing devices e.g. MIT Media Room

• Benefits
  - Natural form of interaction - pointing
  - Enhance communication between signing and non-signing users

• Problems
  - User dependent, variable and issues of coarticulation
Users with disabilities

- visual impairment
  - screen readers, SonicFinder
- hearing impairment
  - text communication, gesture, captions
- physical impairment
  - speech I/O, eye gaze, gesture, predictive systems (e.g. Reactive keyboard)
- speech impairment
  - speech synthesis, text communication
- dyslexia
  - speech input, output
- autism
  - communication, education

... plus ...

- age groups
  - older people e.g. disability aids, memory aids, communication tools to prevent social isolation
  - children e.g. appropriate input/output devices, involvement in design process
- cultural differences
  - influence of nationality, generation, gender, race, sexuality, class, religion, political persuasion etc. on interpretation of interface features
  - e.g. interpretation and acceptability of language, cultural symbols, gesture and colour