Chapter 7  SPEECH COMMUNICATIONS

Speech is an information display in auditory form. Sender and/or receiver may be either human or machine.

nature of speech
criteria to evaluate speech communication
components of speech communication and intelligibility
synthetic speech

I. Speech
A) Nature of Speech
1) Production: Diaphragm & Lungs (produce moving column of air)
   - Larynx (voice box and vocal folds)
   - Pharynx (throat)
   - Mouth (tongue, teeth, and lips)
   - Vocal folds vibrate and impart vibrations to moving air column.
   - Three different resonators: pharynx, oral cavity, nasal cavity

2) Phoneme - basic element of speech
   a) phonemes are different across languages
   b) phonemes -> syllables -> words
   c) English:
      - 13 phonemes from vowels
      - 26 phonemes from consonants
      - a couple phonemes from diphthongs

3) Characteristics: Sinusoidal wave and harmonics
   - Complex composite and waveform envelope
   - Depicting Speech (fig 7.1):
     a) Waveform
     b) Spectrum
     c) Spectrogram
   - Frequency composition

4) Intensity
   - Vowels more intense than consonants
   - Males more intense than females by 3 - 5 dB
   - 45 dBa (weak) and 85 dBa (shouting)

B) Criteria for Evaluating Speech
1) Speech Intelligibility: Nonsense syllables, phonetic balance, sentence
2) Speech Quality: Subjective listener preference

C) Component of Speech Communication System:
   1) Speaker (most intelligible vs. least intelligible)
      - longer syllable duration
      - greater intensity
      - More time on sounds, less time on pauses
      - varied fundamental frequencies

Nature of Speech

2) Phoneme - basic element of speech
   a) phonemes are different across languages

B) Criteria for Evaluating Speech
1) Speech Intelligibility: Nonsense syllables, phonetic balance, sentence

Components of Speech System

2) Message
   a) Phoneme Confusion

DVPBGCT FXSH KJA MN

b) Word Characteristics
   1) More familiar words vs. less familiar
   2) Words more intelligible than letters (Alpha, Bravo, etc.)
Components of Speech System: Message

c) Contextual Features (noisy conditions)
1) Small vocabulary
2) Standard sentence construction (always same order)
3) Avoid short words
4) Familiarization training with vocabulary & structure

3) Transmission system
- Intelligibility vs. fidelity
a) Effects of Filtering (Frequency distortion)
   - Low Pass Filter eliminates high frequencies
   - High Pass Filter eliminates low frequencies
   - Band Pass Filter eliminates frequencies above & below
     - Below 600Hz or above 4000Hz - little effect
     - Between 1000-3000Hz - major loss of intelligibility

b) Effects of Amplitude Distortion (non-linear circuitry)
   - Peak Clipping - no major degradation
   - Center clipping - almost total garble

Components of Speech System: Transmission

b) Effects of Amplitude Distortion (non-linear circuitry)
   - Peak Clipping - no major degradation
   - Center clipping - almost total garble

4) Noise Environment
a) Articulation Index (AI)
   - Predicts speech intelligibility given a knowledge of the noise environment.
   - Methodology of weighted sum articulation indices.

b) Preferred Octave Speech Interference Level (PSIL)
   - Rough estimate of noise effects on speech reception
   - Numeric average of noise levels in 3 bands centered a 500Hz, 1000Hz, 2000Hz.

c) Preferred Noise Criteria Curve (PNC)
   - Noise spectrum plotted against "standard" curve.

d) Reverberation - Reflected (echoed) sound interference.

Components of Speech System: Noise

5) Hearer
a) Hearing ability
   - age
b) Hearing protection
b) Attentiveness
   - Familiarity

Components of Speech System
II. SYNTHESIZED SPEECH

Human Factors Considerations:
1. Determine most appropriate uses.
3. System improvements

A) Types:
1) Analog recordings
   - Mechanical complexities
   - Only pre-recorded messages
   - Time-to-access
2) Digitized Speech
   - Memory Requirements
     - (8-24 Kbyte/sec, 1Mbyte = 40 sec)
   - Fast access (can also be parsed)

B) Methods of Synthesized Speech
1) Analysis-Synthesis
   - Electronic Model (Synthesizer Keyboard)
   - Filters, Modulators, Envelope Generators
   - Requires much less memory
   - Previously analyzed, encoded & stored sounds
   - Co-articulation problem
     (bookcase-book Kase)

2) Synthesis-by-Rule
   - Reproduces phonemes of the language
   - Translates typed text, apply rules, produce sounds
   - Control characteristics:
     - Natural/robot, male/female
     - Speed, frequency, inflection, prosodics
   - English more difficult because of spelling rules

C) Uses of Synthesized Speech

D) Human Performance
1) Intelligibility - Variable (simple words, high S/N, Intelligibility = 99%)
2) Remembering
   - May require more processing capability.
   - Encoding difficulty may disrupt working memory
     - as well as transfer to long-term memory.

E) Guidelines for use of synthesized speech
1) Voice warnings should be qualitatively different
2) If used exclusively for warnings, no pre-alerting
3) If multiple uses, attention direction may be appropriate
4) Maximize intelligibility
5) For GP use, maximize user acceptance via natural sound
SYNTHESIZED SPEECH : Guidelines

6) Replay option

7) Interrupt capability

8) Spelling mode requires higher quality

9) Introductory/familiarization/training message

10) Use sparingly - where appropriate and accepted