BRAIN & COGNITION WEEK 9- speech production and speech perception

SPEECH PRODUCTION

'**Suprasegmental'** features of speech (prosody)

- *Intonation*, *stress/accent*, *tone*, *pauses*, *rate*, *duration*

Phonemes as the basic building blocks

PROSODY refers to suprasegmental speech properties… as above. One way to appreciate prosody is to listen to sentences where the prosody is not quite right e.g. individual durations of certain words are not what is expected.

INTONATION is a change in pitch over time e.g. falling in pitch = statement, rising in pitch = question.

STRESS conveys lexical information- stress or accent refers to the length, loudness or pitch of a syllable. E.g. record an album then release the record…

TONE or accentual language e.g. Japanese, certain syllables are accented or unaccented- and this changes the meaning of the word. Tone language e.g. mandarin, each syllable is lexically marked with one of four lexical tones which all have distinctive pitch contours – altering the pitch contour may change the lexical meaning of a word and sentence.

PAUSE- normal speech has no pause between words- it flows. Strong 'stress' signal 'don't… do that… again!' – alters meaning

RATE- rate is the average number of speech segments per unit time (average 4 syllables per sec)- a faster rate may signal anxiety, slower rate may signal depression.

DURATION- relative duration of segments important semantic cue, key words tend to be pronounced carefully, fillers are not. E.g. 'this is *actually* true' vs 'this is true, *actually'*.

PHONEMES- are distinctive sounds in a language that permit contrasts- e.g. minimal pairs – RAT, HAT, PAT… if you change the phoneme, you change the word. BUT these are abstract representations- categories, NOT sounds- phones.

PHONES & ALLOPHONES- actual sounds are called phones e.g. pin, spin. If you say spin like 'sbin' the two sounds are different (allophones) but swapping them doesn't make it a different word

Conversely a particular **MORPHEME** (unit of written language) can be pronounced differently depending on the speech

HOW ARE CONSONANTS PRODUCED***

3 TYPES OF ARTICULATION

- Bilabial- using both lips
- Labiodental- lip against teeth
- Interdental- between the teeth

Partial obstruction causes turbulence- **FRICATIVES**- airflow partially stopped to produce a rushing sound e.g. the 'fr' in fricative.

SPEECH PERCEPTION

FREQUENCY= number of vibrations per second.

Imagine you put a stick in to water and measure the number of waves that go past the stick each second once you've dropped a stone into the water…

- The number of waves that go past each second= the frequency, a measure of how fast the water is moving.
- The height of the wave is the amplitude, a measure of size… bigger stone = higher waves.

FREQUENCY= how frequent amplitude changes

AMPLITUDE= a measure of size/ how loud the sound is

Any non-sinusoidal waveforms (e.g. speech) can be represented as a collection of sinusoidal waves of different frequencies blended together. The technique of transforming a complex waveform into its sinusoidal components is called FOURIER ANALYSIS.

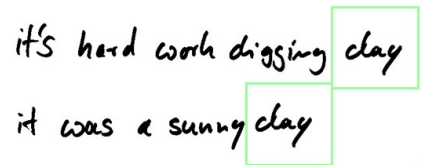
Dark sounds in a spectrogram show the frequencies at which acoustic energy is concentrated i.e. a dark band= It's of LOUD sound at that frequency. These bands are called **formant frequencies**. There may be as many as 4/5 formants, but vowels are usually identifiable by the first two.

PERCIEVING CONSONANTS IN ISOLATION- vowels can be perceived in isolation based on their formant frequencies, but for consonants this is almost impossible because they're produced briefly in bursts of acoustic energy.

PARALLEL TRANSMISSION- most of info about a consonant is encoded in the vowel that follows- this is problematic for a phoneme-based view of speech perception.

CATEGORICAL PERCEPTION- excellent discrimination across boundary OR poor discrimination within boundary.

CONTEXT AFFECTS PERCEPTION-



It's the same for speech- you hear what you expect to hear even if there is something obstructing the full word.

E.g. **phonemic restoration effect**- WARREN & WARREN

E.g. **the auditory vocal channel**- MCGURK