

How do we understand spoken words and continuous speech?

- Speech is fleeting
- Speech flows, continuous – “segmentation” problem
- Speech varies – “invariance” problem
 - Speaker differences
 - Contextual differences, coarticulation
- Humans are still very good at recognizing speech; cannot design robots/computers that recognize speech like humans
 - Perceptual constancy

Segmentation

- Processes in speech segmentation are affected by the ambient language
- Various processes and cues are used to segment speech
- Word recognition in continuous speech – “The word boundary problem” (Cutler)
 - “loose analogy” or “Luce and Allergy”
 - “How big is it?” or “How bigoted?”
 - LMNO
- Stress-based segmentation for English
 - Assumption that strong stress syllables are word initial
- But other processing strategies for other languages (e.g., syllable-based segmentation in French)

Spoken word recognition

- Do we hear each sound? Sound-by-sound processing?
- Each syllable? Is the syllable the fundamental unit?
- Speech or Sound Processing:
 - Prelexical (or phonetic) code
 - Direct analysis of the perceptual information
 - Postlexical (or phonemic) code
 - Identification associated with higher level processes, word recognition or contextual effects
- Research does not support the idea that this dual-code model is operating in parallel.
- Instead, phoneme monitoring tasks are predominantly prelexical, with occasional postlexical support.

Time course for Word Recognition

- Uniqueness point
 - Point at which the sequence for the word does not overlap with any other
- Isolation point
 - Point when we recognize a word in strongly biasing contexts even if it is not really unique yet, so can be before uniqueness
 - May lack confidence
 - Will continue to monitor for correctness
- Recognition point
 - Defined by when listener is confident about word they've heard
- Lexical Access
 - Point when all aspects of the word are available – phonological, semantic, syntactic, pragmatic – all are assisting in the word recognition
 - Next step is one of integration, this is the first step in comprehension

Context Effects in Word Recog

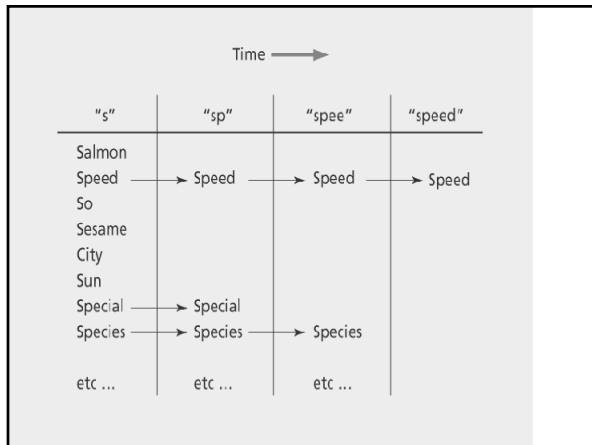
- Controversial, Two possibilities:
- Speech Processing is Autonomous:
 - Word recognition happens independent of context effects
 - Context effects are postlexical or postaccess, so only contribute toward integration and comprehension
- Interactionist perspective:
 - Feedback from higher levels can affect lower levels including phoneme perception

Early Speech Recognition Models

- Template matching
 - Sound targets are stored as templates and searching results in phoneme match
 - PROBLEM: If templates are precise then this model cannot explain speech recognition because of phoneme variation
- Analysis-by-synthesis
 - Model suggests that we synthesize or produce the speech sounds we hear in order to create a match or analyze the sounds.
 - So, speech production assists speech recognition
 - The Motor Theory of Speech Perception (Lieberman)
 - We model the speech movements of the speaker.

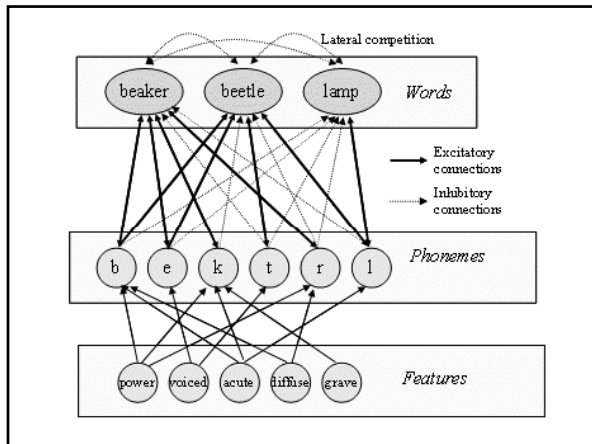
Cohort Model

- Marslen-Wilson & Welsh, 1978; Revised, Marslen-Wilson, 1989
- Bottom-up priority model
- Takes its name from the idea that we develop a “cohort” of possible words to fit the target
- All candidates considered in parallel
- Then we eliminate items from the cohort until we land on the match, see Figure and experimental findings.
- Uniqueness point
- Context is not used in early stages of recognition
 - Need cohorts to remain active for possible errors



TRACE

- McClelland & Elman, 1986
- Connectionist model
- Interactive – not bottom-up or autonomous
- Like the Interactive Activation model of printed word recognition (Yates lecture and chapter 6 in textbook)
- Three levels: Features, Phonemes, Words
- Activation is excitatory and inhibitory
- Main problem with TRACE is context is prelexical allowing top-down effects on phoneme recognition but research does not tend to support this.



Cross-Modal Effect

- Other cues/information in speech processing
- E.g., Visual cues, lip reading
- McGurk Effect
 - [You tube demo](#)
 - [You tube demo](#)
