Models of Word Recognition

I. Stages of Spoken Word Recognition

A Initial contact

- First contact with the lexicon after hearing some speech
- Precise form of contact: depends on theories of speech recognition
    1. LAFS (lexical access from spectra, Klatt): spectrographic
    2. motor theory (analysis-by-synthesis): articulatory gestures
    3. phonemic
- Causes certain lexical representations to "activate"

B. Lexical selection

- Activation continues to accumulate (or, die off)
- Type of activation depends on the model being considered
    1. all-or-nothing activation
    2. better fit leads to higher degrees of activation
    3. may be affected by properties of the words, such as frequency
- Pool of possible candidates being considered therefore changes over time

C. Word recognition

- End point of the selection phase
- Is reached when only a single candidate remains in the pool
- "Uniqueness point"—point in a word at which is can be uniquely identified
    1. null scenario: word recognition occurs at uniqueness point
    2. heavily-biasing contexts: word may be identified before uniqueness point
    3. in some models, certain words might not be recognized until after the uniqueness point

D. Lexical Access & Integration

- Lexical access: Point at which lexically-stored information (phonological, morphological, semantic, etc.) becomes available
- Integration: Working the meaning of the word into the overall meaning of the sentence

II. Models of Word Recognition

A. Context and Word Recognition

- Context: Anything other than what's actually present in the immediate speech stream; includes prior sensory input (previous context) as well as higher knowledge sources (semantic, syntactic, etc.).

1. Autonomous models: Context can affect integration of a word's meaning, but can have no effect on stages prior to word recognition. (bottom-up)
2. Interactive models: Context can affect the stages prior to word recognition (top-down)

- Non-structural context: contextual information from within the same level of processing; does not require top-down processing flow from higher levels. Ex: semantic priming of "university" by "professor."

- Structural context: contextual information from a higher level of processing; requires top-down processing flow. Ex: (speech recognition) phoneme restoration; (word recognition) next word in "I saw the …" must be an adj. or noun (does this affect word recognition?)

  1. Semantic: Words appropriate for a given semantic context are recognized (i.e., responded to) more quickly than inappropriate or neutral words. Question: because of quicker processing, or because of post-perceptual factors?
  2. Interpretive: Refers to use of pragmatic, discourse, or real-world knowledge in word recognition. "She went to the polls and cast a vote for N_____."

☞ Problems with context-based arguments: Finding an experimental paradigm that can differentiate between "real" context effects and post-perceptual effects.

B. Cohort model (Marslen-Wilson)

- Cohort: (gen.) A group of things that share some characteristic (like age)
  (psycholinguistics) A group of words that are in a common candidate set during lexical selection; usu. defined in a L to R manner
- Ex: rabbit [ræbɪt]: Cohort 1: all words beginning with [r]
  Cohort 2: all words beginning with [ræ]
  Cohort 3: all words beginning with [ræb]
  Cohort 4: all words beginning with [ræbɪ]
  Cohort 5: all words beginning with [ræbɪt]
- No. of Cohorts determined by the words uniqueness point
- Evidence provided in support of Cohort:
  1. Shadowing: Speech to shadow had mispronunciations
     Mispronunciations varied according to:
     Number of features changes (= amount of distortion)
     Location (1^{st}, 2^{nd}, 3^{rd} syllable) (= number of candidates)
     Contextual clues to the correct word (appropriateness)
     Speakers are supposed to mimic the errors
     Fluent restorations about 50% of the time
     Fluent restorations more likely with slight distortion, with words highly predictable from context, and when distortion is in final syllable

2. Mispronunciation monitoring
   higher accuracy rates for word-initial mispronunciations
3. Gating: subjects presented with increasingly more and more of a word
   Words without context: need about 333 ms. to identify
   Word in context: need about 199 ms. to identify
   Claimed to show that context does not affect candidate generation, since candidates are freely generated up till the point of recognition

C. TRACE model (McClelland, Elman, Rumelhart)

- Structured as a connectionist model: a set of interconnected nodes
- 1st layer: feature nodes; 2nd layer: phoneme nodes: 3rd layer: word nodes
- Nodes in the same layer: inhibitory connections (competition)
- Nodes in different layers: bidirectional connections (bottom-up and top-down)
- Advantages of the TRACE model:
  Does not require knowing where the left edge of the word is
  Can "recover" if a given segment (even the first one) is missed
  Bidirectional connections account for phoneme restoration & other contextual effects on speech recognition
- Possible disadvantage: Predicts a lot of top-down information flow; should permeate all levels of word recognition.
- Current research: Is this a disadvantage, or are there, in fact, lots of context effects in all levels of word recognition?

BCS 260: Psychology of Language
Dec. 8, 2000

Study Questions: Final Exam

Six of the following questions will be chosen to appear on the question sheet for the final exam. Each student will choose three questions from that set to answer. In addition, there will be a set of review questions from sections 1 and 2—these questions will be drawn from the previous study question sheets. Paper will be provided during the midterm, students need bring only some writing implement.

1. Explain the concept of constituency, and why a grammar that employs constituency is superior in accounting for syntactic competence than a grammar without constituency. Is there any evidence for constituency in performance?
2. Explain what the idea of transformations, and how they relate to the idea of dual-structure syntax. Explain why we use transformations, instead of just making more phrase structure rules. What aspects of a person's syntactic competence are best explained using the concept of transformations?
3. What is the derivational theory of complexity, and what sorts of predictions are made by this theory? Review some of the experimental findings that are pertinent to this theory, and explain why they do/don't support it.
4. What does the term "garden-path phenomenon" refer to, and how does it relate to the idea of syntactic ambiguity? Explain the link between garden path phenomena and parsing heuristics. Give examples of at least three heuristics, providing cases that exemplify how the heuristic is supposed to work.
5. How does the idea of long-distance dependency relate to the idea of transformations? Is there any evidence that long-distance dependencies are psychologically real? Give both clinical and experimental evidence on this topic.
6. Explain the difference between serial and parallel search models of syntactic parsing, and go over whether there is evidence from conceptual/pragmatic and discourse context that favors one or the other.
7. How do the ideas of segmentation and invariance relate to speech recognition?
8. Many psychologists debate whether speech processing is strictly bottom-up, or whether it also contains the possibility to top-down effects. Explain the difference between top-down and bottom-up processing. How do contextual effects on speech perception fit into this debate? Give at least two examples of context effects on speech recognition, explaining each.
9. Give an overview of the Cohort model of word recognition. Provide at leave 2 pieces of experimental evidence that have been offerred in support of this model. Be sure to explain how the experimental findings support the theory.
10. Give an overview of the TRACE model of word recogntion. Explain at least two of the advantages claimed to follow from the TRACE model. How does this model fit in to the debate between autonomous and interactive models for word recognition?