Introduction to Psycholinguistics

Lecture 7

Lexical Processing - Part I

Pia Knoeferle & M. W. Crocker

Department of Computational Linguistics

Saarland University

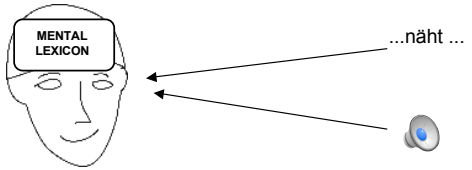
SS 2006

Overview

- Research issues in lexical processing (comprehension)
- Lexical processing: spoken vs. visual word recognition
- Stages of lexical processing
- Serial vs. parallel
- Factors that influence word identification
- Competition
- Some research methods in spoken word recognition
  - Perceptual identification
  - Shadowing
  - Lexical decision
  - Eye tracking: a recent method

Some questions

- When we see or hear a word
  - How do we access its representation within the lexicon
    - In the face of various kinds of ambiguity and noise
  - How do we know whether an item is stored there?
    - *Word* vs. *nonword* (e.g., *lisen*=pronouncable pseudoword, *lrtij*=nonword, *lesen*=word)
  - What are the differences between understanding spoken and visually presented words?
  - Which factors can influence the speed of identifying a word?



Definitions and characteristics

- Spoken/visual word recognition
  - Interface between speech/visual perception & higher levels of cognitive processing
  - Translating acoustic/visual signals into mental representations from long-term memory
- Some characteristics
  - Happens very fast
    - Lexicon of average listener contains around 65.000 words from which to choose during comprehension in real-time
    - Selection of the appropriate word as early as 250 ms into a word
  - Robustness: Errors are rare
    - In a corpus of 200.000 words, 86 lexical errors (< 1 in 2000 words)

Spoken versus visual word recognition

- Early models of word recognition typically were based on insights from studies on *visual* word recognition
  - ➪ Visual word recognition models: often template-matching
- Many models of SWR based on models of visual word recognition
- But: Some differences between written and spoken language

Written text	Speech
Distributed over space	Unfolds over time
Re-fixation of previous words	Re-fixation is not possible
Constant signal	Acoustic-phonetic variation
Linearity: successive sounds represented by strings of letters	Phonemes overlap and are co-articulated

Spoken versus visual word recognition

- Translating from a model of visual word recognition to SWR

> Imagine a row of Easter eggs carried along a moving belt; the eggs are of various sizes, and variously colored, but not boiled. At a certain pint, the belt carries the row of eggs between the two rollers of a wringer, which quite effectively smash them and rub them more or less into each other. The flow of eggs before the wringer represents the series of impulses from the phoneme source; the mess that emerges from the wringer represents the output of the speech transmitter. (Hockett, 1955, p. 216)

  - ➪ Account of SWR must model information integration over time
  - ➪ Transient nature of signal must be taken into account
  - ➪ Invariant mapping of acoustic features to phonemes difficult
    - E.g., co-articulation information would be lost
    - Acoustic-phonetic variation difficult for a template-matching account
- In today's lecture we will focus on spoken word recognition

Stages of lexical processing

- Stages of lexical processing
  - ➪ Identification: Initial contact, lexical selection, word recognition
  - ➪ Lexical access and integration
- *Initial contact* with the lexicon after processing speech input
  - ➪ Theories differ in their assumptions regarding the form of representation that makes contact with the lexicon
- Form of representation
  - ➪ Temporally-defined spectral templates
    - Lexical access from spectra (LAFS) model (Klatt, 1989)
      - ➪ Frequency with which air particles vibrate plus intensity/loudness in a sound wave form pattern that matches items in the lexicon
    - Problems: inter-speaker variability
  - ➪ Motor representation in analysis-by-synthesis models (Halle & Stevens, 1962; Stevens, 1960)
    - Recognize speech by the actions necessary to produce the sound

Stages of lexical processing

- ➪ Motor representation cont.
  - One variant of this model: Motor theory
    - ➪ Extract articulatory gestures ((i.e., lip rounding, tongue position)
    - ➪ Listener models motor movements of the speaker
    - ➪ E.g., all /d/s are made by closing tongue against alveolar ridge
  - Pros
    - ➪ Copes with speaker differences (listeners generate their own candidates)
  - Note: specification of motor movements must be rather abstract
    - ➪ Mute people understand speech & we understand speech that we cannot ourselves produce
- ➪ Abstract units such as phonemes (e.g., Pisoni & Luce, 1987) or syllables (Mehler, 1981)

Stages of lexical processing

- *Lexical selection*
  - Lexical entries that match the representation are “activated”
  - Activation increases/decreases until one lexical entry is selected
  - Depending on the model, the degree of activation is
    - All-or-none
    - Based on word properties such as word frequency or goodness of fit with sensory data
  - Candidate set changes over time
- *Word recognition*
  - End point of selection phase when only one candidate remains
  - Competition process
- *Lexical access*: phonological, semantic, syntactic, and pragmatic information becomes available
- *Integration*: Integrating the word into the sentence context

Factors influencing word identification

- We’ll now consider in more detail the factors influencing lexical access time and the competition process
- Factors influencing the time/accuracy of lexical access
  - E.g., Word length, word frequency, lexical similarity, uniqueness point, semantic priming
  - We review word frequency and lexical similarity here
- Word frequency
  - High vs. low frequency words
  - High frequency words recognized faster and more accurately
    - Savin (1963): high and low frequency words presented in white noise (noise from combining sounds of all frequencies) for perceptual identification
      - HF words recognized at lower signal-to-noise ratios than LF words
      - Errors: higher-frequency words named rather than the target words
    - HF words required less acoustic phonetic info for recognition than LF words in a gating task (Luce et al., 1984)

Factors influencing word identification

- Lexical similarity
  - Measure for lexical similarity in visual word recognition
    - N-metric
      - Two words are visual neighbours if they differ by only one letter
      - Examples: sand, wand (attention: only visual neighbours)
    - Spoken word recognition: Variant of the N-metric
      - Two words are neighbours if they differ by only one phoneme
      - Examples: vote and vogue
    - Dense vs. sparse neighbourhoods
      - E.g., sell has many neighbours (tell, well, bell, sill, till)
  - Neighbourhood size & frequency of words in the neighbourhood affect recognition (above effects of frequency of target word)
    - If a word is phonetically similar to few and/or rare other words
      - Easier recognition than for words with many similar and/or frequent other words

(Luce, 1986; Luce & Pisoni, 1998)

      - High frequency words with few, low-frequency neighbours are most easily recognized

Serial vs. parallel

- Does word recognition take place serially or parallel?
  - Serial models account well for frequency effects
    - Search of words in frequency-ordered way, self-terminating search
  - But have problems accounting for speed of lexical processing
    - Recall: recognition can be as fast as 200-250 ms
    - Number of searched items must be limited
  - Parallel search
    - Most models today agree that there is some amount of parallel competition
    - Unlimited vs. limited capacity
    - How competition is realized depends on individual models

Competition

- Findings on effects of neighbourhood size & the frequency of neighbours
  - Recognition does not solely depend on a match between spoken input and a lexical representation
  - Recognition also depends on the degree to which the input matches representations of alternative words
- If many word candidates match the incoming speech signal
  - Strong competition and slower recognition process
  - Parallel activation of candidates matching in onset (*candle/candy*) or any other part of speech input (rhyme: *speaker/beaker*)

Some ways how competition could take place

(1)

Input	Target	Competitors
/s/	*speed*	*single, sit,spacious, speech, spray, speak, ...*
/sp/	*speed*	*spacious, speech, spray, speak …*
/spi:/	*speed*	*speech, speak*
/spi:d/	*speed*	*speed*

(2)

Input	Target	Competitors
/s/	*speed*	*single, sit,spacious, speech, spray, speak, ...*
/sp/	*speed*	*spacious, speech, spray, speak, pain, peel ...*
/spi:/	*speed*	*speech, speak, peel*
/spi:d/	*speed*	*speed*

Uniqueness point

- Point where a word becomes unambiguous (also *identification point*)
  - ≠ recognition point
  - Examples
    - *cheapness, cheap, cheaper/st: cheap**n**ess*
    - *meeting, meet, meets: meet**i**ng*
    - *Elefant: Ele**f**ant*
  - How to discover the uniqueness point
    - Look up words and their pronunciations in a dictionary
    - Gating paradigm
      - Playing incomplete words to listeners (i.e., with the ending cut off), and measure people's guesses about word identity
  - Faster lexical access for words with an earlier compared to later uniqueness point

Methods in spoken word recognition

- Now we have reviewed some findings. Let's consider the methods used for examine lexical processing
- Perceptual identification
  - Presentation of a (degraded) stimulus to participants
  - Task: identify the stimulus, and respond with a word (open-set: any word; closed-set: response alternatives given prior to a trial)
- Pros
  - Suitable for examining structural relationships among words in the mental lexicon (sensitive to frequency and lexical similarity effects)
- Cons
  - Post-perceptual measure (i.e., end product of lexical processing)
  - Use of a degraded signal
    - Might lead to guessing strategies (not reflect normal processing)
    - Measure of decision rather than recognition processes

Methods in spoken word recognition

- Shadowing and mispronunciation detection
- People repeat spoken utterances in near synchrony with a speaker
  - Marslen-Wilson (1985)
    - Both error rates and shadowing latencies increased the more syntactic semantic, and lexical information in the repeated utterance were anomalous
    - Shadowers do not simply repeat but analyze the utterance
      - Actively engage in syntactic and semantic analysis of the input during shadowing
  - Problems
    - High inter-individual variation in participants
      - Close shadowers (lag of 250 to 300 ms)
      - Distant shadowers (lag of > 500 ms)

Methods in spoken word recognition

- Lexical decision
  - Presentation of isolated words
  - Task: Classify the stimuli as *words* or *non-words*
    - Spoken version: *words* vs. *pronounceable non-words*
    - Written version: *words, non-words, pronounceable non-words*
  - Measure: response latency and accuracy

cotch

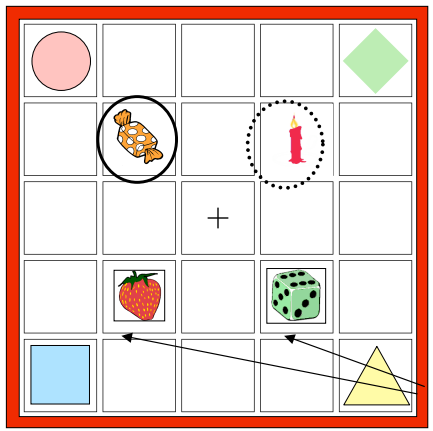
Stimuli	Type	Mean decision time
child	word	708 ms
tree	word	703 ms
csrt	non-word	644 ms
cotch	non-word (pronounceable)	746 ms

  - Response latencies for non-words are faster than words/pronounceable non-words
    - Purely form-based detection of totally illegal words

Eye-tracking spoken word recognition

- Lexical decision
  - "Online" paradigm
  - But reaction times are measured after a word has been presented
    - Do not reveal the time course of processing
- Eye-tracking
  - Enables us to investigate processes during word recognition

Example Tanenhaus (1995)



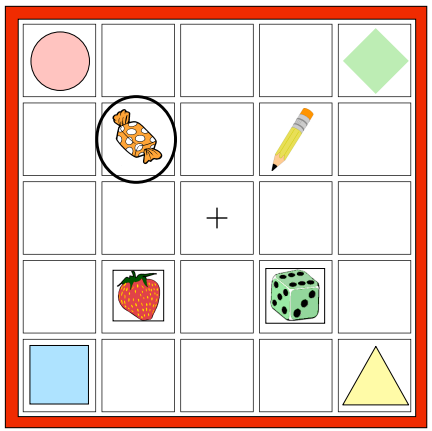
Click on the candy.

target: candy

competitor: candle

distractors: strawberry, dice

Example Tanenhaus (1995)

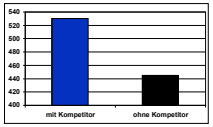


Click on the candy.

target: candy

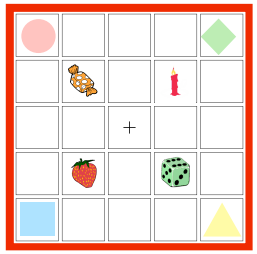
Results Tanenhaus

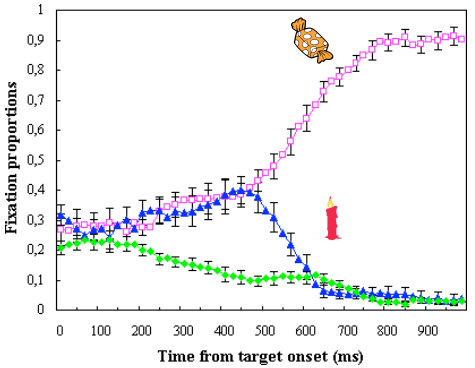
- Method: video-based eye tracker (33 ms sampling rate)
- Analysis of gaze data from target word onset until offset
  - Onset time of first saccade to the target with versus without a competitor



- Findings of a tight time-lock between
  - Eye movements and spoken utterance comprehension
    - Makes it possible to use eye movements to examine comprehension process online
  - Speed of word recognition
    - Influenced by the names of possible referents in the display
      - Incremental interpretation of speech signal with visual information
  - Retrieving lexical information
    - Begins prior to word offset (ca. 200 ms to launch a programmed eye movement)

Fixations over time





Summary

- Spoken vs. written word recognition and implications for modeling
- Stages of lexical processing
  - Identification: Initial contact, lexical selection, word recognition
  - Lexical access and integration
- Factors affecting lexical access
  - E.g., Word frequency and lexical similarity
  - Competition process
- Methods used to study spoken word recognition