PSY 369: Psycholinguistics

Language Comprehension

Word recognition & speech recognition

Lexical access

- How do we retrieve the linguistic information from Long-term memory?
  - How is the information organized/stored?
  - What factors are involved in retrieving information from the lexicon?
  - Models of lexical access

Lexical access

- How do we retrieve the linguistic information from Long-term memory?
  - How is the information organized/stored?
  - What factors are involved in retrieving information from the lexicon?
  - Models of lexical access

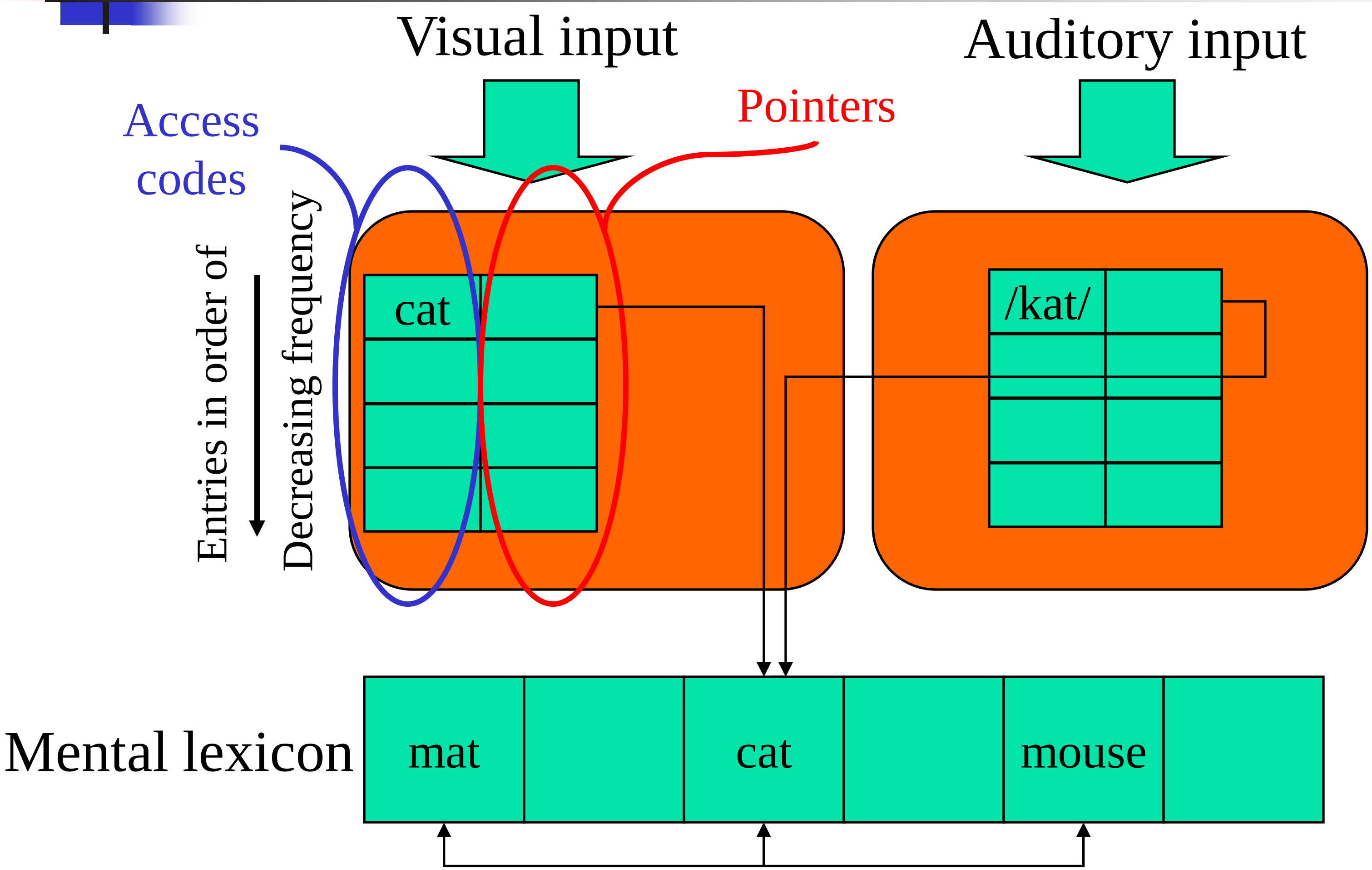
Models of lexical access

- Serial comparison models
  - Search model (Forster, 1976, 1979, 1987, 1989)
- Parallel comparison models
  - Logogen model (Morton, 1969)
  - Cohort model (Marslen-Wilson, 1987, 1990)
- Connectionist models
  - Interactive Activation Model (McClelland and Rumelhart, 1981)

Search model (e.g., Forster, 1976)

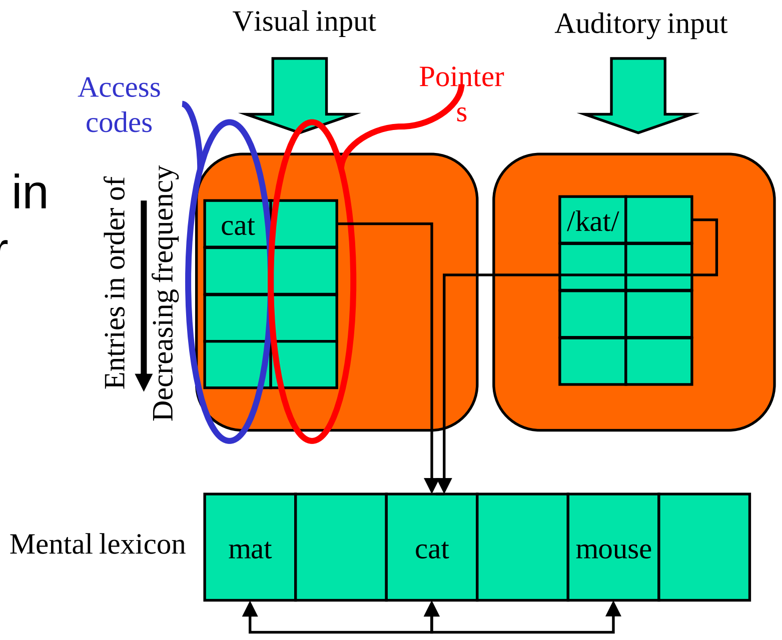
- Access of the lexicon is considered autonomous, independent of other systems involved in processing language
  - A complete perceptual representation of the perceived stimulus is constructed
  - The representation is compared with representations in *access files*
    - Three access files:
      - Orthographic
      - Phonological
      - Syntactic/semantic (for language production)
    - Access files are organized in a series of bins (first syllable or letters)
      - Position within the bins is organized by lexical frequency
      - Access files have “pointers” to meaning information in semantic memory

Search model (e.g., Forster, 1976)



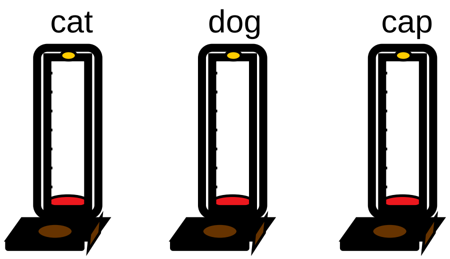
Search model (e.g., Forster, 1976)

- Search model (Forster, 1976, 1979, 1987, 1989)
  - Frequency effects
    - Bin organization
  - Repetition priming effects
    - Temporary reordering of bins in response to recent encounter
  - Semantic priming effects
    - Accounted for by cross referencing in the lexicon
  - Context effects
    - Search is considered to be autonomous, un affected by context (so context effects are "post-access")



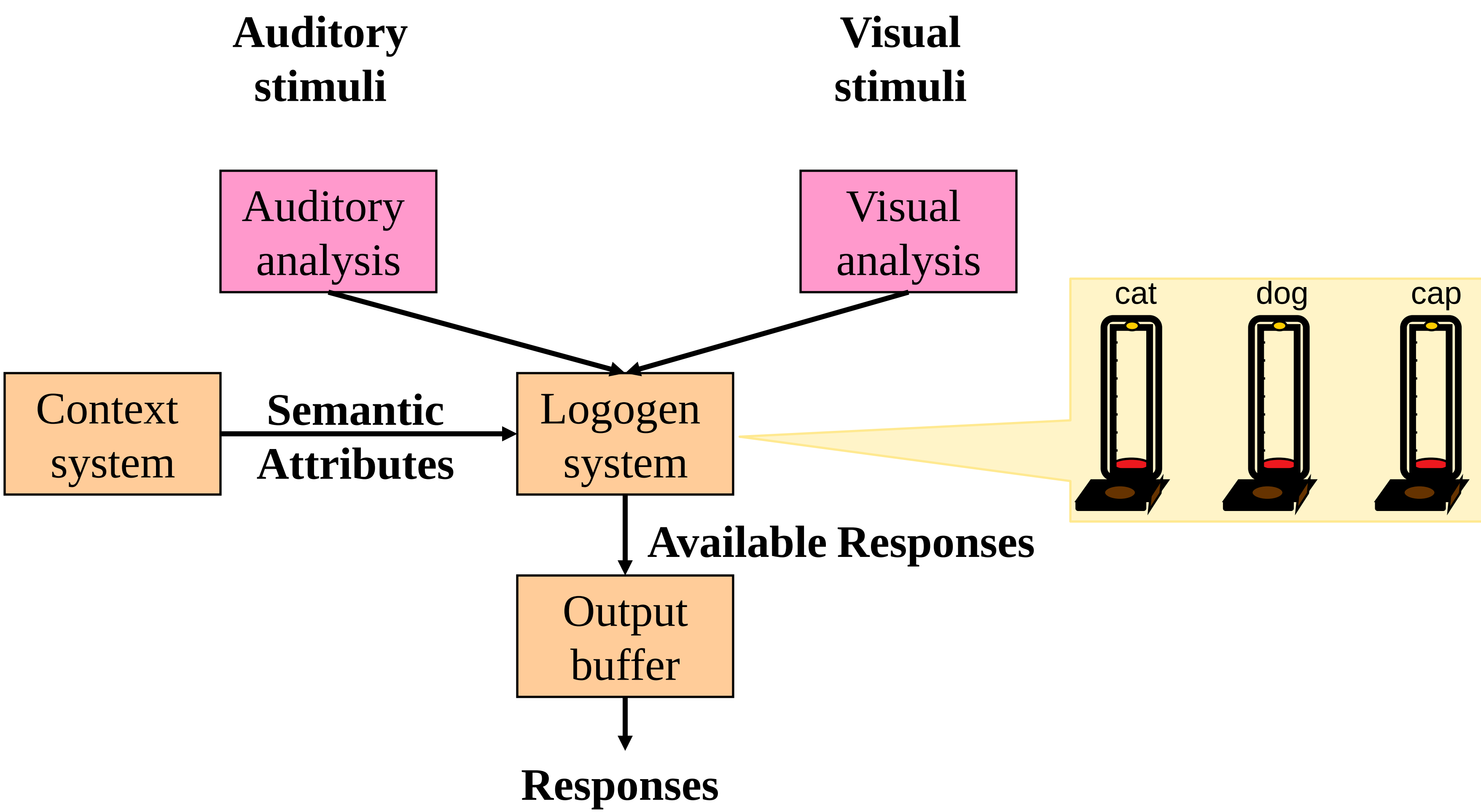
Logogen model (Morton 1969)

- The lexical entry for each word comes with a logogen
  - Logogens specify word's attributes
    - e.g., semantic, orthographic, phonological



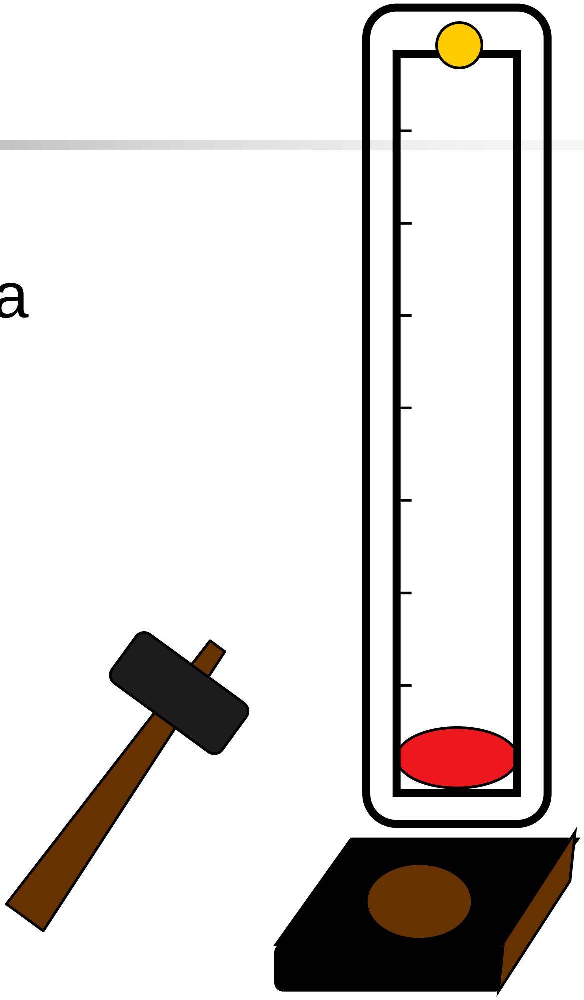
  - *Activated* in two ways
    - *By sensory input*
    - *By contextual information*
  - *Access* (recognition) when reach ***threshold***
    - Different thresholds depending on different factors
      - e.g., frequency
    - Access makes information associated with word available

Logogen model (Morton 1969)



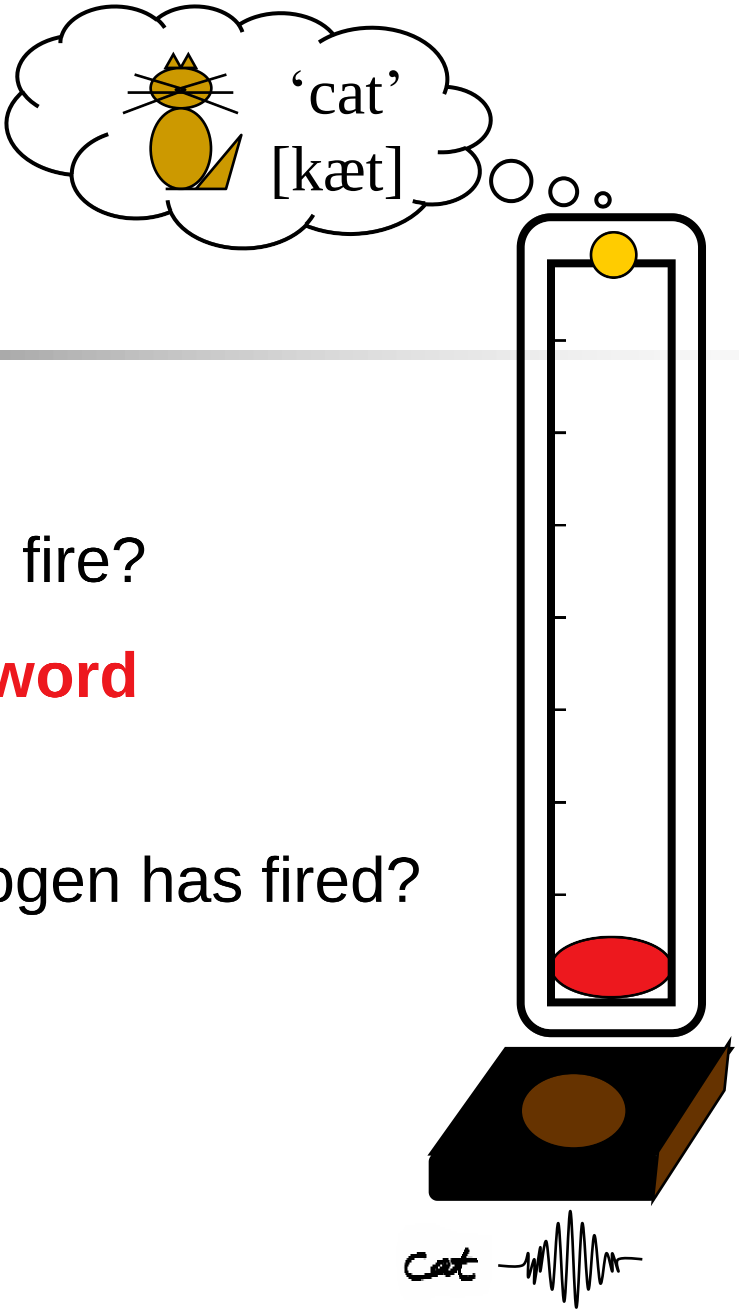
Think of a **logogen** as being like a 'strength-o-meter' at a fairground

When the bell rings, the logogen has '**fired**'

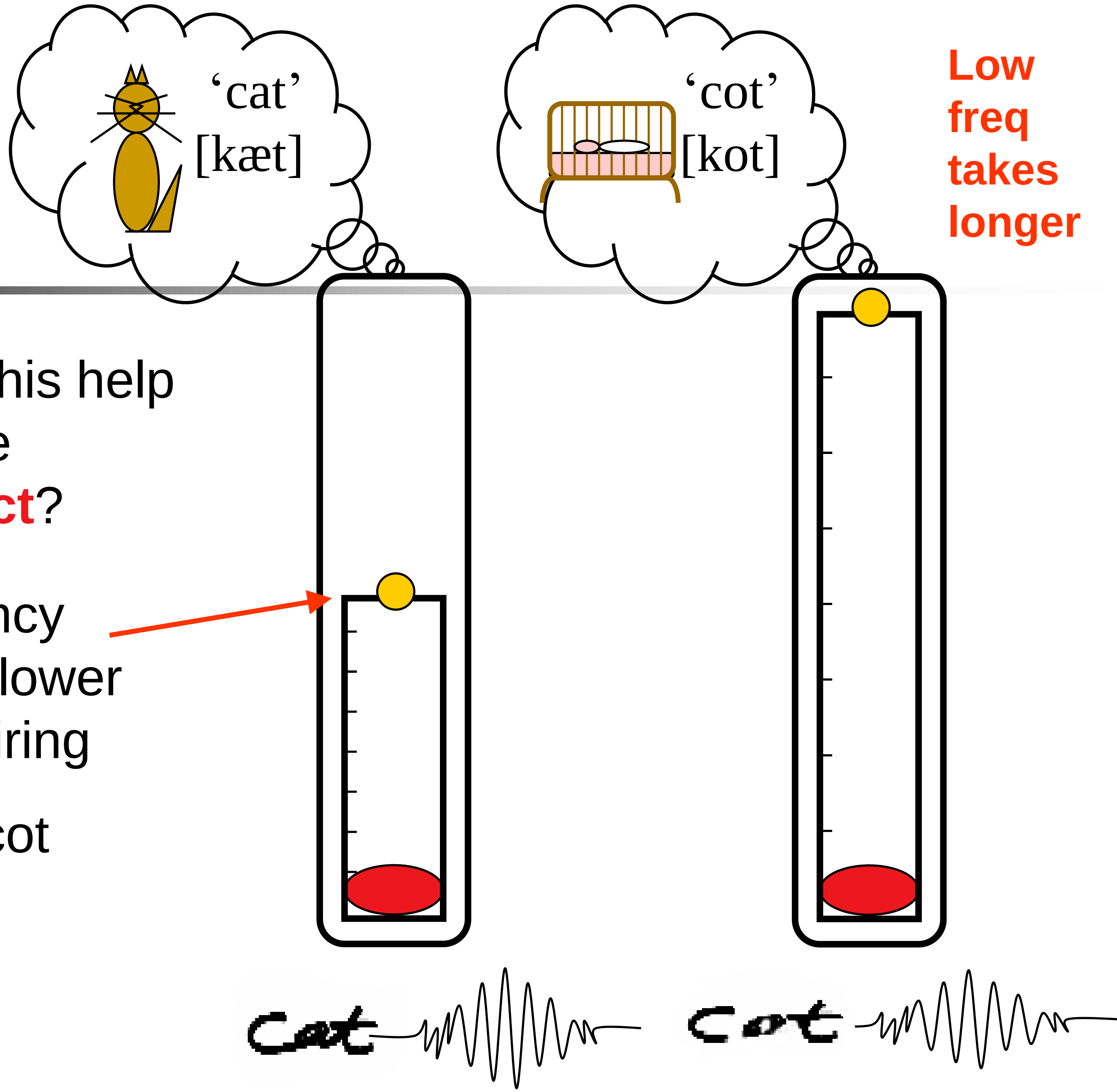


* What makes the logogen fire?
  – **seeing/hearing the word**

* What happens once the logogen has fired?
  – **access to lexical entry!**

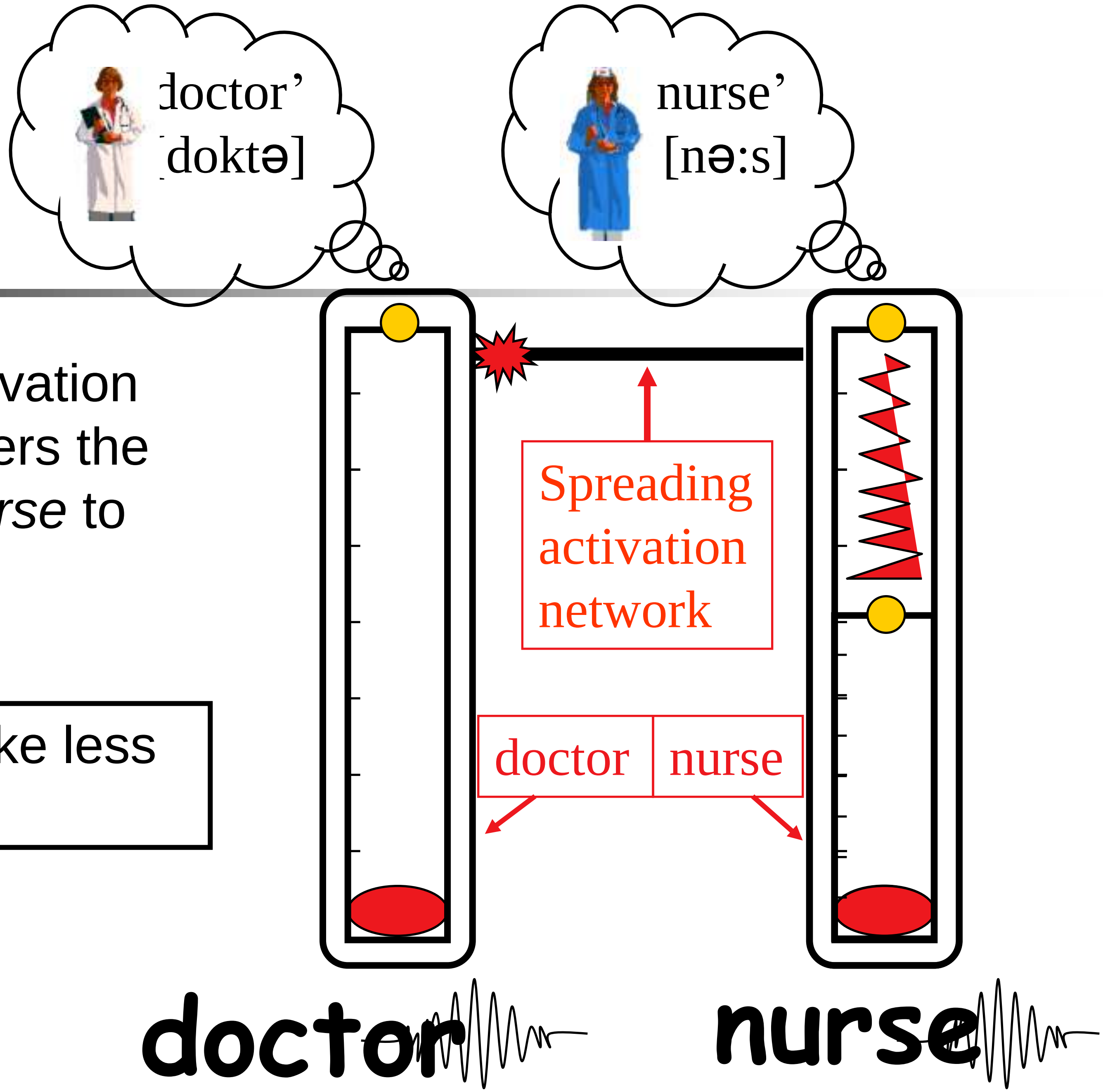


- So how does this help us to explain the **frequency effect**?
  - High frequency words have a lower threshold for firing
  - e.g., cat vs. cot



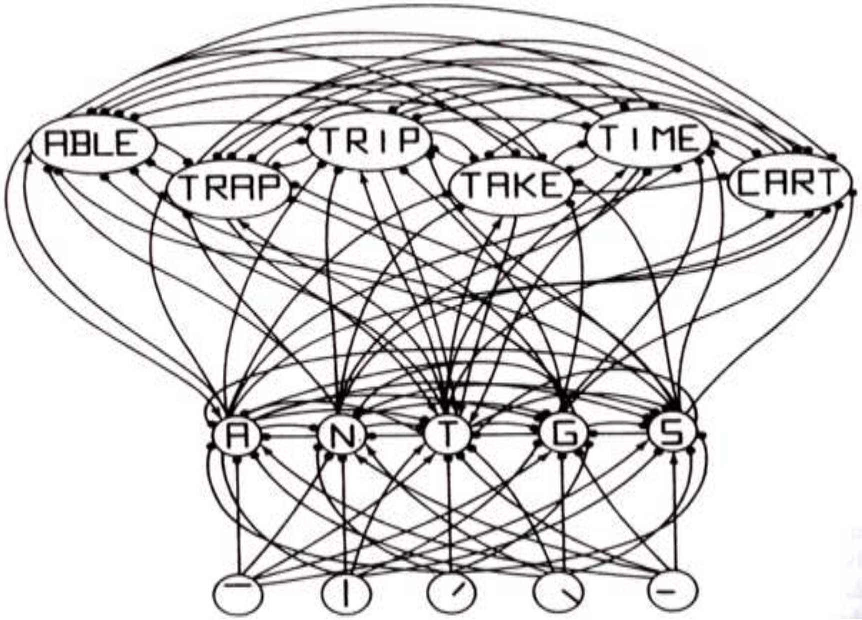
- Spreading activation from *doctor* lowers the threshold for *nurse* to fire

– So *nurse* take less time to fire



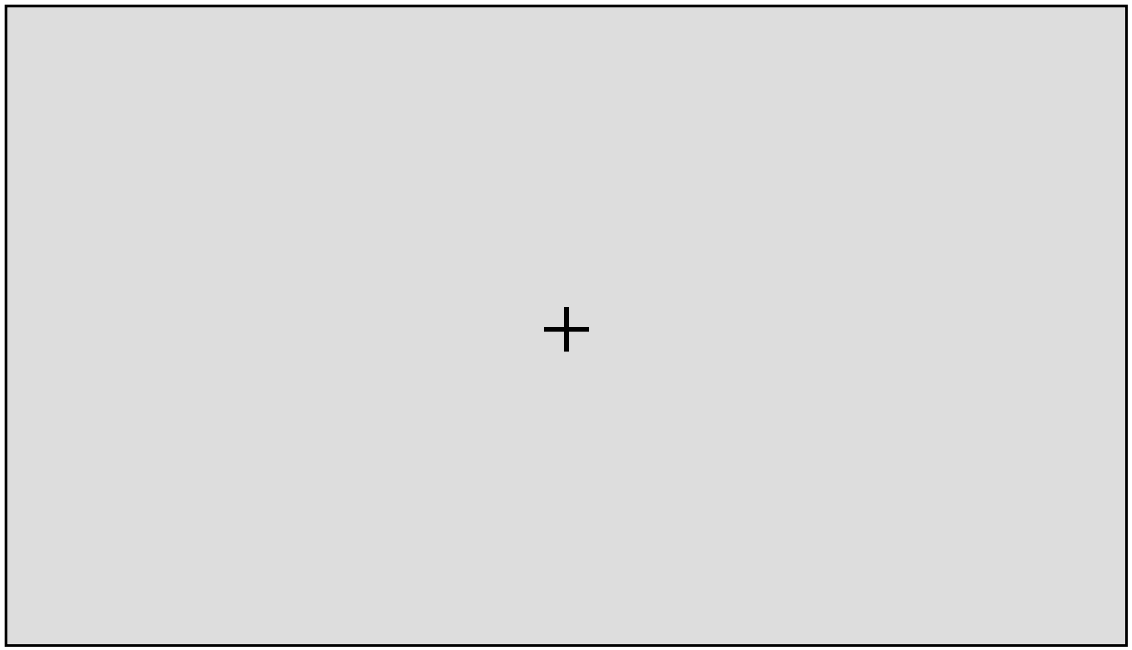
Interactive Activation Model (IAM)

Proposed to account for Word Superiority effect



McClelland and Rumelhart, (1981)

The Word-Superiority Effect (Reicher, 1969)



Until the participant hits some start key

The Word-Superiority Effect (Reicher, 1969)

COURSE

Presented briefly … say 25 ms

The Word-Superiority Effect (Reicher, 1969)

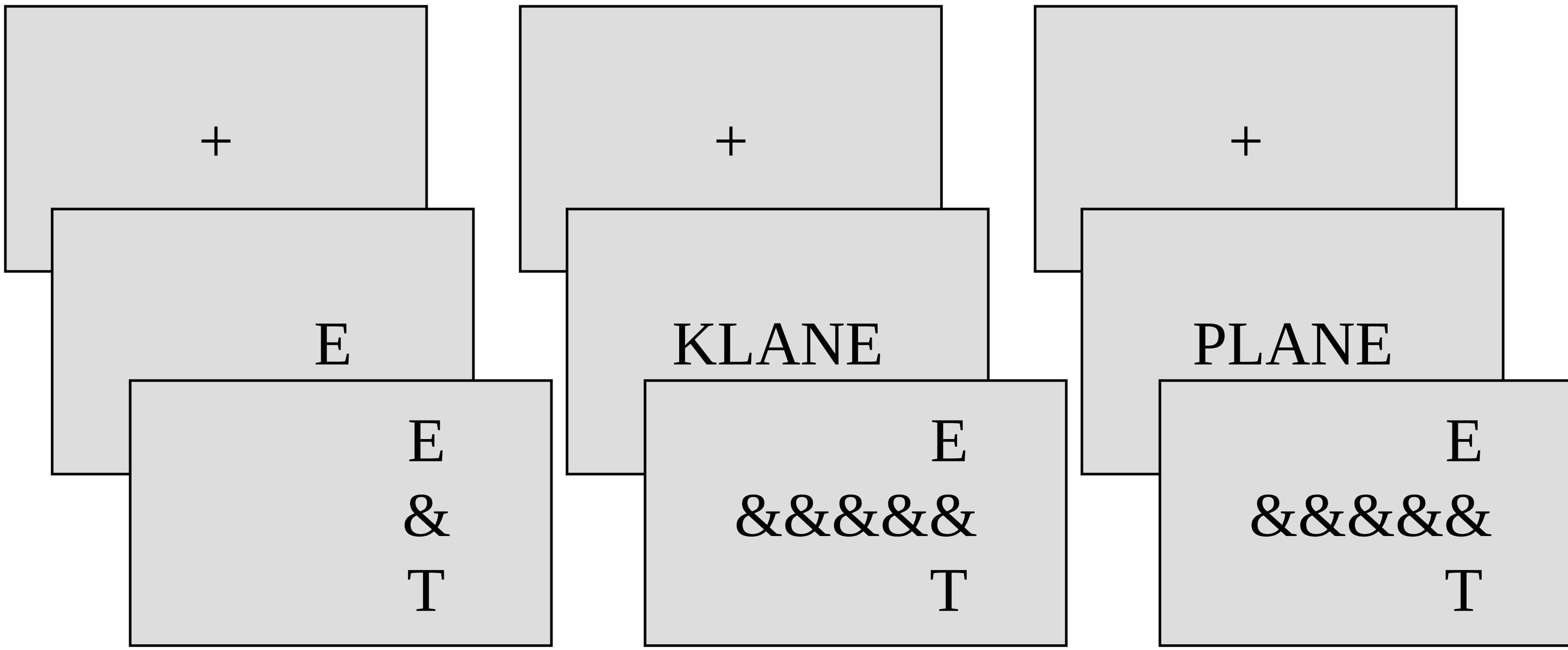
U

&&&&&

A

Mask presented with alternatives above and below the target letter … participants must pick one as the letter they believe was presented in that position.

The Word-Superiority Effect (Reicher, 1969)



Letter only
Say 60%

Letter in Nonword
Say 65%

Letter in Word
Say 80%

Why is identification better when a letter is presented in a word?

Also goes by the name: Interactive Activation and Competition Model (IAC)

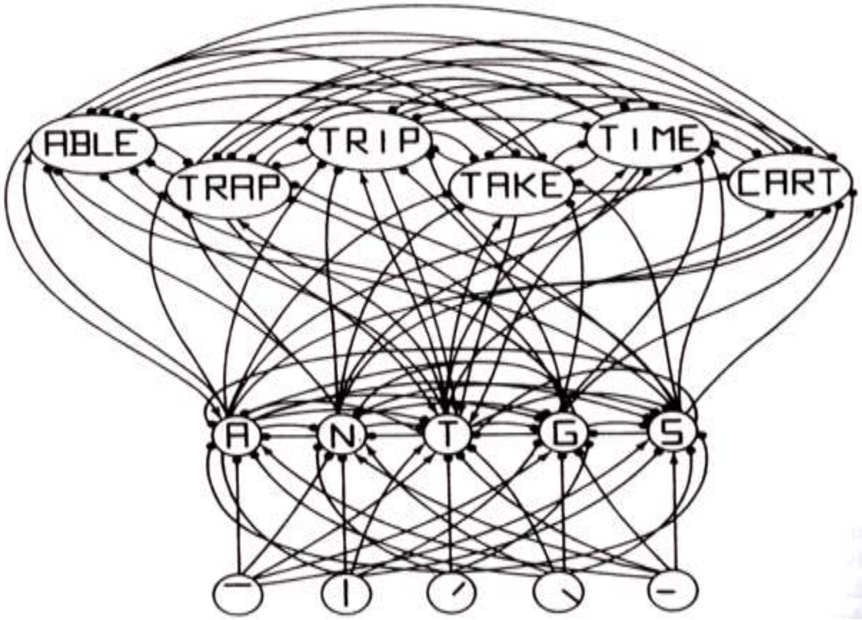
Interactive Activation Model (IAM)

Previous models posed a ***bottom-up*** flow of information (from features to letters to words).

IAM also poses a ***top-down*** flow of information

Nodes:

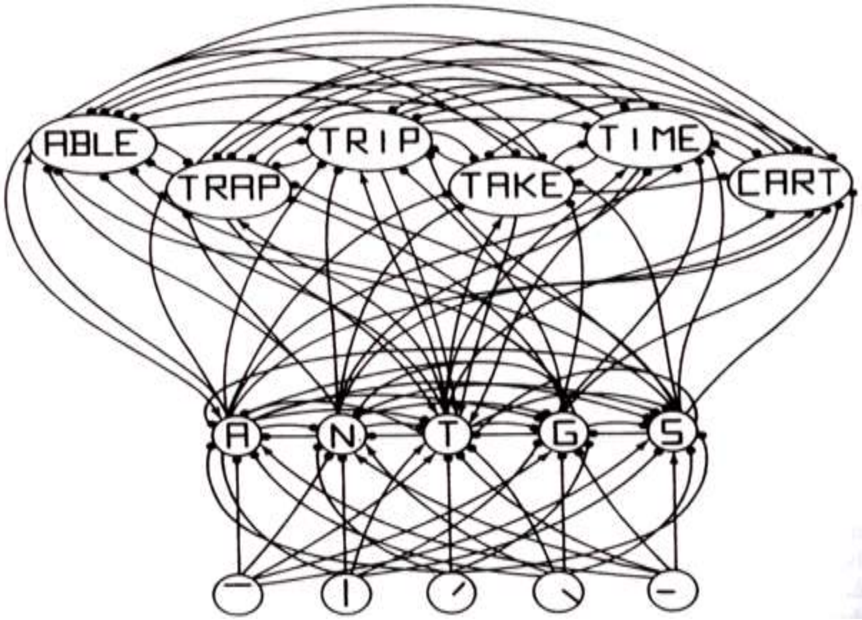
- (visual) feature
- (positional) letter
- word detectors
  - ***Inhibitory*** and ***excitatory*** connections between them.



McClelland and Rumelhart, (1981)

Interactive Activation Model (IAM)

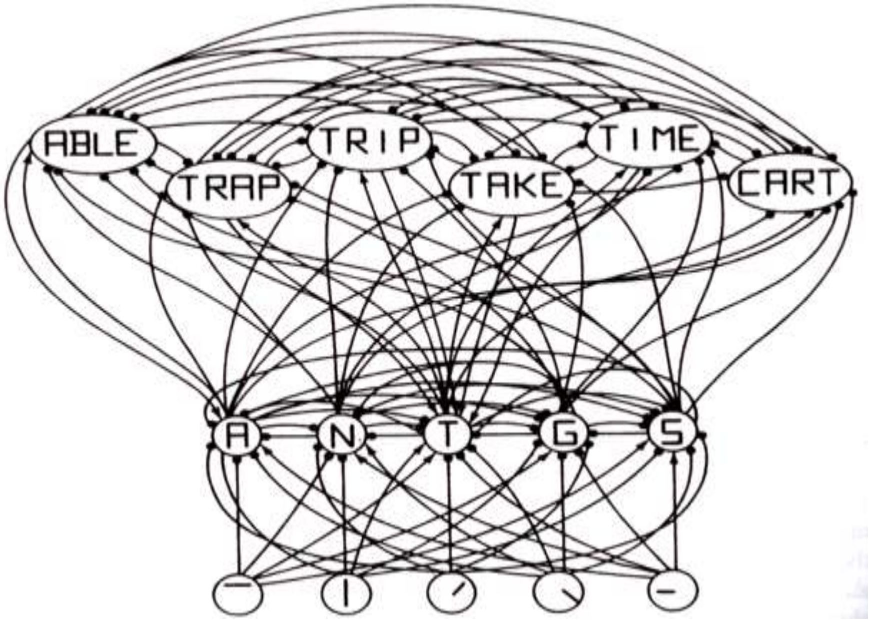
- Inhibitory connections within levels
  - If the first letter of a word is “a”, it isn’t “b” or “c” or …
- Inhibitory and excitatory connections between levels (bottom-up and top-down)
  - If the first letter is “a” the word could be “apple” or “ant” or …., but not “book” or “church” or……
  - If there is growing evidence that the word is “apple” that evidence confirms that the first letter is “a”, and not “b”…..



McClelland and Rumelhart, (1981)

IAM & the word superiority effect

- The model processes at the word and letter levels simultaneously
  - Letters in *words* benefit from bottom-up and top-down activation
  - But letters *alone* receive only bottom-up activation.



McClelland and Rumelhart, (1981)

Cohort model (Marslen-Wilson & Welch, 1978)

- Specifically for auditory word recognition

  (covered in chapter 9 of textbook)

  - Speakers can recognize a word very rapidly
    - Usually within 200-250 msec
  - *Recognition point* (uniqueness point)
    - Point at which a word is unambiguously different from other words and can be recognized (strong emphasis on word onsets)
  - Three stages of word recognition

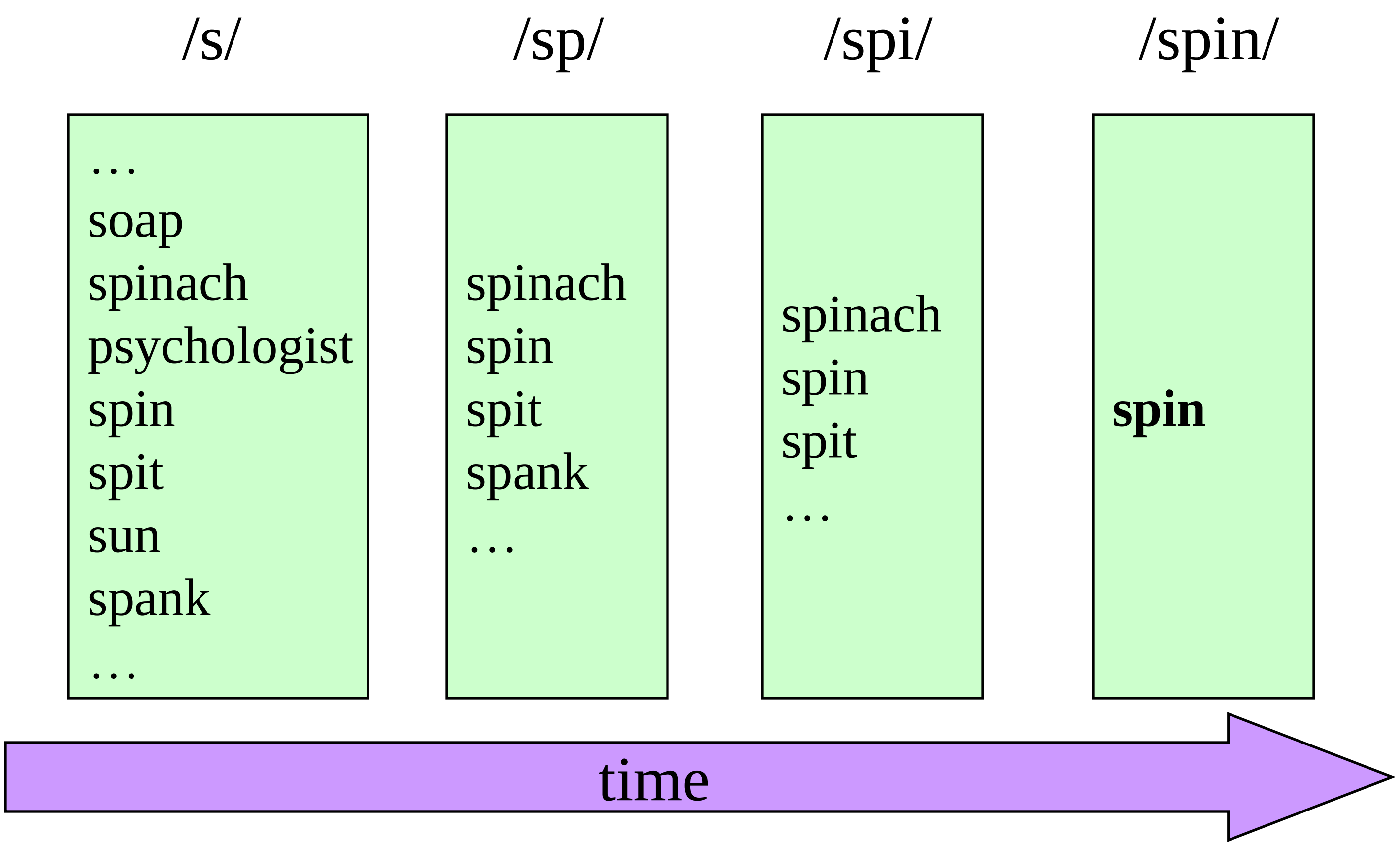
    1) activate a set of possible candidates

    2) narrow the search to one candidate

    3) integrate single candidate into semantic and syntactic context

Cohort model

- Prior context: “I took the car for a …”



Comparing the models

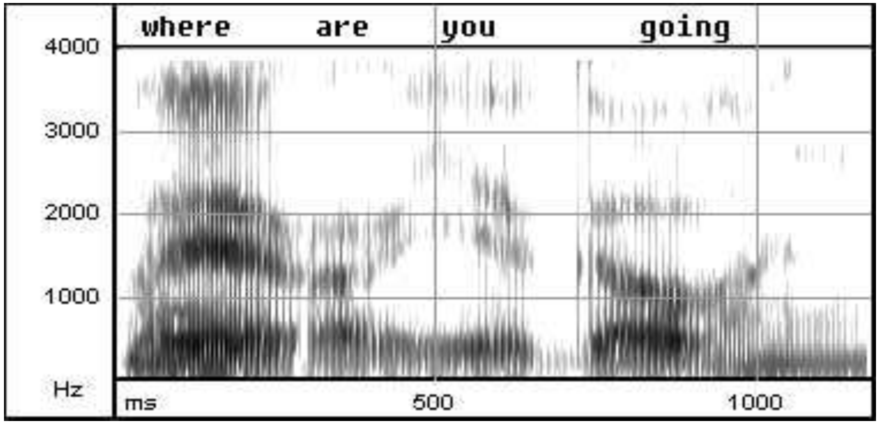
- Each model can account for major findings (e.g., frequency, semantic priming, context), but they do so in different ways.
  - Search model is serial, bottom-up, and autonomous
  - Logogen is parallel and interactive (information flows up and down)
  - AIM is both bottom-up and top-down, uses *facilitation* and *inhibition*
  - Cohort is bottom-up but parallel initially, but then interactive at a later stage

Different signals

<u>Visual word recognition</u>

<u>Speech Perception</u>

Where are you going



- Some parallel input
- Orthography
  - Letters
- Clear delineation
- Difficult to learn

- Serial input
- Phonetics/Phonology
  - Acoustic features
- Usually no delineation
- “Easy” to learn

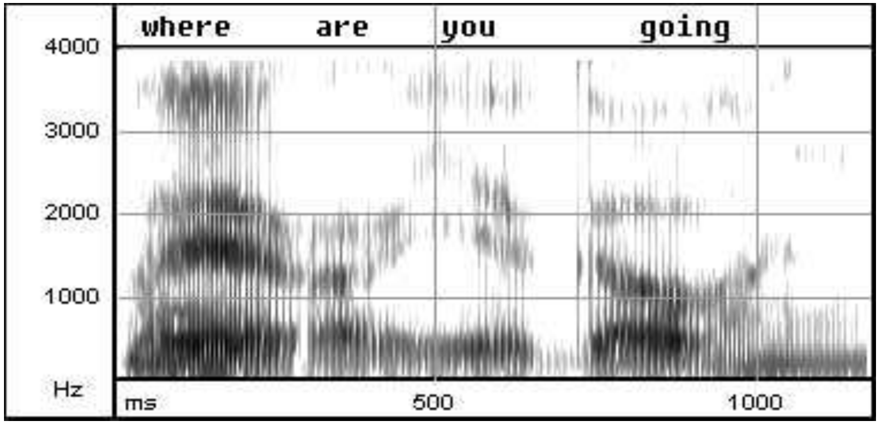
Different signals

Visual word recognition

Where are you going

- Some parallel input
- Orthography
  - Letters
- Clear delineation
- Difficult to learn

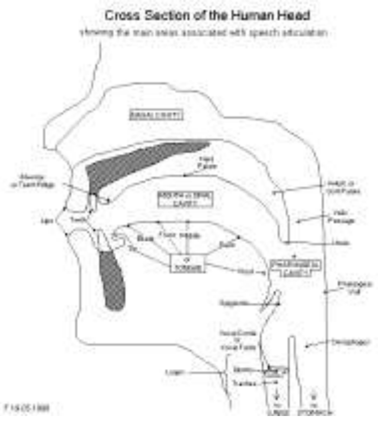
Speech Perception



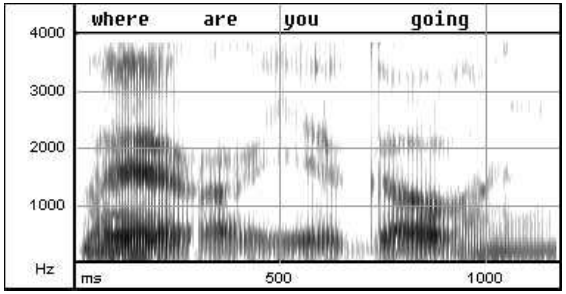
- Serial input
- Phonetics/Phonology
  - Acoustic features
- Usually no delineation
- “Easy” to learn

Speech perception

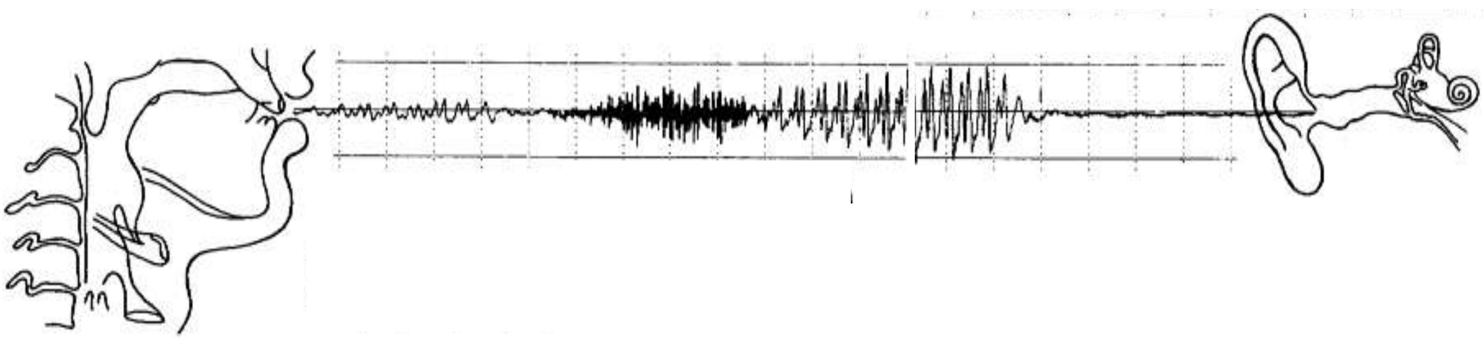
- Articulatory phonetics
  - Production based
    - Place and manner of articulation



- Acoustic phonetics
  - Based on the acoustic signal
    - Formants, transitions, co-articulation, etc.



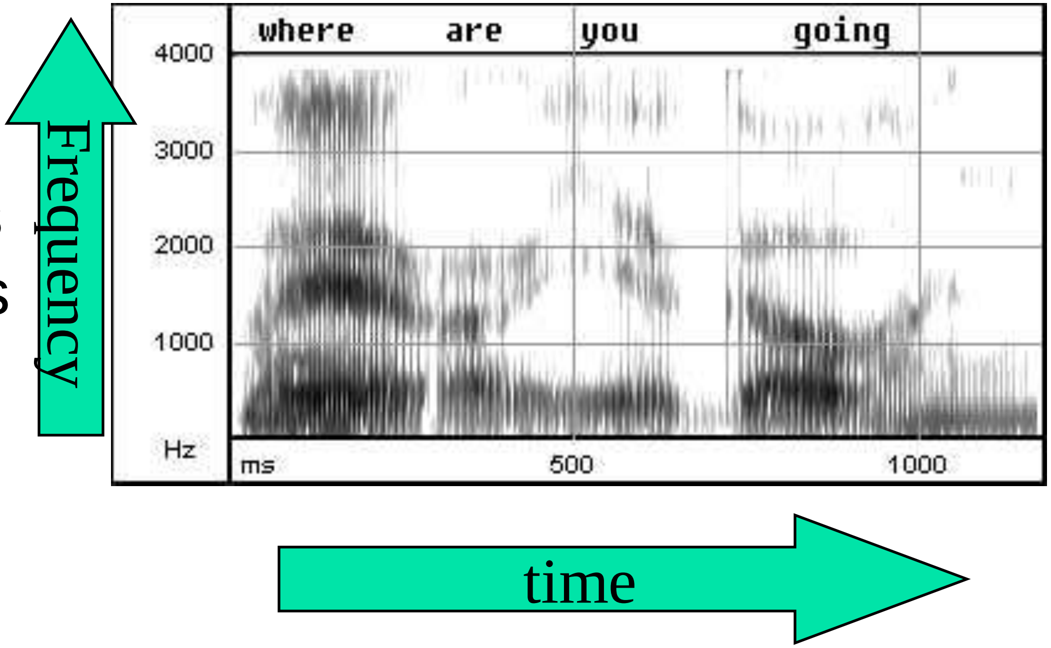
Speech production to perception



- Acoustic cues are extracted and stored in sensory memory and then mapped onto linguistic information
  - Air is pushed into the larynx across the vocal cords and into the mouth nose, different types of sounds are produced.
    - The different qualities of the sounds are represented in ***formants***
  - The formants and other features are mapped onto phonemes

Acoustic features

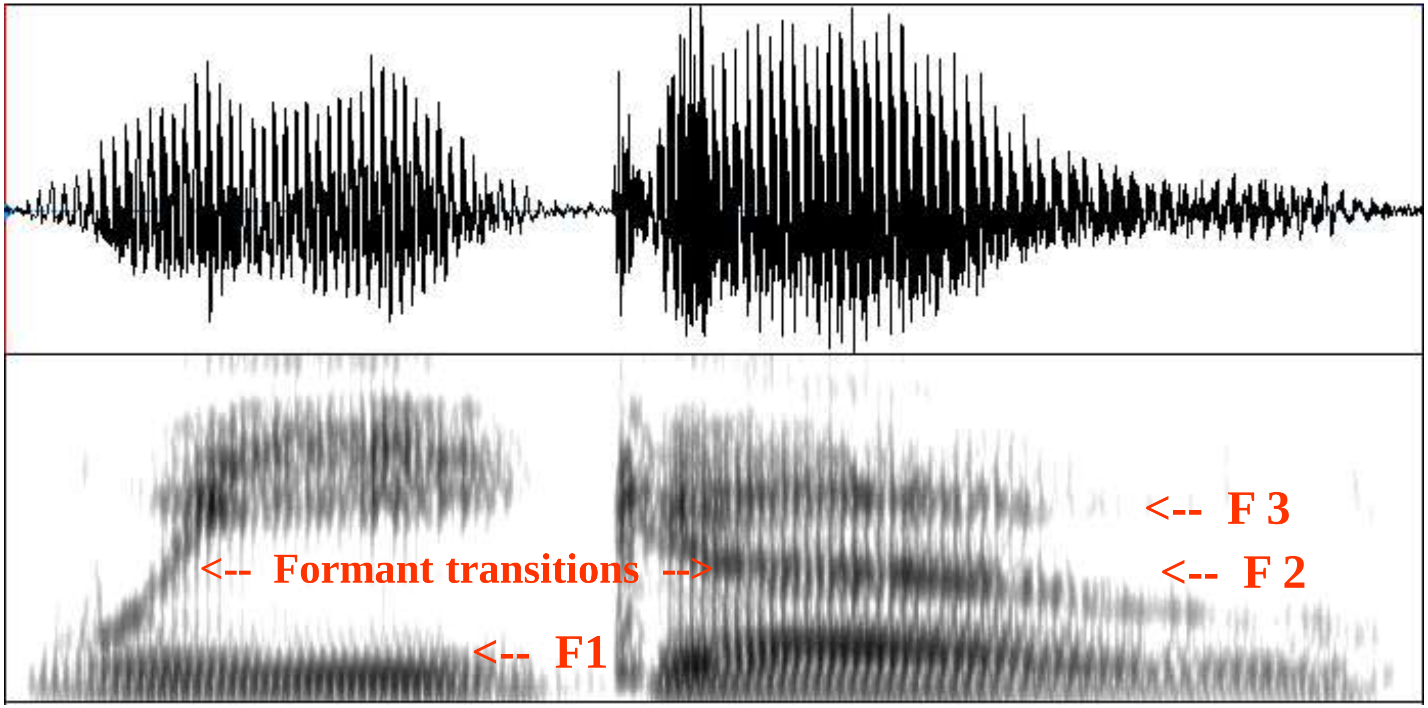
- Spectrogram
  - Time on the x-axis
  - Frequency (pressure under which the air is pushed) on the y-axis
  - Amplitude is represented by the darkness of the lines



Acoustic features

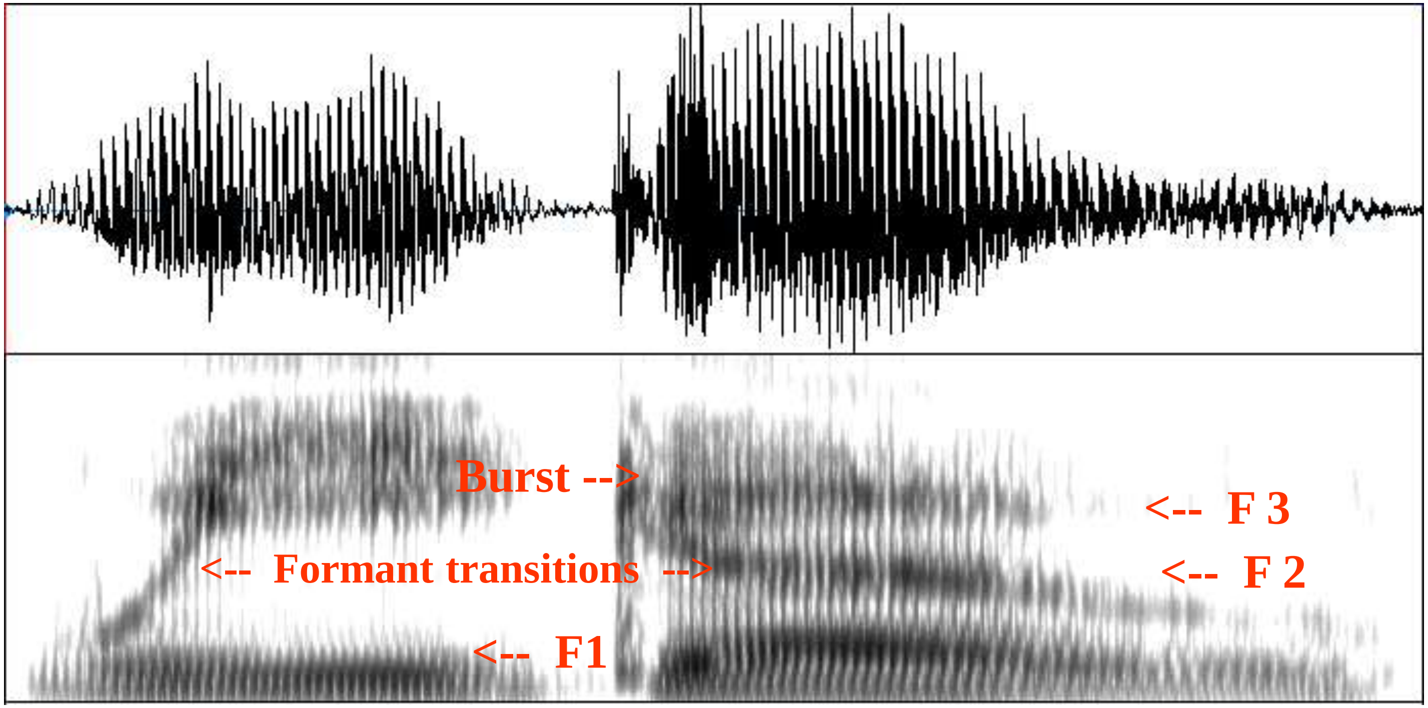
- Acoustic features
  - Formants - bands of resonant frequencies
    - Formant transitions - up or down movement of formants
    - Steady states - flat formant patterns
  - Bursts - sudden release of air
  - Voice onset time (VOT) - when the voicing begins relative to the onset of the phoneme

Formants in a wide-band spectrogram



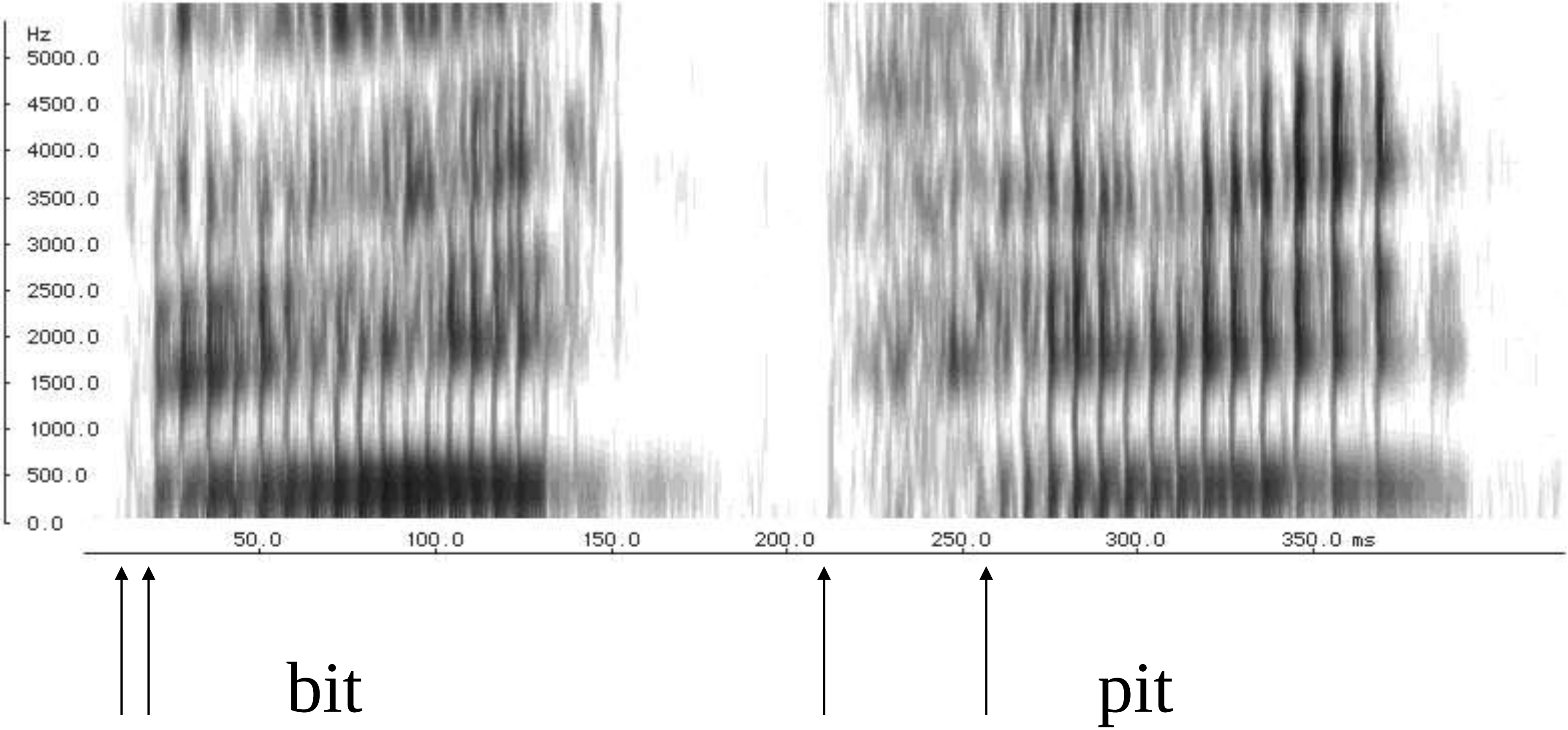
- Formants - bands of resonant frequencies
  - Formant transitions - up or down movement of formants
  - Steady states - flat formant patterns

Formants in a wide-band spectrogram



- Bursts – sudden release of air

Voice-Onset Time (VOT)



bit

pit

5 ms

40 ms

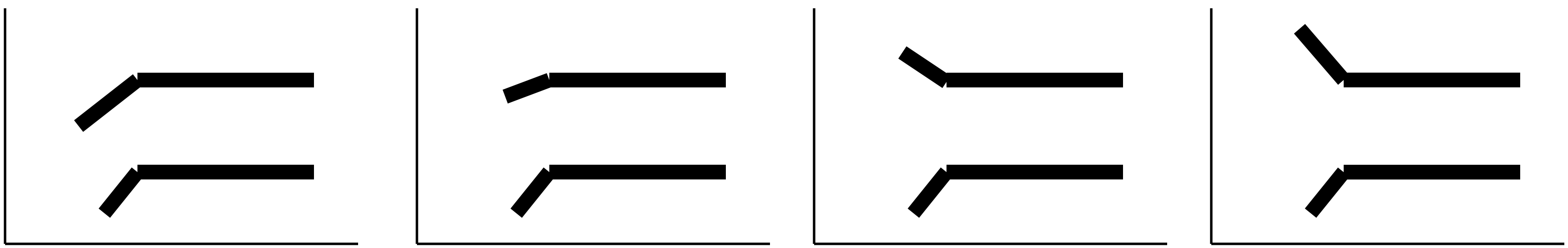
Categorical Perception

- Categorical Perception is the perception of different sensory phenomena as being qualitatively, or categorically, different.
- Liberman et al (1957)
  - Used the speech synthesizer to create a series of syllables panning categories /b/, /d/, /g/ (followed by /a/)
    - Was done by manipulating the F2 formant
  - Stimuli formed a physical continuum
  - Result, people didn’t “hear” a continuum, instead classified them into three categories

Categorical Perception

- Liberman et al (1957)
  1. Set up a continuum of sounds between two categories

/ba/ - /da/

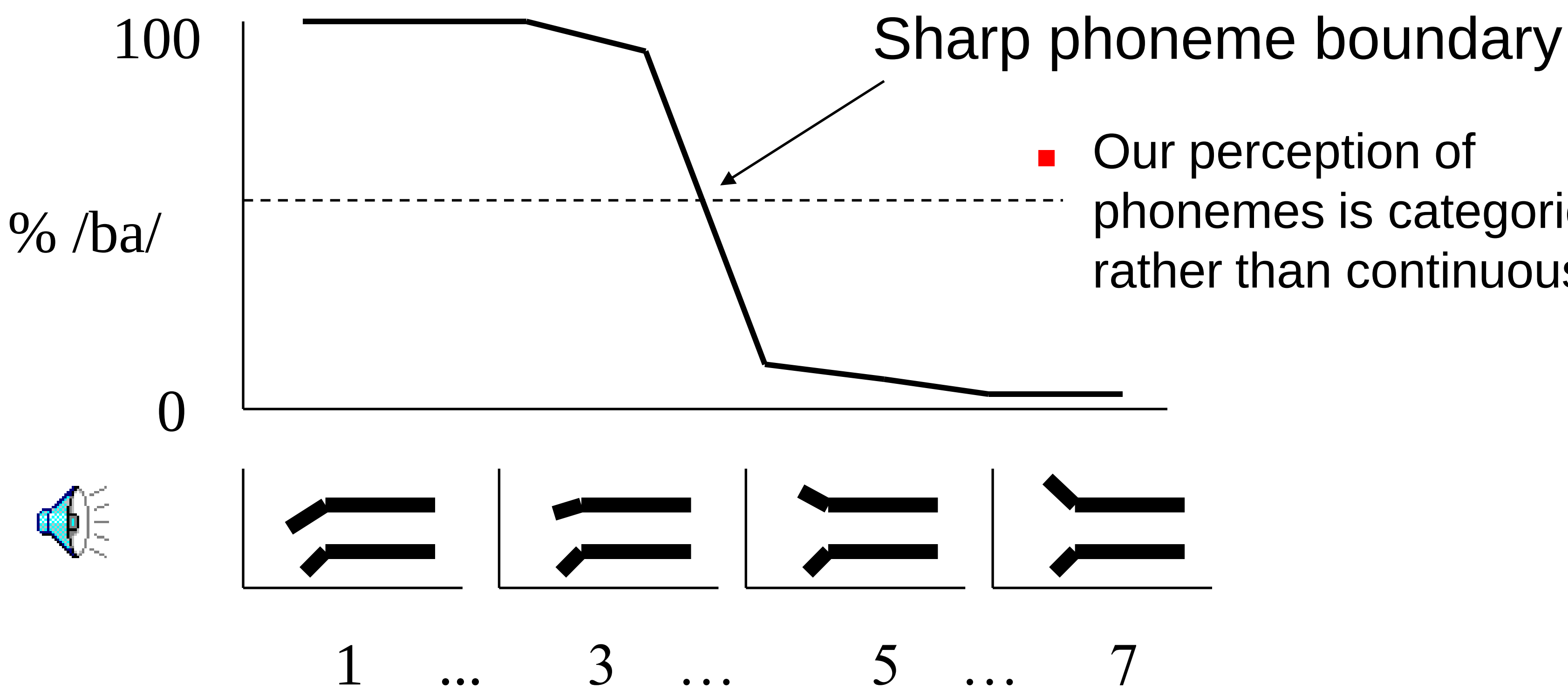


1 ... 3 … 5 … 7

Categorical Perception

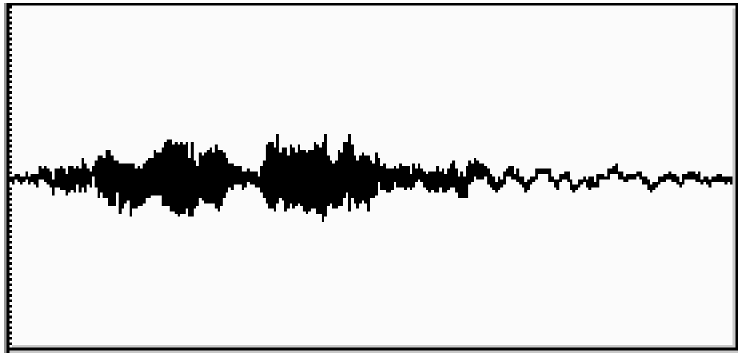
- Liberman et al (1957)

  2. Run an identification experiment



- Our perception of phonemes is categorical rather than continuous.

Hard Problems in Speech Perception

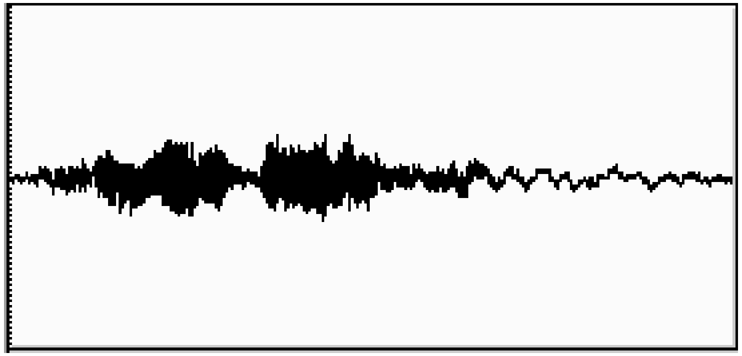


Wave form

- **Linearity (parallel transmission)**: Acoustic features often spread themselves out over other sounds
  - Where does *show* start and *money* end?

- Demo's and info

Hard Problems in Speech Perception

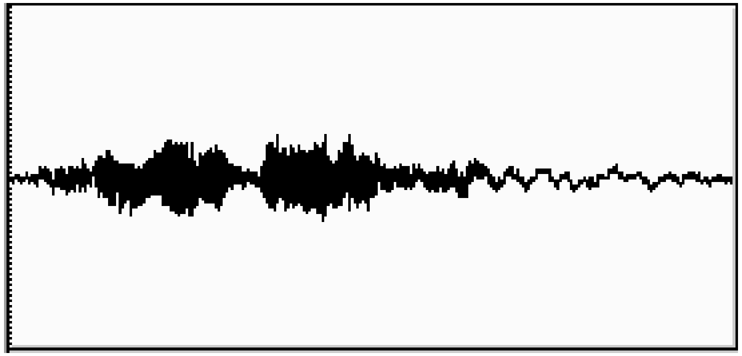


[Wave form]

- **[Invariance]:**
  - One phoneme should have a one waveform
    - But, the /i/ (‘ee’) in ‘money’ and ‘me’ are different
  - There aren’t invariant cues for phonetic segments
    - Although the search continues

- [Demo's and info]

Hard Problems in Speech Perception



[Wave form]

- ***Co-articulation***: the influence of the articulation (pronunciation) of one phoneme on that of another phoneme.
  - Essentially, producing more than one speech sound at once

- [Demo's and info]

Hard Problems in Speech Perception

- **Trading relations**
  - Most phonetic distinctions have more than one acoustic cue as a result of the particular articulatory gesture that gives the distinction.
    - slit–split – the /p/ relies on silence and rising formant, different mixtures of these can result in the same perception
  - Perception must establish some "trade-off" between the different cues.

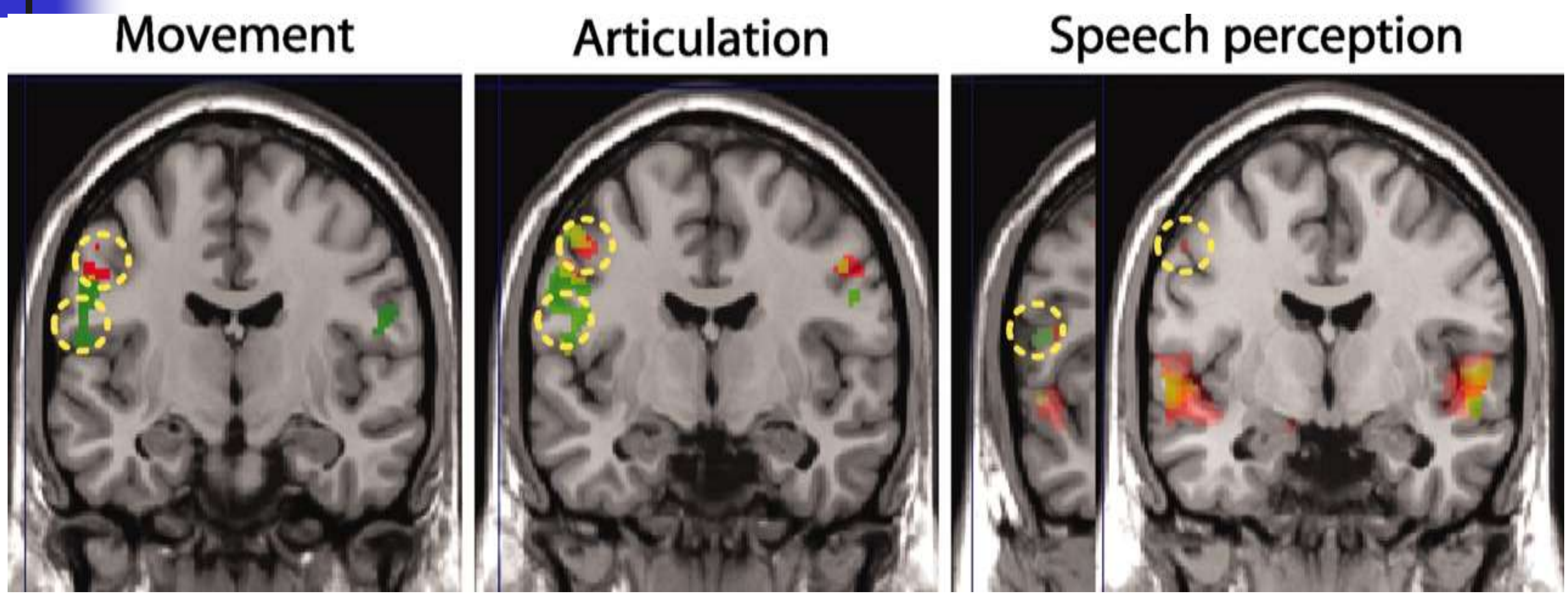
Hard Problems in Speech Perception

- The McGurk effect: McGurk and MacDonald (1976)
  - Showed people a video where the audio and the video don't match
  - Think "dubbed movie"
    - McGurk effect
    - McGurk effect2
  - visual /ga/ with auditory /ba/ often hear /da/
- Implications
  - phoneme perception is an active process
  - influenced by both audio and visual information

Motor theory of speech perception

- A. Liberman (and others, initially proposed in late 50's)
  - Direct translation of acoustic speech into articulatorally defined categories
  - Holds that speech perception and motor control involved linked (or the same) neural processes
    - Theory held that categorical perception was a direct reflection of articulatory organization
      - Categories with discrete gestures (e.g., consonants) will be perceived categorically
      - Categories with continuous gestures (e.g., vowels) will be perceived continuously
  - There is a speech perception module that operates independently of general auditory perception

Frontal slices showing differential activation elicited during lip and tongue movements (Left), syllable articulation including [p] and [t] (Center), and listening to syllables including [p] and [t] (Right)



Pulvermüller F et al. PNAS 2006;103:7865-7870

©2006 by National Academy of Sciences

Motor theory of speech perception

- Some problems for MT
  - Categorical perception found in non-speech sounds (e.g., music)
  - Categorical perception for speech sounds in non-humans
    - Chinchillas can be trained to show categorical perception of /t/ and /d/ consonant-vowel syllables (Kuhl & Miller, 1975)

Other theories of speech perception

- Direct Realist Theory (C. Fowler and others)
  - Similar to Motor theory, articulation representations are key, but here they are directly perceived
  - Perceiving speech is part of a more general perception of gestures that involves the motor system
- General Auditory Approach (e.g., Diehl, Massaro)
  - Do not invoke special mechanisms for speech perception, instead rely on more general mechanisms of audition and perception

- For nice reviews see:
  - Diehl, Lotto, & Holt (2003)
  - Galantucci, Fowler, Turvey (2006)

Top-down effects on Speech Perception

- Phoneme restoration effect
- Sentence context effects

Phoneme restoration effect

Listen to a sentence which contained a word from which a phoneme was deleted and replaced with another noise (e.g., a cough)

The state governors met with their respective legi*latures convening in the capital city.

*** /s/ deleted and replaced with a cough**

Click here for a demo and additional information

Phoneme restoration effect

- Typical results:
  - Participants heard the word normally, despite the missing phoneme
  - Usually failed to identify which phoneme was missing
- Interpretation
  - We can use top-down knowledge to “fill in” the missing information

Phoneme restoration effect

Further experiments (Warren and Warren, 1970):

What if the missing phoneme was ambiguous

The *eel was on the axle.
The *eel was on the shoe.
The *eel was on the orange.
The *eel was on the table.

Results:

Participants heard the contextually appropriate word normally, despite the missing phoneme

Phoneme restoration effect

- Possible loci of phoneme restoration effects
  - *Perceptual loci of effect*:
    - Lexical or sentential context influences the way in which the word is initially perceived.
  - *Post-perceptual loci of effect*:
    - Lexical or sentential context influences decisions about the nature of the missing phoneme information.

Beyond the segment

Shillcock (1990): hear a sentence, make a lexical decision to a word that pops up on computer screen (cross-modal priming)

Hear:

The scientist made a new discovery last year.

NUDIST

Cross-modal priming

Shillcock (1990): hear a sentence, make a lexical decision to a word that pops up on computer screen (cross-modal priming)

Hear:

The scientist made a novel discovery last year.

NUDIST

Cross-modal priming

Shillcock (1990): hear a sentence, make a lexical decision to a word that pops up on computer screen (cross-modal priming)

Hear:

The scientist made a novel discovery last year.
The scientist made a new discovery last year. *faster*

Cross-modal priming

Shillcock (1990): hear a sentence, make a lexical decision to a word that pops up on computer screen (cross-modal priming)

Hear:

The scientist made a novel discovery last year.

The scientist made a new **dis****covery last year.** *faster*

NUDIST gets primed by segmentation error

Although no conscious report of hearing "nudist"

Beyond the segment

- Prosody and intonation
  - English:
    - Speech is divided into phrases.
    - Word stress is meaningful in English.
    - Stressed syllables are aligned in a fairly regular rhythm, while unstressed syllables take very little time.
    - Every phrase has a focus.
    - An extended flat or low-rising intonation at the end of a phrase can indicate that a speaker intends to continue to speak.
    - A falling intonation sounds more final.

Beyond the segment

- Prosodic factors (supra segmentals)
  - Stress
    - Emphasis on syllables in sentences
  - Rate
    - Speed of articulation
  - Intonation
    - Use of pitch to signify different meanings across sentences

Beyond the segment

- Stress effects
  - On meaning
    - “black bird” versus “blackbird”
  - Top-down effects on perception
    - Better anticipation of upcoming segments when syllable is stressed

Beyond the segment

- Rate effects
  - How fast you speak has an impact on the speech sounds
    - Faster talking - shorter vowels, shorter VOT
  - Normalization
    - Taking speed and speaker information into account
      - Rate normalization
      - Speaker normalization