Introduction to Experimental Psycholinguistics

Brian Dillon + Shota Momma
UMass Amherst
LINGUIST611 Spring 2024
- **Today’s question:** How do we retrieve words from the mental lexicon?
- **What’s in a word?**
Questions

→ When does lexical access happen?

→ How are candidate words activated in response to input?
Retrieving words

**Activation / access / generation function:** How are potential word candidates identified, given some sensory input?

**Integration function:** How does output of recognition process relate to higher order levels of analysis (e.g. syntactic or semantic analysis)?

**Selection function:** How do we make a ‘decision’ about the outcome of the recognition process?
Retrieving words

**Activation / access / generation function**: How are potential word candidates identified, given some sensory input?

**Integration function**: How does output of recognition process relate to higher order levels of analysis (e.g. syntactic or semantic analysis)?

**Selection function**: How do we make a ‘decision’ about the outcome of the recognition process?
He thinks he won’t get the letter...
- **When does it happen?** Approximately ~200ms of a word is necessary to identify it in context.
- Linguistic input is analyzed *incrementally*, word by word.
- Analysis occurs at multiple levels in parallel.

Marslen-Wilson 1975
Word frequency effects

→ Measures of lexical access (lexical decision / naming) are strongly impacted by word frequency.

→ Word frequency is the number of times you would count a word if you went and measured a corpus of text in the language.

→ Frequency appears to be a basic organizing principle in the mental lexicon: more frequent words are accessed more quickly than less frequent ones.

Figure 3: The word frequency effect in the English Lexicon Project lexical decision data. The curve shows that above a frequency of 100 per million (log10 = 2), there is no more difference between the stimuli. In contrast, nearly half of the frequency effect is situated below frequencies of 1 per million (log10 = 0). Source: Keuleers et al. (2010a).
Questions

→ **When** does lexical access happen?

→ **How** are candidate words activated in response to input?
Three theories today

→ Logogen models (Continuous evidence accumulation)
→ COHORT (Incrementality)
→ TRACE (Competition between hypotheses)

We will not try to select a correct theory - instead we will see that each gives us some important insight into how lexical access happens.
- **Words are activated in parallel in response to speech input.**

- A logogen is an individual **evidence accumulation unit** associated with a particular word. Evidence accumulates over time.

- When a logogen reaches its activation threshold, the lexical item has been accessed.

- Words differ in their activation thresholds: **frequent words have lower thresholds**.
Decisions take time, and as a rule more difficult decisions take more time. But this only raises the question of what consumes the time. For decisions informed by a sequence of samples of evidence, the answer is straightforward: more samples are available with more time. Indeed, the speed and accuracy of such decisions are explained by the accumulation of evidence to a threshold or bound. However, the same framework
Evidence accumulation: *Big idea!*
Cohort model (Marslen-Wilson, 1987)

- All possible matching candidates are activated in parallel as the word unfolds. The set of matching candidates is called the **cohort**.
- Candidates are eliminated from consideration once they no longer match the input.

\[
/k/ \Rightarrow /k\lambda/ \Rightarrow /k\lambda p/ \Rightarrow /k\lambda p.l/ \]

<table>
<thead>
<tr>
<th>cat</th>
<th>cut</th>
<th>cup</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>cup</td>
<td>cup</td>
</tr>
<tr>
<td>candy</td>
<td>couple</td>
<td>couple</td>
</tr>
<tr>
<td>candle</td>
<td>cuddle</td>
<td>cupping</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Cohort model (Marslen-Wilson, 1987)

- All possible matching candidates are activated in parallel as the word unfolds. The set of matching candidates is called the cohort.
- Candidates are eliminated from consideration once they no longer match the input.
- Candidates are generated bottom-up (i.e. from the speech stream alone), but context / semantics may select candidates from cohort.

/k/ ➞ /kʌ/ ➞ /kʌp/ ➞ /kʌp.l/

cat  cut  cup  cup
catch  cup  couple  couple
candy  cuddle  cupping...
candle  cull  couple  ...
Incrementality: *Big idea!*

- **Word uniqueness point:** The point at which the cohort only has a single member.

\[
/k/ \rightarrow /k\emptyset/ \rightarrow /k\emptyset p/ \rightarrow /k\emptyset p.l/
\]

<table>
<thead>
<tr>
<th>cat</th>
<th>cut</th>
<th>cup</th>
<th>cup</th>
<th>couple</th>
</tr>
</thead>
<tbody>
<tr>
<td>catch</td>
<td>cup</td>
<td>couple</td>
<td>cupping</td>
<td>couple</td>
</tr>
<tr>
<td>candy</td>
<td>cuddle</td>
<td>cupping</td>
<td></td>
<td>word uniqueness point</td>
</tr>
<tr>
<td>candle</td>
<td>cull</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Cohort model (Marslen-Wilson, 1987)

- COHORT proposes that word activation operates **without feedback or guidance from the context**.
- The cohort is identified purely on the basis of bottom-up information: words are automatically activated, **regardless of whether they are appropriate in the context or not!**

/k/ → /kʌ/ → /kʌp/ → /kʌp.ɪ/  

cat  catch  candy  candle  …  
cut  cup  cuddle  cull  …
cup  couple  cupping  …
couple  word uniqueness point
Prediction: Non-word judgments

/s/  ➔  /sθ/

sing  

sat

spell  

…
Prediction: Non-word judgments
Non-word judgments (Marslen-Wilson, 1984)
- Word candidates are eliminated from consideration as soon as they are no longer consistent with the bottom-up input.
- Word candidates are identified strictly left-to-right.
- Word candidates are identified on the basis of phonetic information alone.
→ **Lexicon** can be modeled as a network: **Long-term memory** encodes links between words, between words and phonemes, and between phonemes and features.
→ **Nodes** in the network have gradient levels of **activation**. Activation can spread through **excitatory** links...
... and can be used to suppress or inhibit other notes through inhibitory connections.
Network dynamics (t = 1)
Network dynamics (t = 2)
Network dynamics (t = 3)
Network dynamics ($t = 4$)
Network dynamics (t = 5)
Because of lateral inhibitory connections, word candidates compete for activation in a winner-take-all fashion.
Because of lateral inhibitory connections, word candidates **compete** for activation in a **winner-take-all** fashion.

Words receive more or less activation from input, but they are not **removed** entirely from candidate set.
Eye-tracking

→ Visual world paradigm: Experimental paradigm where participants are given instructions on how to interact with a display (e.g. please click on the picture of a beetle), while their eye movements are continuously tracked.

**FIG. 3.** An example of a stimulus display presented to participants.
→ **Fixations** are points of relative stability, where the eye is located on one position in the text. A fixation typically lasts around 150-250 milliseconds.

→ **Saccades** are rapid ballistic movements that last around 25-40 milliseconds, during which you are functionally blind (**saccadic suppression**).

→ It takes about **150ms** to plan and execute a saccade.
The eyes (probabilistically) reflect the contents of attention. Where people look is therefore one index of what is activated / in attention: the eye-mind hypothesis.
Because of lateral inhibitory connections, word candidates compete for activation in a winner-take-all fashion.

Words receive more or less activation from input, but they are not removed entirely from candidate set.

Computer simulations with TRACE

**FIG. 3.** An example of a stimulus display presented to participants.

**FIG. 2.** Predicted response probabilities converted from TRACE using the scaled Luce choice rule.
Rhyme competitor can be activated with later input.

**FIG. 3.** An example of a stimulus display presented to participants.

**FIG. 2.** Predicted response probabilities converted from TRACE using the scaled Luce choice rule.
Rhyme competitor can be activated with later input.

**Allopenna et al (1998)**

**FIG. 4.** Probability of fixating each item type over time in the full competitor condition in Experiment 1. The data are averaged over all stimulus sets given in Table 1; the words given in the figure are examples of one set.

**FIG. 2.** Predicted response probabilities converted from TRACE using the scaled Luce choice rule.

Multiple candidates activated

**FIG. 4.** Probability of fixating each item type over time in the full competitor condition in Experiment 1. The data are averaged over all stimulus sets given in Table 1; the words given in the figure are examples of one set.

**FIG. 2.** Predicted response probabilities converted from TRACE using the scaled Luce choice rule.

Humans activate rhyme competitor!  
Rhyme competitor can be activated with later input.

**FIG. 4.** Probability of fixating each item type over time in the full competitor condition in Experiment 1. The data are averaged over all stimulus sets given in Table 1; the words given in the figure are examples of one set.

**FIG. 2.** Predicted response probabilities converted from TRACE using the scaled Luce choice rule.

Dahan et al. (2001)
Critically: ‘speaker’ was not categorically ruled out at /b/...

This is one example of a right context effect: When later input overrides or suggests a reinterpretation of previous input.
Coarticulatory cues on vowels

→ **Vowels contain coarticulatory cues**: Formant transitions preceding the stop are one cue to the place of articulation of a stop.

Dahan et al. (2001)
Coarticulatory cues on vowels

→ ... and listeners are sensitive to these cues.

Dahan et al. (2001)
Gradient phonetics and lexical access

McMurray et al. (2009)
Gradient phonetics and lexical access

[?]ɛər...
Gradient prototypicality effects observed at disambiguation point: Looks to target were slower as the VOT was more distant from the VOT continuum end point.
→ Multiple candidates are activated and maintained: Both ‘parakeet’ and ‘barricade’ are activated and maintained in working memory.

→ Amount of activation for each depends on how strongly the input matches the prototype.
Right context effects

Does right context influence interpretation of phonetic ambiguities?

<table>
<thead>
<tr>
<th>Context</th>
<th>Distance</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tent-biasing</td>
<td>Near (3 syllables)</td>
<td>When the [t/d]ent in the forest was ...</td>
</tr>
<tr>
<td>Dent-biasing</td>
<td>Near (3 syllables)</td>
<td>When the [t/d]ent in the fender was ...</td>
</tr>
<tr>
<td>Tent-biasing</td>
<td>Far (6-8 syllables)</td>
<td>When the [t/d]ent was noticed in the forest, ...</td>
</tr>
<tr>
<td>Dent-biasing</td>
<td>Far (6-8 syllables)</td>
<td>When the [t/d]ent was noticed in the fender, ...</td>
</tr>
</tbody>
</table>

Table 1: Example stimuli from the experiment in each biasing context and distance condition.
Right context effects

Consistent effect of right context: Bushong and colleagues observe effect of biasing context at near and far distances, across all VOTs.

Table 1: Example stimuli from the experiment in each biasing context and distance condition.

<table>
<thead>
<tr>
<th>Context</th>
<th>Distance</th>
<th>Sentence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tent-biasing</td>
<td>Near (3 syllables)</td>
<td>When the [t/d]ent in the <strong>forest</strong> was ...</td>
</tr>
<tr>
<td>Dent-biasing</td>
<td>Near (3 syllables)</td>
<td>When the [t/d]ent in the <strong>fender</strong> was ...</td>
</tr>
<tr>
<td>Tent-biasing</td>
<td>Far (6-8 syllables)</td>
<td>When the [t/d]ent was noticed in the <strong>forest</strong>, ...</td>
</tr>
<tr>
<td>Dent-biasing</td>
<td>Far (6-8 syllables)</td>
<td>When the [t/d]ent was noticed in the <strong>fender</strong>, ...</td>
</tr>
</tbody>
</table>

Bushong et al., 2017
Listeners are sensitive to potential ambiguity of stimulus, and are more likely to commit to one word before additional context if there is less perceptual ambiguity.

Figure 4: Proportion of responses before biasing context by perceptual ambiguity. Error bars are 95% confidence intervals.
→ **Logogen models**: We gather evidence over time for multiple words simultaneously, they differ in their thresholds.

→ **COHORT**: We incrementally narrow the space of possible words as input comes in.

→ **TRACE**: Lexical access relies on a network architecture that allows interactive activation of words in parallel over time.
→ **Incrementality**: We activate word candidates incrementally; word recognition can happen with as little as 200ms of a word.

→ **Multiple candidates**: Multiple different candidates are activated in response to input, and held in memory.

→ **Gradient activation**: Strength of activation of candidates depends on how closely each matches the cues - affected by fine phonetic detail.

→ **You can go back if you need to**: The current context could trigger reanalysis / reinterpretation of previous commitments (rhyme effects).