The path to words

Jonathan R. Brennan
Previously Speech and phonemes in 0.1 sec. What next? Accessing words.

Today The time-course of accessing word meanings, and what this can tell us about how words are mentally represented.

Next time Where are word meanings stored, and what can this tell us about what word meanings are?
Agenda

» 🔄 Outlining the time-course of visual word recognition
» Timing evidence for decompositional lexical representations
» The time-course of spoken word recognition
» Non-linear lexical representations?
Take-away Visual properties modulate activity ~0.1 s, while linguistic properties (like being a letter or word) modulate activity ~0.2 s
High Frequency (*water*)  
Medium Frequency (*amaze*)  
Low frequency (*yodel*)
Word frequency affects brain response early ~200 ms, and late, ~300–400 ms.

Later affect called N400: A negative going ERP to meaningful stimuli that peaks around 400 ms

**Challenge**: teasing apart lexical from sub-lexical properties like the frequency of letter combinations
Three stages in lexical access:

activation  
\[\downarrow\]  
competition  
\[\downarrow\]  
selection
**Question** Which stage of *lexical access* is reflected in brain activity at 300–400 ms?

» *Activation* will be indicated by faster brain responses for *high probability* words

» *Competition* will be indicated by slower brain responses to word with *many neighbours*

\[ \uparrow \text{probability} \propto \downarrow \text{neighbours} \]

High probability & density (BELL, LINE) vs. Low probability & density (PAGE, DISH)

*Pylkkanen et al. 2002 Brain Lang*
<table>
<thead>
<tr>
<th></th>
<th>Word Frequency</th>
<th>Letter-pair frequency</th>
<th>Neighbours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higher</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BELL</td>
<td>27</td>
<td>4,209</td>
<td>13</td>
</tr>
<tr>
<td>LINE</td>
<td>55</td>
<td>11,918</td>
<td>17</td>
</tr>
<tr>
<td>Lower</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISH</td>
<td>16</td>
<td>5,932</td>
<td>6</td>
</tr>
<tr>
<td>PAGE</td>
<td>26</td>
<td>3,093</td>
<td>11</td>
</tr>
</tbody>
</table>

Pylkkanen & Marantz, 2003 *Trends Cogn Sci*
Clues to the timing of lexical activation form the $M_{350}$: A peak in MEG activity occurring 300–400 ms after visual word presentation.
The **M350** component measured by MEG indicates that lexical activation, but not lexical competition, has been completed by 300–400 ms
Clues towards when lexical activation begins from comparing words with pseudowords

"beak choot dirt heap jep pait park noke ripe tip vap "

MacGregor et al. 2012 Nat Comm
What sets *lexical activation* apart from *form properties*?

**Heteronyms** Words with identical spelling but different meanings

*Heteronym frequency ratio* relative frequency of one meaning or another can only be calculated as part of *lexical activation*!

\[
\frac{\text{wind}_{\text{breeze}}}{\text{wind}_{\text{clock}}} = \frac{10}{3} = 3.3, \quad \frac{\text{bank}_{\text{money}}}{\text{bank}_{\text{river}}} = \frac{8}{.1} = 80, \ldots
\]

So… test whether activity correlates with this *lexical property* or with a *form-based property* like letter sequence probability.
Form properties (orthographic sequence probability) @ 150–200 ms

Lexical properties (heronym frequency ratio) @ ≈350 ms
We can use this consensus to answer questions about lexical representations
Agenda

» Outlining the time-course of visual word recognition

» 🔔 Timing evidence for decompositional lexical representations

» The time-course of *spoken* word recognition

» Non-linear lexical representations?
This is a WUG

Now there is another one.
There are two of them.
There are two _____.

Wug test!
This is a GUTCH.

Now there is another one.
There are two of them.
There are two _____.
This is a man who knows how to RICK.
He is Ricking. He did the same thing yesterday. What did he do yesterday?
Yesterday he __________.
This is a man who knows how to SPOW.
He is SPOWING. He did the same thing yesterday. What did he do yesterday?
Yesterday he ________.
Consensus that inflection is based on *compositional rules*...

But **irregular morphology** does not follow simple compositional rules

- **Past**: teach → taught (*teached*), give → gave (*gived*), find → found (*finded*)
- **Plural**: mouse → mice (*mouses*), fish → fish (*fishes*)

Also **derivational words** mix composition with idiosyncratic meanings (Compare "adap-er" ← "adapt" with "farmer" ← "farm")

**Dual-route** theories of lexical access state that *regular* morphology is recognized in pieces, or *decomposed*, but *irregular* morphology is recognized as *whole-words*.

On the other hand, **Full decomposition** theories hold that all complex words are recognized in terms of their parts, e.g. *tought = teach + [past]*

Taft 1979 Mem Cognit; Pinker & Ullman 2002 Trends Cog Sci
**Priming** is a speed-up in processing the second of two related stimuli

- Regular morphology lead to *repetition priming* ("dated" primes "date")
- Irregular morphology does not lead to priming → more evidence for *whole*-w … *wait!*

Stockall & Marantz 2006 *Mental Lexicon*
**Priming** is a speed-up in processing the second of two related stimuli.

dated... DATE... [word or not?]
church... DATE... [word or not?]

» Regular morphology lead to repetition priming ("dated" primes "date")

» Irregular morphology does not lead to priming → more evidence for whole-w ... wait!

\[
give \quad \rightarrow \quad GAVE
\]

\[
\begin{align*}
\text{semantic similarity} & \quad \rightarrow \quad \text{phonological similarity} \\
\text{priming} & \quad \leftarrow \quad \text{competition}
\end{align*}
\]
If only there were some way to test whether lexical activation has been *primed* prior to the influence of competition effects... the M350
<table>
<thead>
<tr>
<th>condition</th>
<th>prime</th>
<th>target</th>
</tr>
</thead>
<tbody>
<tr>
<td>low overlap</td>
<td>teach</td>
<td>taught</td>
</tr>
<tr>
<td>high overlap</td>
<td>give</td>
<td>gave</td>
</tr>
<tr>
<td>regular</td>
<td>date</td>
<td>dated</td>
</tr>
<tr>
<td>similar semantics &amp; form</td>
<td>boil</td>
<td>broil</td>
</tr>
</tbody>
</table>

Stockall & Marantz 2006 *Mental Lexicon*
Table 4. Mean M350 Latencies (in ms) Averaged Across Items in Experiment 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Overall</th>
<th>Rel.(SD)</th>
<th>Unrel.(SD)</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular (eg., jump–jumped)</td>
<td>338.4(26.0)</td>
<td>331.4(24.2)</td>
<td>345.4(27.1)</td>
<td>-16*</td>
</tr>
<tr>
<td>Hi-Overlap Irr (eg., give–gave)</td>
<td>342.3(44.6)</td>
<td>327(40.8)</td>
<td>357.6(44.9)</td>
<td>-30.6*</td>
</tr>
<tr>
<td>Lo-Overlap Irr (eg., teach–taught)</td>
<td>347.3(32.7)</td>
<td>334.4(28.2)</td>
<td>360.2(33.1)</td>
<td>-25.8*</td>
</tr>
<tr>
<td>+S+O-M (eg., boil–broil)</td>
<td>358.13(28.1)</td>
<td>358.7(36.9)</td>
<td>357.2(23.6)</td>
<td>+1.5</td>
</tr>
<tr>
<td>Average</td>
<td>345.92</td>
<td>335.72</td>
<td>355.1</td>
<td>-17.73</td>
</tr>
</tbody>
</table>

*p<0.05
SD = standard deviation

Table 5. Mean Lexical Decision Time (ms) Averaged Across Items in Experiment 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Overall</th>
<th>Rel.(SD)</th>
<th>Unrel.(SD)</th>
<th>Diff.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular (eg., jump–jumped)</td>
<td>637.85(51.12)</td>
<td>649.43(51)</td>
<td>625.33(51.37)</td>
<td>-24.1</td>
</tr>
<tr>
<td>Hi-Overlap Irr (eg., give–gave)</td>
<td>600.48(46.15)</td>
<td>587.9(45.53)</td>
<td>613.06(44.05)</td>
<td>-25.16</td>
</tr>
<tr>
<td>Lo-Overlap Irr (eg., teach–taught)</td>
<td>600.32(44.01)</td>
<td>586.5(36.9)</td>
<td>614.14(46.8)</td>
<td>-27.64</td>
</tr>
<tr>
<td>+S+O-M (eg., boil–broil)</td>
<td>666.05(55.85)</td>
<td>666.13(50.44)</td>
<td>665.96(61.99)</td>
<td>-0.17</td>
</tr>
<tr>
<td>Average</td>
<td>624.10(55.75)</td>
<td>613.83(55.48)</td>
<td>634.18(54.41)</td>
<td>-13.78</td>
</tr>
</tbody>
</table>

*p<0.05
SD = standard deviation
Irregular verbs show **morphological priming** at the M350, but not in reaction times. Consistent with *shared morphological structure* for irregulars, just like regular verbs!
Bigger picture: comprehension as rebuilding

Information scattered across physical stimulus and transducing neurons

» Speech: place-code for different spectro-temporal distribute information about the same features/words across thousands of neurons

» Reading: retinal place-code similarly distributes information for the same letter → word across many thousands of neurons

Picture credit: Dehaene 2010 Reading in the Brain Ch. 3
Bigger picture: comprehension as rebuilding

Information scattered across physical stimulus and transducing neurons

Aligns with morphological representations that are decomposed to facilitate look-up based on parts (not wholes) which are then recombined

e.g. Taft & Forster 1975 J Verb Learn Behav; Halle & Marantz 1993 View from Bldg 20
Distinct stimulus properties affect these separate stages, we can tease out brain activity associated with each property to estimate *where* and *when* each stage occurs.

<table>
<thead>
<tr>
<th>stage</th>
<th>stimulus feature</th>
<th>when</th>
<th>where</th>
</tr>
</thead>
<tbody>
<tr>
<td>decomposition</td>
<td>pseudo-affixation</td>
<td>150–200 ms</td>
<td>left inferior temporal</td>
</tr>
<tr>
<td>look-up</td>
<td>family size</td>
<td>250–350 ms</td>
<td>left superior temporal</td>
</tr>
<tr>
<td>activation-selection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>recombination</td>
<td>semantic coherence</td>
<td>~400 ms?</td>
<td>inferior frontal?</td>
</tr>
</tbody>
</table>
If we decompose words automatically, then it should apply just as often to pseudo-affixed items as to truly affixed complex words.

Can check this with pseudo-morpheme priming!

<table>
<thead>
<tr>
<th>condition</th>
<th>prime</th>
<th>target</th>
</tr>
</thead>
<tbody>
<tr>
<td>true affix</td>
<td>alarming</td>
<td>ALARM</td>
</tr>
<tr>
<td>fake affix</td>
<td>archer</td>
<td>ARCH</td>
</tr>
<tr>
<td>no affix</td>
<td>brothel</td>
<td>BROTH</td>
</tr>
</tbody>
</table>

Rastle et al. 2000 *Cognition*; Lehtonen et al. 2011 *J Cog Neurosci*
Look-up is affected by the number of morphologically related items that are competing for activation, or the family entropy of the word. This correlates with left temporal brain activity at ~250 ms.

Recombination is sensitive to the semantic coherence of root + suffix, which correlates with activity >350 ms.
Summary so far

Identify **neural time-course** for reading: occipital lobe at ~100 ms → inferior temporal lobe by ~200 ms ("M170") → superior temporal lobe by ~300 ms ("N400", "M350")

In parallel, identified **stages of lexical processing**, from recognizing letters that comprise morphemes → **activating** morphemes → reconciling **competition** and selecting the best fit → finally **recombining** morphemes together into **complex words**

Aligned neural time-course with lexical stages via **correlations** with stimulus information: form-properties like probability of some letter sequence) (~170 ms), lexical properties like the frequency ratio for ambiguous words (~300 ms), and semantic properties like the fit between a stem and its affix (>300 ms)

Example of using these details to test what sorts of information are activated for debated representations, like **irregular verbs**
Agenda

» Outlining the time-course of visual word recognition
» Timing evidence for decompositional lexical representations
» 📽️ The time-course of *spoken* word recognition
» Non-linear lexical representations?
Results from reading connect well with spoken word recognition?

Long story short: we see the same general time-course when we look at speech, not reading.

As before, define different stimulus properties that reflect separate stages of processing and correlate with MEG activity while *listening* to words.
We also decompose spoken words for lexical access.

Unexpected phonemes increase auditory cortex activity:

\[
\begin{align*}
[b... \ ɚ ...b... n] & \Leftrightarrow ? \wedge \\
[b... \ ɚ...b...l] & \Leftrightarrow ? \wedge ?
\end{align*}
\]

<table>
<thead>
<tr>
<th></th>
<th>expected ending</th>
<th>unexpected ending</th>
</tr>
</thead>
<tbody>
<tr>
<td>one morpheme</td>
<td>bourbon</td>
<td>burble</td>
</tr>
<tr>
<td>two morphemes</td>
<td>breather</td>
<td>breathing</td>
</tr>
</tbody>
</table>

...and do so more if the phonemes are in a new morpheme.
Agenda

» Outlining the time-course of visual word recognition
» Timing evidence for decompositional lexical representations
» The time-course of *spoken* word recognition
» 🎨 Non-linear lexical representations?
English regular morphology is *linearly separable*: *separ* – *able*

Evidence for decomposition in irregulars suggests that linearity is not a condition on morphological processing...

...test this more robustly in languages with non-concatenative morphology like *Arabic*
<table>
<thead>
<tr>
<th>Root: <em>ktb</em></th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>kaataba</strong></td>
<td><em>corresponded</em></td>
</tr>
<tr>
<td><strong>kutiba</strong></td>
<td><em>was written</em></td>
</tr>
<tr>
<td><strong>kitaab</strong></td>
<td><em>book</em></td>
</tr>
<tr>
<td><strong>kutub</strong></td>
<td><em>book</em></td>
</tr>
<tr>
<td><strong>kuttaab</strong></td>
<td><em>writers</em></td>
</tr>
</tbody>
</table>

Test for competition effects based on **whole-word** or **root only**

Gwilliams & Marantz 2015 *Brain Lang*
Take-aways:

» Word recognition in 1/3 sec

» Stages as a window into representations: (visual) word recognition is maximally decomposed

» Similarities between spoken/visual despite different pathways

» Lots of questions about generalizability of these findings
  » visual vs. spoken recognition
  » regular & irregular morphology
  » concatenative and non-concatenative morphology