Clusters of Information in the Reading Brain

Davide Crepaldi

International School for Advanced Studies (SISSA), Trieste
Milan Center for Neuroscience (NeuroMi)

(Most work done at Milano Bicocca)

miLanguage 2015
The claim

- Scripts can be seen as fully-fledged visual systems
- They can be studied as such, without language
- The way we learn to deal with them can be captured through statistical learning
- The way we learn to map them onto language can be captured through statistical learning
The claim

- Scripts can be seen as fully-fledged visual systems
- They can be studied as such, without language
- The way we learn to deal with them can be captured through statistical learning
- The way we learn to map them onto language can be captured through statistical learning
Different parts of the form convey different parts of the meaning (DEAL-ER vs. ELEPHANT)

These parts are called morphemes

Morphemes can also be words in their own right (e.g., CAT), but not always are (e.g., -ER)

All about the relationship between form and meaning
Masked priming

500 ms

30-55 ms

2000 ms

dealer

DEAL

press button

Davide Crepaldi

Clusters of Information in the Reading Brain
Masked priming

Davide Crepaldi

Clusters of Information in the Reading Brain
Corners who corn

dealer–DEAL vs. corner–CORN vs. dialog–DIAL
Corners who corn

dealer–DEAL vs. corner–CORN vs. dialog–DIAL

(Rastle et al., 2004)
The gold in the corner

- Theoretical linguistics define morphology as correlation between form and meaning, but form has a stance on its own.

- Stress the regularity issue: if form is critical, how could we ever parse things like BOUGHT, or FELL?
Masked irregular priming

fell–FALL vs. fill–FALL vs. pair–FALL

(Crepaldi et al., 2010a)
What’s regularity, really?

- FEED, BLEED, MEET and BREED all become FED, BLED, MET and BRED

- SPEND, SEND, LEND and BEND all become SPENT, SENT, LENT and BENT

- Is the system picking up this regularity in pattern alternation? (Marcus et al., 1992)
Not just morpho-orthography

bell-BALL vs. bull-BALL vs. rope-BALL
All about form?

bell-BALL vs. bull-BALL vs. rope-BALL

(Crepaldi et al., 2010a)
Two mechanisms

- Irregular priming is sensitive to semantics (works for fell–FALL, but not for bell–BALL), while regular priming is not (works for dealer–DEAL, but also for corner–CORN).

- Regular priming is based on parsability (it works for BROTH-ER, but not for BROTH-EL), while irregular priming is not.

- A tale of two morphologies.
The lemma model

(Crepaldi et al., 2010a)
Two predictions

- Regular priming should be larger than irregular priming because it has two sources vs. one.

- Regular priming should emerge earlier than irregular priming because it is based on an earlier level of processing.
ASKS is more than SEEN

(asks–ASK vs. ends–ASK) vs. (seen–SEE vs. mind–SEE)
ASKS is more than SEEN

(asks–ASK vs. ends–ASK) vs. (seen–SEE vs. mind–SEE)

(Rastle et al., 2015)
ASKS is earlier than SEEN

(Rastle et al., 2015)
Morpheme positional constraints

- KINDNESS and NESSKIND
- PREHEAT and HEATPRE
- CATWALK and WILDCAT
- OVERHANG and HANGOVER
Blind to suffixes

- (GASFUL vs. GASFIL) vs. (FULGAS vs. FILGAS)

(Crepaldi et al., 2010b)
Blind to prefixes

- (PREHOSE vs. PLEHOSE) vs. (HOSEPRE vs. HOSEPLE)
Stems everywhere

(fishgold–GOLDFISH vs. kacnvrqw–GOLDFISH) vs. (tonebari–BARITONE vs. suyzchmw–BARITONE)

(Crepaldi et al., 2013)
How far do these constraints go?

- Word boundaries vs. local constraints
- All-or-none vs. graded constraints
My stronger supporters

Davide Crepaldi

Clusters of Information in the Reading Brain
Baboons

- Baboons can learn English words
- Baboons have no human–like language
- So, baboons think I’m right!
Baboons learn words

(Grainger et al., 2012)
Baboons extract knowledge about letter stats

**A**  All first words

- **Nonword responses (%)**
- **First words**
- **Nonwords**
- **Difference**
- **DAN**
- **ART**
- **CAU**
- **DOR**
- **VIO**
- **ARI**

**B**  Last 50 first words

- **Nonword responses (%)**
- **First words**
- **Nonwords**
- **Difference**

Davide Crepaldi  Clusters of Information in the Reading Brain
Baboons extract knowledge about letter stats

A

Nonword accuracy (Monkeys)

B

Nonword accuracy (Humans)

Orthographic distance to words (OLD20)

Orthographic distance to words (OLD20)
### Transparent stems?

<table>
<thead>
<tr>
<th>Study</th>
<th>Experimental</th>
<th>Control</th>
<th>Mean difference</th>
<th>MD</th>
<th>95%−CI</th>
<th>W(fixed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total Mean SD</td>
<td>Total Mean SD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>50 560 40</td>
<td>50 572 34</td>
<td></td>
<td>-12.00</td>
<td>[-26.55; 2.55]</td>
<td>16.6%</td>
</tr>
<tr>
<td>2</td>
<td>34 643 56</td>
<td>34 648 50</td>
<td></td>
<td>-5.00</td>
<td>[-30.23; 20.23]</td>
<td>5.5%</td>
</tr>
<tr>
<td>3</td>
<td>46 604 34</td>
<td>46 624 35</td>
<td></td>
<td>-20.00</td>
<td>[-34.10; -5.90]</td>
<td>17.6%</td>
</tr>
<tr>
<td>4</td>
<td>57 578 36</td>
<td>57 590 29</td>
<td></td>
<td>-12.00</td>
<td>[-24.00; 0.00]</td>
<td>24.3%</td>
</tr>
<tr>
<td>5</td>
<td>71 599 29</td>
<td>71 609 31</td>
<td></td>
<td>-10.00</td>
<td>[-19.87; -0.13]</td>
<td>36.0%</td>
</tr>
<tr>
<td>Fixed effect model</td>
<td>258</td>
<td>258</td>
<td></td>
<td>-12.31</td>
<td>[-18.23; -6.38]</td>
<td>100%</td>
</tr>
</tbody>
</table>

*Heterogeneity: I−squared=0%, tau−squared=0, p=0.7945*
Form as a cue to meaning

CORN

- Get all words that start with CORN
- Take their semantic representations
- Compute their similarity
- Take the mean

Orthography–Semantic Transparency (OSC)

- The internal consistency of the “form” family in terms of meaning
- How similar in meaning are words similar in form
- How good of a cue to meaning is form
OSC gets unique variance

Table 6. Results of the regression analysis on the lexical decision latencies extracted from the BLP for a large set of random words

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>6.5922</td>
<td>0.0109</td>
<td>602.89</td>
<td>0.0001</td>
</tr>
<tr>
<td>Word frequency</td>
<td>-0.0308</td>
<td>0.0009</td>
<td>33.41</td>
<td>0.0001</td>
</tr>
<tr>
<td>Word FS</td>
<td>-0.0041</td>
<td>0.0021</td>
<td>1.97</td>
<td>0.0495</td>
</tr>
<tr>
<td>Word length</td>
<td>0.0035</td>
<td>0.0013</td>
<td>2.74</td>
<td>0.0061</td>
</tr>
<tr>
<td>OSC</td>
<td>-0.0254</td>
<td>0.0066</td>
<td>3.84</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

(Marelli et al., 2014)
The claim

- Scripts can be seen as fully-fledged visual systems
- They can be studied as such, without language
- The way we learn to deal with them can be captured through statistical learning
- The way we learn to map them onto language can be captured through statistical learning
The claim

- Scripts can be seen as fully-fledged visual systems
- They can be studied as such, without language
- The way we learn to deal with them can be captured through statistical learning
- The way we learn to map them onto language can be captured through statistical learning
Acknowledgments

- The Ministry of Education, University and Research in Italy; the ESRC, the British Academy and the Leverhulm Trust in the UK.
- Simona Amenta (Milano Bicocca), Marco Marelli (Trento), Kathy Rastle (Royal Holloway, UK), Colin Davis (Bristol, UK), Max Coltheart and Lyndsey Nickels (CCD, Australia), and Aureliu Lavric (Exeter, UK).

davide.crepaldi@sissa.it
www.davidecrepaldi.net


