Morphology in the Decomposing Brain: Correlational Analyses of Single Trial MEG Data

Alec Marantz, Olla Somolyak, Ehren Reilly
Bill Badecker, Asaf Bachrach
KIT/NYU MEG Joint Research Lab
Departments of Psychology and Linguistics NYU

Map

- Simple Models of Lexical Access Rejected
  - simple dual route model (Pinker)
  - simple obligatory decomposition model
  - simple whole word access model
- Real Issues Identified
  - modality specific access “lexicon”?  
- what properties modulate affix-stripping/decomposition, stem access and recombination on an obligatory decomposition model?

Completed Experiment
- evidence for, at least, full decomposition or interactive dual route model
- no support for whole word access route
- support for early effects of surface frequency of affixed forms relative to stem frequency at decomposition stage
- Experiment Underway
  - tracking the –able in amiable and stable

Competing Models of Lexical Access – according to Pinker’s Words and Rules Cartoon Version

Saturday Morning Models of Lexical Access

- Full storage model: all complex words (walked, taught) stored and accessed as wholes
  - only surface frequency effects predicted
- Full decomposition model: no complex words stored and accessed as wholes
  - only stem frequency effects predicted

Dual Route Model: irregular complex forms are stored and accessed as wholes; regular complex forms are not:
- surface frequency effects for irregulars (and high frequency regulars)
- stem frequency but no surface frequency effects on access for regulars
Facts Support Dual Route Model, if Alternatives are these Cartoon Versions

- Fact: Stem Frequency effects in access for complex words
- Fact: These effects are not attributable to post-access decomposition
  - masked priming studies showing morphological (stem) priming when neither form nor semantic priming are found

- But, fact: Surface frequency effects in lexical access are found in wide variety of cases, including completely regular morphology (e.g., for most inflected words in Finnish)

Problems for Dual Route Model

- The representation of irregular derived or inflected forms must be complex
  - from the grammatical point of view, felt is as complex as walked
    - e.g., behavior with respect to do support, impossibility of further inflection or derivation, rigidity of meaning...
  - from the psycho and neurolinguistic point of view, irregulars contain the stem in the same way that regulars do
    - taught-teach identity priming in long-lag priming and for M350 brain response

Whole Word “Representations” for Regulars

- Surface frequency effects on access are seen for a variety of completely regular derivations and inflections.
- Obligatory decomposition:
  - surface frequency effects could be tied to decomposition (the more you’ve decomposed a particular letter/sound sequence into stem and affix, the faster you are at it) and/or
  - recombination (the more often you’ve put together a particular stem and affix, the faster you are at it)
- Against Pinker’s dual route model, such effects imply representation of whole word as complex structure, regardless of regularity.

Lexical Morphology and Lexical Access

Comparing surface and base frequency effects: Level 1 vs Level 2 Morphology:
surface frequency effects even for transparent productive regular morphology, and for same words that yield base frequency effects

<table>
<thead>
<tr>
<th>Suffix</th>
<th>High frequency roots</th>
<th>Low frequency roots</th>
<th>High frequency whole words</th>
<th>Low frequency whole words</th>
</tr>
</thead>
<tbody>
<tr>
<td>-less</td>
<td>2.36</td>
<td>1.32</td>
<td>1.20</td>
<td>0.15</td>
</tr>
<tr>
<td>-ity</td>
<td>2.32</td>
<td>1.16</td>
<td>1.82</td>
<td>0.60</td>
</tr>
<tr>
<td>-ation</td>
<td>2.23</td>
<td>1.03</td>
<td>1.48</td>
<td>0.26</td>
</tr>
</tbody>
</table>

FIG. 2. Response times for Whole Word-Control Pairs in Experiment 1.
**Base Frequency Effect:**
here only for productive transparent “-less” type

![Bar chart showing mean response times for root-contrast pairs in Experiment 1.](image)

**“Representations”**

- Saying that every combination of morphemes in perception or production, no matter how regular, leaves a trace in the language system of the speaker is saying that frequency information is part of the grammar and that all combinations of morphemes are stored in some sense.

**Beyond the Cartoons**

- *walked* may “stored” as a complex form with a certain frequency in the same way that a saying like, *And now for something completely different,* is.
- Both must be composed with the grammar when heard or produced, but both may have frequency and special meaning information associated with them that have no implications for the grammar whatsoever.

**Realistic Full Decomposition Models Must...**

- Recognize that complex words, both regular and irregular, are stored in some sense, leading to surface frequency effects (but this is true of phrases and sentences as well as words and holds no implications for the grammar)
- Investigate the role of surface frequency in
  - decomposition
  - stem access
  - recombination

**Realistic Dual Route Models Must...**

- Recognize that all complex forms must be representationally complex, containing structures of morphemes and contrasting with monomorphemic constituents
- Explore the possible existence of stored “whole word” word form representations at modality-dependent “access lexicons” to distinguish themselves from obligatory decomposition models which deny such representations
**Simplistic Prediction of Hay Model**

- Affix dominant words should show surface frequency effects but no stem frequency effects since they are accessed via the whole word route.
- Stem dominant words should show stem (cumulative) frequency effects since they are accessed via the decomposition route.

**Effect of “Dominance” on Lexical Access:**
view from interactive dual route model

- Jen Hay: importance of the relative frequency of a morphologically complex form with respect to the frequency of its stem.
- Complex words with high frequencies relative to their stems are “affix-” or “surface dominant”; those with low frequencies are “stem” or “base dominant”
- Hay: affix dominance leads to difficulty in parsing/decomposition, thus reliance on whole-word recognition and suppression of decomposition in favor of whole-word route.

**How can we distinguish these accounts of RT differences?**

Reilly, Badecker & Marantz 2006 (Mental Lexicon):
by determining the point of lexical access via brain monitoring.

---

**Taft (2004): "Morphological Decomposition and the Reverse Base Frequency Effect"**

Obligatory decomposition makes same predictions as Hay for RT

- Base frequency effects...
  - RT to complex word correlates with freq of stem
  - ...reflect accessing the stem of morphological complex forms whereas
- Surface frequency effects...
  - RT to complex word correlates with freq of complex word
  - ...reflect the stage of checking the recombination of stem and stripped affix for existence and/or well-formedness.
Sequential processing of words

Tracking the time course of word recognition with MEG

Linna Pyllänen and Alex Marantz

1Department of Linguistics, New York University, 721 Broadway, 4th Floor, New York, NY 10003, USA
2Department of Linguistics and Philosophy, MIT-MEG Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Repetition Frequency

1: 700
2: 140
3: 30
4: 6
5: 1
6: .2

Frequency Category (Frequent -- Infrequent)

Latency of M350 sensitive to lexical factors such as lexical frequency and repetition: reflects stage of lexical access

Stimuli: 3 Lexical Categories

- Nouns: singular/plural
  - bone
  - bones
- Verbs: stem/progressive
  - chop
  - chopping
- Adjectives: adjective/-ly adverb
  - clear
  - clearly

Experiment: parallel behavioral and MEG processing measures

- Lexical Manipulation (Baayen, Dijkstra & Schreuder, 1997, JML)
  - Lemma/stem frequency (CELEX database)
  - Stem vs. affix dominance

<table>
<thead>
<tr>
<th>Stem Frequency</th>
<th>Stem Dominant</th>
<th>Affix Dominant</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>desk – desks</td>
<td>crop – crops</td>
</tr>
<tr>
<td>Mid</td>
<td>deck – decks</td>
<td>cliff – cliffs</td>
</tr>
<tr>
<td>Low</td>
<td>chef – chefs</td>
<td>chord – chords</td>
</tr>
</tbody>
</table>

Experiment: behavioral measures

- Reliable effect of stem frequency in RT

Response Time (ms)

High | Medium | Low
Experiment: behavioral measures

- Interacting effects on RT of affixation (base vs. affixed) and dominance (base-dominant vs. affix-dominant)

Analysis of M350 peak latency (brain index of lexical access)

- Reliable effect of Stem frequency for unaffixed words and for affixed words

Analysis of M350 peak latency

- Reliable effect of Affixation (base vs. affixed)

Analysis of M350 peak latency

- No effect of Dominance (base-dominant vs. affix-dominant) on M350 peak latency

Analysis of M350 peak latency

- No interaction between Dominance (base-dominant vs. affix-dominant) and Affixation (base vs. affixed)

Analysis of M350 peak latency

- Evidence that early stages of access for affixed words is based on full parsing: Stem frequency affects M350/lexical access while whole word frequency affects post-access (recombination) stage of word recognition.
Problem with this Conclusion

• No acknowledgement of the effects of dominance and/or surface frequency on parsing stage of decomposition

Possible effects of dominance at different stages in word recognition

<table>
<thead>
<tr>
<th></th>
<th>parsing affix</th>
<th>stem access</th>
<th>recombination and checking</th>
<th>RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>affix dominant merely</td>
<td>harder, since tighter connection (possibly only at high surface freq values)</td>
<td>based on stem frequency, possibly speeded if high conditional probability of stem given affix</td>
<td>faster, should correlate with surface frequency at high freq values</td>
<td>might correlate (-) with surface frequency, given speed up in recombination</td>
</tr>
<tr>
<td>stem dominant sanely</td>
<td>easier than for affix dominant (lower transition probability)</td>
<td>based on stem frequency</td>
<td>at lower surface freq values, no effect of surface freq</td>
<td>at lower surface freq value, should correlate (-) with stem freq</td>
</tr>
</tbody>
</table>

RMS Correlations Across Subjects

• Let’s examine these possibilities with correlational analyses

• For some set of sensors, calculate at each time point in each experimental "epoch" the root mean square (RMS) = the square root of the mean of the squares of the values at each sensor (after normalization of values)
• So, for each subject, for each item, an RMS "wave" can be provided for the correlational analysis
• At each time point, the RMS value for each stimulus is correlated with a stimulus variable

Grand Average All Stimuli All Subjects (11)

M170 Sensors Chosen on the basis of field pattern, subject by subject
M350 sensors chosen subject by subject

M170 Correlation with Dominance: Significant “parsing” effect

No M170 word frequency effect for unaffixed words

No M170 Word/Non-Word “lexicality” effect

Although surface frequency correlates with affix dominance in these words, no surface frequency M170 effect

Recombination Effect?: Correlation with Conditional Probability of Stem, Given Affix, for Affixed Words
Summary of Correlations

- **M170 amplitude:**
  - Dominance
    - For affixed words: \( r = 0.07; p < 0.05 \)
    - Affixed surface dominant words: \( r = 0.11; p < 0.01 \)
- **M350 Latency:**
  - Stem Frequency
    - For all words: \( r = -0.06; p < 0.05 \)
    - For affixed words: not significant (neg. cor.; \( p = 0.21 \))
    - For unaffixed words: \( r = -0.07; p < 0.05 \)
- **Post M350 activity and affix frequency:**
    - Left Sensors:
      - W1: \( r = 0.07; p < 0.05 \)
      - W2: \( r = 0.08; p < 0.01 \)
      - W3: \( r = 0.08; p < 0.01 \)

Evidence for an orthographic word form lexicon

- Frequency of stem relative to full affixed form – affix dominance – correlates with M170 amplitude; implies access to some kind of stem representation
- Zweig & Pylkkänen (2006) show M170 effect of decomposition in the contrast between *farmer* (complex) and *winter* (simple), where the contrast implies access to a representation of *farm* at the M170

Zweig & Pylkkänen (2006)

- Bimorphemic: refill, Monomorphemic Orth: resume
- Bimorphemic: farmer, Monomorphemic Orth: winter

Modality-Specific Access Lexicon?

- **For:** “Parsing” at the M170 may require access to “lexicalized” word forms (or to high-n n-grams, or what)?
- **For:** Dominance effects at the M170 suggest frequency information associated with word-forms
  - dominance reflects the conditional probability of the affix given the stem
- **Against:** At the M170 it’s difficult to find word-form frequency effects or lexicality effects

Sophisticated Interactive Dual-Route Models predict Whole Word Word Form Effects

- However, there is no evidence from the current study that whole word word forms are available for complex words
- To the contrary, complex words with high affix dominance – the very words that should have whole word word forms according to Hay – show the biggest effect of complexity at the M170, where word form effects are expected
New Experiment:
Investigate Factors Influencing Stages of Processing for Morphologically Complex Words AND Provide Neural Indices of Morphological Decomposition for Controversial Cases

- Initial parsing and decomposition
  - affix frequency?
  - transition probabilities, computed over morphemes or over strings?
- stem activation
  - frequency
  - “family” structure (family size, frequency, semantic transparency, etc.)
  - conditional probability from affix recognition?
- re-merger of pieces
  - surface frequency and conditional probabilities?
  - evaluation of combined structure
  - semantic transparency?

status of bound stems

- durable
  - same root in duration
  - predicts durability
- amiable
  - no other uses of root
  - but, predicts amiability (stability)
- cable
  - false parse, compared to stable

tracking the -able in amiable

- If words like tolerable with a recurring root and amiable with a unique root nevertheless are parsed and computed as is workable with a word root, then
- M170 “parsing” effects should be visible for these “opaque” words, since effects are strong for affix-dominant words
- M350 effects should be modulated by morphological family and conditional probability of word given the affix
- correlational contrasts with “matched” unaffixed words should be seen at all stages of access

Categories of Affixed Words for New Experiment

- 1. Word-Affix
  - taxable
- 2. Root-Affix
  - tolerable
- 3. Pseudo-Affix (mono-morphemes)
  - capable

- Morphological parsing as from English Lexicon Project

Nine Affixes

- able
- ary
- ant
- ity
- ate
- ic
- er
- al
- ion

More Examples

<table>
<thead>
<tr>
<th>Word-Affix</th>
<th>Root-Affix</th>
<th>Pseudo-Affix</th>
</tr>
</thead>
<tbody>
<tr>
<td>rationality</td>
<td>quantity</td>
<td>vicinity</td>
</tr>
<tr>
<td>destroyer</td>
<td>sorcerer</td>
<td>character</td>
</tr>
<tr>
<td>classic</td>
<td>specific</td>
<td>empiric</td>
</tr>
</tbody>
</table>
The 27 subgroups...

- All groups are equivalent and well-distributed over:
  - length
  - mean bigram count
  - frequency
- All word-affix groups are well-distributed over:
  - surface vs stem frequency dominance

Matched Words

- Matched on frequency of ending:
  - Calculated affix frequencies from CELEX database
  - For each affix, found a list of mono-morphemes with roughly the same ending frequency. (log frequency of +/- 0.3 from affix log frequency)
  - Chose 18 matched words for each affix that had same parameter distributions as affixed words.

Examples

- er
  - evening, lightning, mediocre
- ity
  - caricature, terrain, pertain
- able
  - sentence, thought, profound

Parsability

- Affixed and Matched words are equally distributed over “parsability” – transitional probability between the last 2 letters of the stem and the first two letters of the affix (or affix-match).

- Example – for “ability” this would be: given that you see “il” 5 letters from the end of a word, what is the probability that “it” will follow it?

Analysis in brain space

- A distributed source model for each subject is constructed based on the structural MRI of the S’s brain
- Brain regions and time windows of interest (ROIs) are identified based on the grand average of responses over all stimuli
- Correlations are computed with stimulus variables within each ROI

ROIs (green) for M170, left and right hemispheres for one subject
**Preliminary results (N=2)**

- Word+affix words yield more/later M170 activity than root+affix words
- And, M170 responses correlate with affix dominance, as in previous experiment
- Lemma to suffix transitional probability correlations with M170 activity for root+affix words, not word+affix words

**Representative single subject correlations**

- **M170 left hemisphere:**
  - Root+affix:
    - Latency by TransProb.lemma: $r = 0.3193$ $p = 0.022$
  - Word+affix:
    - Latency by TransProb.lemma: $r = 0.2739$ $p = 0.052$
- **M170 right hemisphere:**
  - Root+affix:
    - Latency by TransProb.lemma: $r = 0.2739$ $p = 0.052$
  - Word+affix:
    - Amp by Surface Dominance: $r = -0.2473$ $p = 0.074$
    - Amp by Stem Word Freq: $r = 0.2613$ $p = 0.058$

**Leading interpretation:**

- For word+affix words, the most important variable is the relative frequency of the stem as an independent word compared to the frequency of the word+affix; here affix dominance inhibits parsing
- For root+affix, the predictability of the affix given the stem (like affix dominance) seems to affect parsing
- More evidence in favor of full decomposition along with some sort of orthographic access lexicon for morphemes

**That's all, fo..(oh, you know the rest...)**
marantz@nyu.edu