Morphology as a simple ... or complex system

Jim Blevins (jpb39@cam.ac.uk)

Li1: (Sounds &) Words, 2016-03-10
Two perspectives on morphological organization

1. Bloomfieldian constructivism: The structure of a system is derivative of units and combinatoric principles that define individual forms.

2. Neo-Saussurean resonance: A system is characterized by an ‘organized complexity’ arising from “a sizable number of factors which are interrelated into an organic whole” (Weaver 1948: 539)
The decompositional programme

The inflectional forms are relatively easy to describe, since they occur in parallel paradigmatic sets; the traditional grammar of familiar languages gives us a picture of their inflectional systems. It may be worth noticing, however, that our traditional grammars fall short of scientific compactness by dealing with an identical feature over and over again as it occurs in different paradigmatic types. Thus, in a Latin grammar, we find the nominative-singular sign -s noted separately for each of the types *amīcus* ‘friend’, *lapis* ‘stone’, *dux* ‘leader’, *tussis* ‘cough’, *manus* ‘hand’, *faciēs* ‘face’, when, of course, it should be noted only once, with a full statement as to where it is and where it is not used. (Bloomfield 1933: 238, emphasis added)
The goal of an analysis is to eliminate redundancy by reducing systems to inventories of minimal units and general rules. Yet there is little evidence that morphological systems are organized in ways that facilitate disassembly and reassembly.

- It may be possible to disassemble a word form into minimal units.
- However there are often no effective criteria for selecting ‘correct’ segmentations into stable roots, stems and inflections.
- Moreover, in all but the simplest systems it is not possible to reconstitute word forms from inventories of minimal units, because disassembly has lost information about the distribution of units.
Challenges to decomposition

- The economy achieved by a decompositional account relies on the idealization that forms can be segmented into minimal parts and these parts can be assigned constant meanings.
- Yet this kind of approach confronts a number of challenges:
  1. Contrasts that discriminate a set of word forms do not always determine unique segmentations into minimal elements.
  2. Contrastive elements cannot always be associated with properties.
  3. Even in cases where a unique segmentation is possible and the elements can be associated with properties, an associative analysis may misrepresent a learning context as the structure of the system.
Problems of segmentation: Fusional indeterminacy

In a fusional language, if one seeks to arrive at constant segments in such a manner, conflicts arise in the placing of the cuts. One comparison of forms suggests one placement, while another comparison suggests another. Often, in fact, no constant segment can be isolated at all which corresponds to a given constant meaning. Situations of this kind often permit of more than one solution according to different manners of selecting and grouping environments. (Lounsbury 1953: 172)
Morphological processes in English

- What is the morphemic structure of English noun plurals?

  books:  book + s
  children:  child + ren
             childr + en
             child + r + en
  men:  man + a∼e
        men + Ø
  knives:  ...
Morphological processes in English

- What is the morphemic structure of English noun plurals?

  - **books:** book + s
  - **children:** child + ren
    - childr + en
    - child + r + en
  - **men:** man + a~e
    - men + Ø
  - **knives:** ...
Morphological processes in English

- What is the morphemic structure of English noun plurals?

  *books*:  $book + s$

  *children*:  $child + ren$

    *childr + en*

    *child + r + en*

  *men*:  $man + a\sim e$

    *men + Ø*

  *knives*:  ...
What is the morphemic structure of English noun plurals?

*books*:  \(book + s\)

*children*:  \(child + ren\)
\(childr + en\)
\(child + r + en\)

*men*:  \(man + a~e\)
\(men + \emptyset\)

*knives*:  ...
What is the morphemic structure of English noun plurals?

books:  \( \text{book} + s \)

children:  \( \text{child} + \text{ren} \)
- \( \text{childr} + \text{en} \)
- \( \text{child} + r + \text{en} \)

men:  \( \text{man} + \text{a} \sim \text{e} \)
- \( \text{men} + \emptyset \)

knives:  ...

Morphological processes in English

Decompositional analysis

Problems of segmentation
Morphological processes in English

- What is the morphemic structure of English noun plurals?

  - **books**: book + s
  - **children**: child + ren
    - *childr + en*
    - *child + r + en*
  - **men**: man + a∼e
    - *men + Ø*
  - **knives**: ...
Morphological processes in English

What is the morphemic structure of English noun plurals?

books:  book + s

children:  child + ren
           childr + en
           child + r + en

men:  man + a∼e
       men + Ø

knives:  ...
Morphological processes in English

▶ What is the morphemic structure of English noun plurals?

books: \( \text{book} + s \)

children: \( \text{child} + ren \)

\( \text{childr} + en \)

\( \text{child} + r + en \)

men: \( \text{man} + a\sim e \)

\( \text{men} + \emptyset \)

knives: ...
Morphological processes in English

▶ What is the morphemic structure of English noun plurals?

books: \( \text{book} + s \)

children: \( \text{child} + ren \)

\( \text{childr} + en \)

\( \text{child} + r + en \)

men: \( \text{man} + a \sim e \)

\( \text{men} + \emptyset \)

knives: ...
What is the morphemic structure of English noun plurals?

- **books**: book + s
- **children**: child + ren
  - childr + en
  - child + r + en
- **men**: man + a~e
  - men + Ø
- **knives**: ...
Morphological processes in English

What is the morphemic structure of English noun plurals?

*books:*  book + s

*children:*  child + ren
  childr + en
  child + r + en

*men:*  man + a~e
  men + Ø

*knives:*  ...
Problems of segmentation: Romance stems

- The segmentation of Romance stems remains recalcitrant:

  The order of morphemes is fixed: (derivational prefix(es)) + lexical stem + theme vowel + tense marker (sometimes including an empty morph) + person marker. Some forms, however, have fused in the course of history and a neat segmentation is not always possible. The preterit is the most difficult paradigm to analyse, since the theme vowel is sometimes indistinguishable, and segmenting the second and third person plural markers in the regular way, /-is, -n/, leaves an awkward residue that occurs nowhere else in the system. (Green 1997: 99)

- There is still no accepted solution to this problem (Spencer 2012).
The difficulties that arise in dividing forms into “minimal segments” have been a prominent concern of associative models, from Harris (1942) and Hockett (1947) through Spencer (2012).

The fact that these difficulties arise in fully discriminated subsystems suggests that they do not reflect underlying variability in the language but are artifacts of an inapplicable method of analysis.
First declension partitive constructions in Estonian

<table>
<thead>
<tr>
<th></th>
<th>Nom SG</th>
<th>Part SG</th>
<th>Gen SG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sg</td>
<td>kukk</td>
<td>lukk</td>
<td>pukk</td>
</tr>
<tr>
<td></td>
<td>kukke</td>
<td>lukku</td>
<td>pukki</td>
</tr>
<tr>
<td></td>
<td>kuke</td>
<td>luku</td>
<td>puki</td>
</tr>
<tr>
<td></td>
<td>‘rooster’</td>
<td>‘lock’</td>
<td>‘trestle’</td>
</tr>
</tbody>
</table>

- *kukke, lukku, pukki* and *sukka* all realize partitive singular.
  - Each form contains a (lexically-specified) theme vowel that distinguishes them from the consonant-final nominative singular.
  - Each form also contains a ‘strong’ stem that distinguishes them from the corresponding ‘weak’ genitive singular form.
First declension paradigms contain forms that (i) realize partitive singular and (ii) are composed of recurrent stems and vowels.

But neither the the vowel nor the stem realizes partitive in isolation:

- The strong stems *sukk, kukk, pukk* and *lukk* are not partitive, since they also realize the nominative.
- The theme vowels *-a, -e, -i* and *-u* are not partitive, since they are lexically specified and occur in the genitive.
- Instead, it is the combination of a strong stem and a theme vowel that realizes partitive singular in a first declension noun.

Hence the recurrent stem-vowel patterns are non-redundant.
‘Priscianic’ constructions in Latin

There are a few exceptions; but, in general, if the stem of the Past Participle is \( x \), no matter how irregular it may be, that of the Future Participle is \( x \) with \(-\text{ūr}-\) added. (Matthews 1991: 200)

<table>
<thead>
<tr>
<th>Verb</th>
<th>SUPINE</th>
<th>PAST PASS PART</th>
<th>FUT ACT PART</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMO</td>
<td>amāt-um</td>
<td>amāt-us</td>
<td>amāt+ūr-us</td>
<td>‘to love’</td>
</tr>
<tr>
<td>SECO</td>
<td>sect-um</td>
<td>sect-us</td>
<td>sect+ūr-us</td>
<td>‘to cut’</td>
</tr>
<tr>
<td>MONEO</td>
<td>monit-um</td>
<td>monit-us</td>
<td>monit+ūr-us</td>
<td>‘to advise’</td>
</tr>
</tbody>
</table>

- The supine stem recurs in each of these forms.
- But it cannot be assigned morphosyntactic features in isolation.
Inanimate ‘soft stem’ declensions in Russian (Timberlake 2004)

- Inflection class morphology is recurrent but nonredundant.

<table>
<thead>
<tr>
<th></th>
<th>1st (Masculine)</th>
<th>2nd (Feminine)</th>
<th>3rd (Feminine)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>Plu</td>
<td>SG</td>
</tr>
<tr>
<td>Nom</td>
<td>slovar’</td>
<td>slovari</td>
<td>nedelja</td>
</tr>
<tr>
<td>Acc</td>
<td>slovar’</td>
<td>slovari</td>
<td>nedelju</td>
</tr>
<tr>
<td>Gen</td>
<td>slovarja</td>
<td>slovarej</td>
<td>nedeli</td>
</tr>
<tr>
<td>Dat</td>
<td>slovarju</td>
<td>slovarjam</td>
<td>nedele</td>
</tr>
<tr>
<td>Inst</td>
<td>slovarëm</td>
<td>slovarjami</td>
<td>nedelej</td>
</tr>
<tr>
<td>Loc</td>
<td>slovare</td>
<td>slovarjax</td>
<td>nedele</td>
</tr>
</tbody>
</table>

’dictionary’  ‘week’  ‘bone’

- The different declensions are not defined by class-specific endings, but by the class-specific distribution of a set of common endings.
The interpretive role of paradigmatic context

- The Estonian and Latin and Russian patterns present different interpretive challenges for a decompositional analysis.
- They also show that the interpretation of a form may be externally motivated, determined by contrasts with other forms of the same item, rather than by internally motivated by properties of its parts.
- At the same time, the contrasts with other forms identify the class of an item and facilitate the deduction of the rest of its paradigm.
Why do inflectional patterns resist compositional analyses?
Because inflectional systems are not organized in ways that facilitate the derivation of words from inventories of minimal units:

1. It is not always possible to motivate a unique segmentation.
2. It is not always possible to assign constant properties to segments.
3. Even in cases where conditions (1) and (2) are both satisfied, it may still not be possible to define a morphological system in terms of the disassembly and reassembly of forms, since the minimal parts obtained by segmentation are tuned to contexts.
The discriminative function of sub-phonemic contrasts

- Davis et al. (2002) found suggestive differences in duration and fundamental frequency between words like *captain* and a morphologically unrelated onset word such as *cap*.
- Baayen et al. (2003) found that speakers produced Dutch nouns with a longer mean duration when they occurred as singulars than as when they occurred as the stem of the corresponding plural.
- Kemps et al. (2005) tested speakers’ sensitivity to prosodic differences, and concluded that “acoustic differences exist between uninflected and inflected forms and that listeners are sensitive to them” (Kemps et al. 2005: 441)
In summarizing experiments that tested speaker sensitivity to differences between the phonemically ‘same’ form, in isolation and as the stem for a plural form, Kemps et al. (2005: 441) conclude

The prosodic mismatch effect documented in this study has important consequences for our understanding of the morphological structure of complex words. The way words are written in languages such as Dutch and English suggests that they consist of stems and affixes that are strung together as beads on a string. Phonemic transcriptions convey the same impression. Our experiments show that this impression is wrong. **Plurals are not just singulars with an additional suffix.** The precise acoustic realization of the stem provides crucial information to the listener about the morphological context in which the stem appears.
Decompositional analysis

Problems of interpretation

The costs of decomposition

- Decompositional accounts have assumed that, where segmentation is possible, forms can be disassembled into minimal meaningful parts with no loss of information.
- Even this assumption turns out to be questionable.
- It has long been known that information about the frequency of a form cannot be recovered from the frequencies of its parts.
- However, since many models treat frequency as a matter of ‘usage’ rather than ‘structure’, this has not necessarily been regarded as a serious problem for decompositional analyses.
The costs of decomposition

- Yet frequency, among other factors, influences sub-phonemic properties, and speakers appear to be sensitive to these properties.
- The effect is that seemingly ‘recurrent’ elements are not identical. Instead, they are ‘tuned’ to the contexts in which they occur.
Uncertainty and uncertainty reduction

- A classical WP model implicitly approaches variation in terms of uncertainty and structure in terms of uncertainty reduction.
- The choice of units of analysis (words and paradigms) is guided by the goal of minimizing grammatical uncertainty, not inventory size.
Why words?

- The uncertainty that arises in associating a word form with grammatical properties is less than (or equal to, in the case of a simple word form) the sum of the uncertainty that arises in associating its component morphs with properties.

The word is a more stable and solid focus of grammatical relations than the component morpheme by itself. Put another way, grammatical statements are abstractions, but they are more profitably abstracted from words as wholes than from individual morphemes. (Robins 1959: 128).
Why paradigms?

Paradigms exhibit the most reliable patterns of interpredictability because they are defined over a closed, uniform feature space:

The most general insight is that one inflection tends to predict another ... This insight can be incorporated into any model. Traditionally, it is the basis for the method of exemplary paradigms. If the alternations were independent, these would have to be numerous ... But since they are interdependent, their number can be very small ... It is more attractive to learn paradigms as wholes than each alternation separately. (Matthews 1991: 197f)
Why information theory?

- Information theory provides frequency-weighted measures of uncertainty (surprisal/entropy) and uncertainty reduction.
- Entropy-based complexity metrics discriminate between patterns in finite systems (unlike metrics based on the Chomsky hierarchy).
- Conditional entropy (or normalized mutual information) provides a measure of remaining complexity given partial information and thus provides a realistic measure of the difficulty of using (or acquiring) a morphological system (rather than describing it ‘from scratch’).
- Entropy measures correlate closely with response latencies.
Saussure did not propose cutting words up into smaller constituents; the syntaxlike view of morphology of the Decade of the Morpheme was mainly an American notion. Saussure treated morphology in terms not of dissection but association ... His “rapports associatifs” (which later came to be called PARADIGMATIC) are just our “resonances”. They hold between a form actually spoken and an infinite number of more-or-less similar forms in the user’s internal storage; and Saussure contrasted these with the RAPPORT SYNTAGMATIQUE which hold between different forms in what is actually said ... Thus most, but not all, syntagmatic connections are between words and phrases and are therefore syntactic ... Similarly, most paradigmatic effects are resonances evoked by individual words, and are thus morphological. (Hockett 1987: 95f.)


