Morphological Theory:
Goals & Problems

Andrea D. Sims
The Ohio State University
Something ‘easy’: Affix ordering

- Chumbivilcas Quechua (Peru): Causative and desiderative affixes are ordered according to meaning (semantic scope) (Muysken 1988: 278)

a. mikhu-**naya-chi**-wa-n
   eat-**DESID-CAUS**-1.OBJ-3.SBJ
   ‘It causes me to feel like eating.’

b. mikhu-**chi-naya**-wa-n
   eat-**CAUS-DESID**-1.OBJ-3.SBJ
   ‘I feel like making someone eat.’

---

desiderative = 'want to do'
causative = 'caused version of base verb/state'
Something ‘easy’: Affix ordering

- Choguita Raramuri (Mexico, Uto-Aztecan): Mostly semantically-based ordering, but evidential marker must attach to bases that (on the surface) have final-syllable stress and are consonant final (Caballero 2010: 183-4)

a. buyá-n-čan-a
go.out.sg-DESID-EVID-PROG
‘[It sounds like [she wants to go out]]’

b. atís-ča-nare
sneeze-EVID-DESID
‘[It sounds like [they want to sneeze]]’

Evidentiality indicates the source of the action
Something ‘easy’: Affix ordering

Fula (Guinea; Niger-Congo, Atlantic): Subject and object markers order are conditioned by morphosyntactic values (Stump 2001: 151, citing Arnott 1970)

a. mball-u-(no-)*daa-mo’
help-RELPST-ACT-(PRET-)2SG.SUBJ-3SG.OBJ
‘You (sg) (had) helped him.’

b. mball-u-(no-)*moo-mi’
help-RELPST-ACT.--(PRET-)3SG.OBJ-1SG.SUBJ
‘I (had) helped him.’
Affix order summary, v. 1

- Ordering of affixes may reflect semantic structure (Chumbivilcas Quechua)...

- ... Or phonological constraints (Choguita Rarámuri)

- ... Or morphosyntactic feature structure constraints (Fula)
What do we need our theory to do?

- Identification of form units

- **Order of morphophonological (form) units**

- Constraints on operation of affixation – affixes select bases with particular properties
  - ‘Selectional restrictions’: Semantic, lexical, morphological, phonological

- Semantic relationship between base and derived form
Major questions for any theory

- What are the **primitive elements** of word structure?
- What type(s) of **operations** over those elements exist?
- What type(s) of **conditions** govern or constrain the application of those operations?
- What is the **output** of an operation?

There are fundamental disagreements in the field of morphology about how to answer these questions.
What is morphology?

- Two (simplified) alternatives for how to think about morphology:

  - Definition 1: "Morphology is the study of the systematic covariation in the form and meaning of words." (Haspelmath and Sims 2010: 2)
    - Implies a ‘Word and Paradigm’ perspective

  - Definition 2: “Morphology is the study of the combination of morphemes to yield words.” (Haspelmath and Sims 2010: 3)
    - Implies an ‘Item and Arrangement’ perspective

- These aren’t the only possibilities! Endpoints on a scale...
Item-and-Arrangement (IA)

Russian:

čitatelju  ‘reader (dat.sg.)’

→ ‘dat sg’

→ ‘person who’

→ ‘read’
Item and Arrangement framework

- **Primitive elements** = morphemes
  - Morphemes = lexical bundles of form + meaning

- **Operation type** = concatenation

- **Conditions** = mostly affix-driven selectional restrictions

- **Output** = meaning-adding (incremental)
An inflected word-form (e.g., čitatelju) is not compositional in the sense relevant to the Item-and-Arrangement approach.

Instead, inflectional values are associated with the word as a whole, not (directly) with individual subcomponents.
Word and Paradigm framework

Each word-form's inflectional values are defined by paradigmatic contrast -- the set of oppositions -- not by morphological operations that apply to build the form of the word.

<table>
<thead>
<tr>
<th></th>
<th>I čitatel’</th>
<th>‘reader’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>nom</td>
<td>čitatel’</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>čitatelja</td>
</tr>
<tr>
<td></td>
<td>gen</td>
<td>čitatelja</td>
</tr>
<tr>
<td></td>
<td>dat</td>
<td>čitatelju</td>
</tr>
<tr>
<td></td>
<td>loc</td>
<td>čitatele</td>
</tr>
<tr>
<td></td>
<td>inst</td>
<td>čitatelem</td>
</tr>
<tr>
<td>singular</td>
<td>nom</td>
<td>čitatel’</td>
</tr>
<tr>
<td></td>
<td>acc</td>
<td>čitatelej</td>
</tr>
<tr>
<td></td>
<td>gen</td>
<td>čitatelej</td>
</tr>
<tr>
<td></td>
<td>dat</td>
<td>čitateljam</td>
</tr>
<tr>
<td></td>
<td>loc</td>
<td>čitateljx</td>
</tr>
<tr>
<td></td>
<td>inst</td>
<td>čitateljami</td>
</tr>
</tbody>
</table>
Word and Paradigm framework

- **Primitive elements** = words
  - No morphemes in any classical sense

- **Operation type** = processes
  - Functions over stems that may include concatenation, but are not limited to this

- **Conditions** = affix-driven selectional restrictions, but less limited by this

- **Output** = meaning-realizing (realizational)
Returning to affix ordering...

- Is the order of morphophonological units governed fundamentally by semantic or syntactic structure, or fundamentally by independent principles of morphological organization?

- Item and Arrangement: affix order is fundamentally governed by syntax or semantics

- Word and Paradigm: affix order is fundamentally governed by independent principles of morphology
Things we will think about this month

- What are the primitive elements, operation types, and conditions on and outputs of morphological operations?

- In other words, what does a good theory (of morphology) look like? And how can we evaluate (morphological) theories?
A related but different question...
More facts about affix ordering

English

No violations of acyclicity!

Table Key
y-axis = 1st affix
x-axis = 2nd affix
# = word types attested with affix combo

Plag and Baayen (2009:123)
More facts about affix ordering

- The set of English derivational suffix combinations is statistically acyclic.
- In other words, there is some constraint on affix combinability that is not simply a function of meaning, or the properties of individual affixes.

<table>
<thead>
<tr>
<th>Affixes</th>
<th>Attested combos</th>
<th>Violations of acyclicity</th>
<th>Likelihood due to chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>English suffixes (Plag &amp; Baayen 2009)</td>
<td>31</td>
<td>161</td>
<td>10*</td>
</tr>
</tbody>
</table>

* Questionably valid words; ultimately rejected
One hypothesis

- Complexity-based Ordering (CBO) Hypothesis (Hay 2003): More parsable affixes occur external to less parsable affixes
  - The order of affixes on the previous screen reflects their parsability

- In other words, Hay posits a cognitively-rooted explanation for why there are no affix combinations ‘below the diagonal’: those combinations are difficult to process
More facts about affix ordering

Russian

Violations of acyclicity!

Table Key
y-axis = 1\textsuperscript{st} affix
x-axis = 2\textsuperscript{nd} affix
1 = word with combo is attested (not a type frequency count)

<table>
<thead>
<tr>
<th></th>
<th>ar'</th>
<th>tel'</th>
<th>anin</th>
<th>ak</th>
<th>ist</th>
<th>ant</th>
<th>ov</th>
<th>im</th>
<th>sk</th>
<th>ičn</th>
<th>liv</th>
<th>čik</th>
<th>stv</th>
<th>ik</th>
<th>izm</th>
<th>ničaj</th>
<th>išč</th>
<th>acij</th>
<th>ost'</th>
</tr>
</thead>
<tbody>
<tr>
<td>ar'</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>tel'</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>anin</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ak</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ist</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ant</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ov</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>im</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>sk</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ičn</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>liv</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>čik</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>stv</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ik</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>izm</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ničaj</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>išč</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>acij</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ost'</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Sims and Parker (2015)
Affix order summary, updated

- Ordering of affixes may reflect semantic structure (Chumbivilcas Quechua)...
- Or phonological constraints (Choguita Rarámuri...)
- Or morphosyntactic feature structure constraints (Fula)...
- But affix combinations that obey selectional restrictions of individual affixes may nonetheless fail to exist (English)
  - Suggests a more general constraint – maybe rooted in lexical processing?
  - But not a universal one (Russian)
This raises the question...

- Should our theory be required to account for all (un)attested combinations of morphological units, or does explanation for some facts fall outside of the grammar?

- A more general version of this question (important for any theory):

  What are we trying to describe (and explain)?
  What is the scope of the data under investigation?
Basic concepts:
Selectional restrictions
Semantic restrictions

- German -in does not like to create female nouns for things low on the animacy scale

  König-in  ‘queen’
  Löw-in  ‘lioness’
  ??Käfer-in  ‘female beetle’
  ??Würm-in  ‘female worm’

- English de-

  ??deassassinate
  ??deincinerate
Semantic restrictions

- **English un-**
  
  unhappy vs. ?unsad (but cf. not unsad)
  unwell vs. ??unill
  uncheerful vs. ??unsorrowful
  unoptimistic vs. ??unpessimistic

- **English -ed for inalienable compounding**
  
  blue-eyed
  three-legged
  red-roofed
  *two-carred man
  ??black-shoed lady
Lexical restrictions

- *-ous + -ity* (*-ous* sometimes disappears – lexically determined)

curious $\rightarrow$ curiosity
voracious $\rightarrow$ voracity
Morphological restrictions

- English -ment + -al

<table>
<thead>
<tr>
<th>ornament</th>
<th>ornamental</th>
</tr>
</thead>
<tbody>
<tr>
<td>employment</td>
<td>*employmental</td>
</tr>
</tbody>
</table>

- Example of more general constraints on combos of ‘Germanic’ and ‘Latinate’ suffixes in English
Morphological restrictions

- Latinate suffixes (e.g. -ify, -ate, -ity) – tend to be vowel-initial and secondarily stressed
  - Prefer Latinate bases
  - Often have bound roots as bases
  - Roots are mostly polysyllabic (if not bound)
  - Not likely to attach to a base with a Germanic suffix (e.g. *-less-ity)
Morphological restrictions

- Germanic suffixes (e.g. -ship, -ful, -ness, -hood, -en) – tend to be consonant-initial and unstressed
  - Tolerate both kinds of bases
  - Usually the base is also itself viable as a free word
  - Roots are mostly monosyllabic or disyllabic with an unstressed second syllable
  - Can be attached to a base with a Latinate suffix (e.g. -ive-ness).
Phonological restrictions

- English expletive infixation -- comes immediately before stressed syllable; prefers 3-syllable words with non-initial stress

  lickety-fucking-split (*lick-fucking-etysplit)
  im-fucking-possible
  kanga-bloody-roo
  propa-fucking-ganda
  BUT atmo-fucking-sphere
Phonological restrictions

- **German diminutive affixes -chen and -lein**
  - After ...l(e), only -chen (Bällchen ‘little ball’; Schälchen ‘little scarf’)
  - After /x, ñ, g/, only -lein (Bächlein ‘little stream’, Ringlein ‘little ring’)
  - After bases ending in other consonants, free variation (Briefchen/Brieflein ‘little letter’; Häuschen/ Häuslein ‘little house’)

The fact about science is that everyone who has made a serious contribution to it is aware, or very strongly suspects, that the world is not only queerer than anyone has imagined, but queerer than anyone can imagine. This is a most disturbing thought, and one flees from it by stating the exact opposite.

J.S. Haldane, as cited by R. G. Reid (2007: 4311)