Ling 580e: Computational Morphology
September 30, 2004
Introduction, Montage, Slave morphology
Overview

- Introduction: course goals, assignments
- Who’s here?: discussion assignments
- Montage
- Slave morphology
Introduction

- Course goals:
  - Learn about finite-state morphology
  - Design an interface between LKB and \texttt{xfst}
  - Learn about (and reflect on) the research process

- Course requirements:
  - Participation (15/30%)
  - Leading a discussion (15/30%)
  - KWLH paper (20/40%)
  - Term paper or project (50/0%)

- Who’s here?
KWLH paper

- What you already know (~1 page)
- What you want to learn (~1 page)
- What you learned (~3 pages)
- How you’ll apply it in your research/studies (~2 pages)
- Write the “K” and “W” parts by next week.
- Keep notes along the way for L and H.
- Whole paper (7 pages) due 12/9.
**Term paper/project**

- Anything related to computational morphology
- (Doesn’t have to be Montage-related)
- Papers should be ~15-20 pages, the kernel of something that could be extended to a conference paper.
- Projects should be accompanied by a 5 page description.
- Choice of final type and topic due 10/28
- Term paper outlines, term project specs due 11/24 (Wed. before Thanksgiving)
Montage (1/2)

- Markup for ONTological Annotation and Grammar Engineering
- Software to support language documentation
- Fits in with existing systems:
  - DoBeS’s Elan for transcription
  - E-MELD’s FIELD for lexica
  - E-MELD’s GOLD (general ontology for linguistic description)
Montage (2/2)

- Leverage advances in computational linguistics to benefit descriptive/documentary linguistics
- Exploit synergy between descriptive and formal grammars
- Create discoverable/accessible resources almost as a side-effect
Figures

- How Montage fits in
- Workflow with Montage
Ramping up to implemented formal grammars

- Electronic descriptive grammars, linked to GOLD-annotated texts
- Underspecified formal grammars, Matrix-based “wizards”, ??
- Implementations of linguistic hypotheses, testable against corpus data (Matrix-based)
Morphology in Montage

- Independent morphophonological analysis
- Morphophonological analysis attached to LKB morphosyntactic (“lexical”) rules
Two approaches to morphology

- Item-and-Arrangement: each morpheme has an underlying form, put together in a string or a tree.
- Item-and-Process: roots have underlying forms, all other morphemes correspond to processes which affect the phonological form of the stem (or not).
- Sub-word trees v. spindly chains of lexical rules.
- Hybrids? Other possibilities?
Montage: Summary

- Need a morphological component that is up to the task
- Reuse of morphophonological rules in different stages of analysis
- This quarter: Design the interface between LKB and xfst.
- Do so with four ornery cases in mind.
Case 1: Slave (Athabaskan)

- Up to sixteen prefixes on a verb
- Incorporation of two different kinds of open-class stems in verbs (adverbs and nouns)
- Lexical entries (‘verb themes’) consisting of discontinuous strings of stem+prefix
- Elaborate phonological rules
Verb prefixes

obj= pp# adv#dist#cust#stem#no.+DO+deic+theme+asp+conj+mode+subj=cl
se- e ná- yá- na- kwi ĺe- se- ts’e ne- de- φ φ h- φ
ne- -ch’a ní- gó- go- ne- ke- de- ne- n- ñ- ne- h-
be- -tá ni- tsih- be- ye- í- w- ghu- íd- d-

* * open-class slot

stem: optional possessive pronoun + stem
cl: ‘classifiers’, order: d- h- thematic
Example verb

-ná- ?e- ne- l- nih ‘wrap oneself’

PP DO asp cl stem

?edená?enehndih ‘I wrap myself’
?edená?enendih ‘s/he wraps him/herself’
?edená?enįndih ‘you sg wraps yourself’
Incorporation

- There are two positions for open-class incorporation:
  - One takes only adverbs (and aspect markers)
  - One takes nouns, ‘action stems’, and adverbs. These can be subjects, objects (direct or oblique), and adverbs.
- Some stems have a special form when incorporated, others don’t change, still others only appear as incorporated stems.
- Incorporated body parts must be possessed by subject.
Adverb-only position: examples

?a # go % ‘go nonstop, continue’ (go is DO)
?a odé?e ‘s/he went nonstop by boat’
?a odedéhdhe ‘s/he flew nonstop’
di # de+ ‘into fire’ (de is aspect)
didedadhé ‘s/he flew into fire’
obj didiŋla ‘s/he put obj in fire’
**Anything goes position: examples**

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>sa-</td>
<td>‘sun’</td>
</tr>
<tr>
<td>rásayiʔo</td>
<td>‘the sun set’</td>
</tr>
<tr>
<td></td>
<td>(s/he placed 3D obj down)</td>
</tr>
<tr>
<td>keeshį</td>
<td>‘shoelace’</td>
</tr>
<tr>
<td>rakeeshįdéhyá</td>
<td>‘they (shoes) are tied’</td>
</tr>
<tr>
<td>shį</td>
<td>‘song’</td>
</tr>
<tr>
<td>k’í shinededa</td>
<td>‘s/he walks around singing’</td>
</tr>
</tbody>
</table>
Verb themes

- A verb theme is the basic lexical entry for a verb.
- It consists of the stem, the classifier, and any prefixes that must occur with the verb.
- Thematic prefixes can be incorporated postpositions, adverbials, incorporated stems, number prefixes, direct object markers, and ‘themes’.
**Verb theme examples**

<table>
<thead>
<tr>
<th>theme</th>
<th>go-ϕ-deeh</th>
<th>‘talk’ (go is DO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>godee</td>
<td>‘s/he talks’</td>
<td></td>
</tr>
<tr>
<td>base</td>
<td>da-go-ϕ-dee</td>
<td>‘stutter’ (da is inc. stem)</td>
</tr>
<tr>
<td>dagodee</td>
<td>‘s/he stutters’</td>
<td></td>
</tr>
<tr>
<td>base</td>
<td>-éh-go-d-ϕ-dee</td>
<td>‘tell a story’ (-éh is pp, d is cl)</td>
</tr>
<tr>
<td>ségadee</td>
<td>‘s/he told me a story’</td>
<td></td>
</tr>
</tbody>
</table>
**Elaborate phonological rules**

- **The D-Effect Rule**

  \[
  \begin{align*}
  d + ? & \rightarrow t' \quad d + z \quad \rightarrow \quad dz \\
  d + zh & \rightarrow j \quad d + gh \quad \rightarrow \quad g \\
  d + l & \rightarrow dl \quad d + w \quad \rightarrow \quad gw/b \\
  d + n & \rightarrow d \quad d + m \quad \rightarrow \quad b
  \end{align*}
  \]

- Applies when classifier *d-* or 1pl *íd-* precede a stem-initial consonant.
Slave – Summary

- The morphosyntactically ‘easy’ stuff is buried inside the hard stuff.
- Underlying forms are remote from surface forms.
- Incorporation suggests an item-and-arrangement approach.
- D-Effect and other rules lend themselves to item-and-process.
- Rice’s approach is entirely item-and-arrangement.
Overall Summary

- Introduction: course goals, assignments
- Who’s here?: discussion assignments
- Montage
- Slave morphology