Implementing a Syntax-Morphology Interface for Athabaskan

Emily M. Bender and Jeff Good
ebender@u.washington.edu and good@eva.mpg.de
University of Washington and MPI Leipzig
LSA, Oakland CA, 1/7/05

Introduction: Montage

- Suite of tools to assist in the documentation of underdescribed languages (Bender et al 2004)
- Focus on grammar (especially morphology and morphosyntax)
- Integrate with other initiatives building tools for transcribed texts and lexicons (e.g., ELAN, FIELD, AGTK)

Overview

- Introduction/goals
- Terminological distinctions
- Reject two possible interfaces
- Proposed interface design: run-time and development
- Conclusions
  ... Illustrated with examples from Slave (Rice 1989)

Some terminology

- Morphophonology:
  - Morphotactics (e.g., position classes)
  - Morph.-conditioned phonological rules
  - General phonological rules
  - Mapping to abstract morphemes
- Morphosyntax:
  - Syntactic-semantic representations built from analysis of strings of abstract morphemes

Possible interfaces

- Morphophonology in morphosyntax
- Morphosyntax in morphophonology
- Independent morphophonology and morphosyntax

Introduction: Montage

- Overarching goal: Allow the “ordinary working linguist” to make use of sophisticated grammar engineering tools without being grammar engineers themselves
- This talk: the Montage model for morphological analysis, and the morphology-syntax interface
Morphophonology in Morphosyntax

- Morphosyntactic rules associated with morphophonological effects
- Standard in HPSG, perhaps most thoroughly worked out in Orgun 1996
- Assumed in current version of the LKB (Copestake 2002)

Morphosyntax in Morphophonology

- Interpret abstract morphemes as actual feature bundles
- Output of morphophonology is a lexical edge which can be used directly by the morphosyntactic parser
- Doesn’t generalize to morphosyntactically complicated cases

Morphophonology in Morphosyntax

- Hard to reuse morphophonological work in morphosyntax
- Hard to push all morphophonology into one efficient machine
- Particularly awkward for strictly phonological effects

Slave Morphological Causatives

<table>
<thead>
<tr>
<th>hedenět̪i</th>
<th>hdenět̪̣t̪i</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘s/he fell asleep’</td>
<td>‘s/he put him/her to sleep’</td>
</tr>
</tbody>
</table>

(Rice 1989:454)

- Syntactically and semantically, the causative form cannot be produced merely by adding features to the intransitive form.

Epenthesis in Slave

- An epenthetic “peg element” is inserted before verb stems if they would not otherwise be preceded by some syllable in their word (Rice 1989:133)
- hehji ‘I sing’ vs. nejji ‘you sing’

Theoretical conclusion

- A computational system should allow morphophonology and morphosyntax to be modeled as independent, articulated systems
- The point of interface is the abstract morpheme
Independent
Morphophonology and Morphosyntax

- Morphophonology: maps surface forms to strings of abstract morphemes
- Morphosyntax: maps strings of abstract morphemes to syntactic/semantic information (feature structures)

cf. Kaplan et al 2004, Siegel and Bender 2002

Runtime Interface

Development interface

- Spell each underlying stem only once
- Define default morphotactic/syntactic pairings
- Allow multiple continuation classes for the same word sense and vice versa

Slave Verb Classifiers

- Verb classes based on the “classifiers” a verb theme contains (Rice 1989:439–470)
  - Ø-classifier, Ø-ʔah ‘eat, chew’
  - h-classifier, h-ʔô ‘suck’
  - d-classifier, d-shin ‘sing’
  - l-classifier, ná-l-séh ‘hunt’

Bipartite lexical database design

One morphosyntactic entry::many morphophonological entries

- “Each classifier has a basic function, although they must be considered as part of the verb theme since this semantic content is not always clear.” (Rice 1989:453)
- In some cases, verbs can alternate in their choice of classifier, with no non-morphophonological consequences (Rice 1989:449–50)
One morphophonological entry::many morphosyntactic entries

- Homophony within the same morphological class
- Multiple valence patterns, not predicted by a productive valence alternation

Correlated morphophonological and morphosyntactic choices

- ṭe ‘ice’ (as a noun), -ṭe ‘freeze’ (as a verb) (Rice 1989:161)
- kāṭedijste ‘s/he broke through the ice’ (incorporated noun) (Rice 1989:653)
- These stems will be associated with multiple morphophonological and morphosyntactic classes
- Handle correlation explicitly or implicitly

Conclusions

- Morphology and syntax are best treated as independent of one another
- Point of interface is abstract morphemes
- Two interfaces are required: run time and development
- Morphologically exuberant languages like Athabaskan are informative

Acknowledgments

Thanks for helpful discussion to:
Duane Blanchard, Anya Dormer, Scott Drellishak, Ann Gaponoff, David Goss-Grubbs, Jeremy Kahn, Bill McNeill, Matty Noble, Laurie Poulson

References

AGTK: Annotation Graph Toolkit. http://www.ldc.upenn.edu/Projects/AG/