

Morphology (2)

Definitions

- Morphology: study of word formation
 - Morph: shape → "shape" of individual words in a language
- Morpheme
 - Smallest meaning-bearing unit of language
 - Pieces that make up words
- Processes: methods used to form and combine words in language(s)

Classifying words

- Part-of-speech category
 - Noun, verb, adjective, adverb, etc.
- Simple word vs. complex word
 - One morpheme vs. more morphemes
- Word classes
 - Open class vs. closed class
(also called lexical vs. functional)

Classifying morphemes

- Root: most important morpheme of the word
- Affix:
 - Prefix: precedes the root
 - Suffix: follows the root
 - Infix: splits apart the root
 - Circumfix: surrounds the root
- Base: word that undergoes affixation
- Free/bound morphemes
 - Whether or not can stand alone

English word structure

- Pluralization of nouns
 - dog+s, bat+s, walk+s
- Conjugation of verbs
 - walk+∅, walk+s, walk+ed, walk+ing
- Adverbialization of adjectives
 - careful+ly, reckless+ly
- Other possibilities
 - out+swim/eat/run, re+do/think/release, non-negotiable/returnable
 - big+er, big+est

English word structure (2)

- English is not complicated, yet nontrivial:
 - skies ← sky+s, flies ← fly+s, keys ← key+s
 - forgetting ← forget+ing, targeting ← target+ing
 - baking ← bake+ing, tragically ← tragic+ly
- Lots of exceptions: went ← go+ed, automata ← automaton+s, child(ren), corpora ← corpus+s
- Ambiguity
 - axes ← axe+s OR axis+s
 - runs ← verb OR noun
 - leaves ← leaf+s OR leave+s

Portuguese word structure

- Verb conjugation
 - 63 possible forms
 - 3 major conjugation classes, many sub-classes
 - Over 1000 verb ending types
- Noun pluralization
 - Almost as simple as English
- Adjective inflection
 - Number
 - Gender

Portuguese verbs (falar)

falando

falado

falar falares falar falarmos falardes falarem

falo falas fala falamos falais falam

falava falavas falava falávamos faláveis falavam

falei falaste falou falamos falastes falaram

falara falaras falara faláramos faláreis falaram

falarei falarás falará falaremos falareis falarão

falaria falarias falaria falaríamos falaríeis falariam

fala falai

fale fale falemos faleis falem

falasse falasses falasse falássemos falásseis falassem

falar falares falar falarmos falardes falarem

Finnish complexity

- Nouns
 - Cases, number, possessive affixes
 - Potentially 840 forms for each noun
- Adjectives
 - As for nouns, but also comparative, superlative
 - Potentially 2,520 forms for each
- Verbs
 - Potentially over 10,000 forms for each

Archi verb complexity

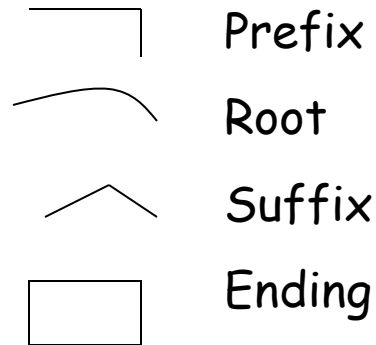
- Mood: 223 forms
- Negation: 48 forms
- Tense/aspect: 34 forms
- Participles: 470 forms
- Gerunds: 929 forms
- ...
- Grand total: 1,502,839 forms/verb (!)

Armenian word forms

հայր
հօր
հօրմով
հայրական
հայրենակից
հայրենասեր
հայրենասիրական
հայրենասիրութիւն
հայրենիք
հայրութիւն
նախահայր

Armenian traditional analysis

դժբախտութիւններով



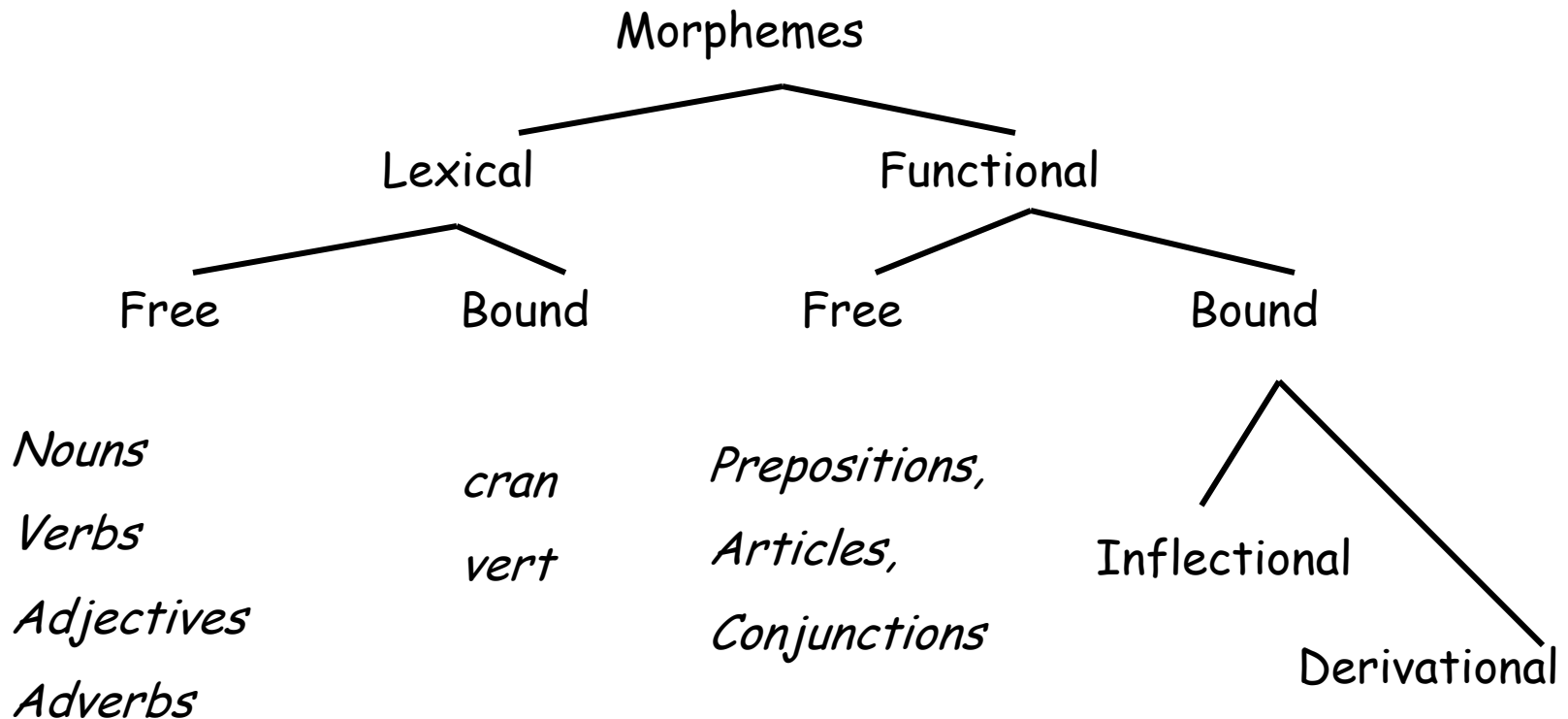
Linguistic complexity

- Varying degrees of word complexity across languages
 - qasuiirsarvingssarsingitluinarnarpuq
 - "someone did not find a completely suitable resting place"
 - Dampfschiffahrtsgesellschaftsdirektorsstellvertretergemahlin

English complexity

superconductivity's	disproportionately	overspecialization
telecommunications	constitutionality	counterproductive
misrepresentations	superconductivity	administration's
biotechnological	deoxyribonucleic	enthusiastically
immunodeficiency	mischaracterizes	nonmanufacturing
nonparticipation	pharmaceuticals'	recapitalization
responsibilities	superspecialized	unapologetically
unconstitutional	administrations	anthropological
capitalizations	cerebrovascular	competitiveness
computerization	confidentiality	confrontational
congressionally	criminalization	discombobulated
discontinuation	dispassionately	dissatisfaction
diversification	entrepreneurial	experimentation
extraordinarily	inconsistencies	instrumentation
internationally	liberalizations	micromanagement
microprocessors	notwithstanding	pharmaceuticals
philosophically	professionalism	proportionately

Morpheme classes



Base (citation) form

- Dictionaries typically don't contain all morphological variants of a word
- Citation form: base form, lemma
- Languages differ on citation form
 - Armenian: verbs listed with first person sg.
 - Semitic languages: triliteral roots
- Dictionaries differ on citation form

Some properties (morphemes)

- Ambiguity (some morphemes)
 - -er: agentive suffix (e.g. singer, kicker, ...)
 - -er: comparative suffix (e.g. bigger, hotter)
 - -s: present tense conjugation (e.g. sings)
 - -s: plural (e.g. cats)
- Productivity
 - How widely can it be used?
- Allomorphy
 - Different forms of the same morpheme occur in various situations

Allomorphy

- Allomorphs: variant forms for related morphemes
 - a/an
 - -ed, -t
 - il-, in-, ir-, im-
- Selection of appropriate form is often based on surrounding sounds .

Some properties (morphemes)

- Ambiguity (some morphemes)
 - -er: agentive suffix (e.g. singer, kicker, ...)
 - -er: comparative suffix (e.g. bigger, hotter)
 - -s: present tense conjugation (e.g. sings)
 - -s: plural (e.g. cats)
- Productivity
 - How widely can it be used?
- Allomorphy
 - Different forms of the same morpheme occur in various situations

Analyzing morphological data

- Find regularities
- Identify morphemes and their types
- Group possible allomorphs together
- Keep processes, classes, constraints in mind
- Glosses can be both helpful and misleading!
- Practice, practice, practice!

Word formation processes

- Affixation
- Internal change
- Suppletion
- Stress/tone placement
- Reduplication (full, partial)
- Interleaving
- Conversion
- Clipping
- Blends
- Acronyms (e.g. NATO)
- Initialisms (e.g. FBI)
- Coinage
- Backformation