BBN–FLN11–001/eng
FOUNDATIONS OF LINGUISTICS
5. MORPHOLOGY

Péter Szigetvári, szigetvari@elte.hu, Department of English Linguistics, Eötvös Loránd University

http://seas.elte.hu/w/!fol

morphology and morphs

- morphology is about (word) forms (Greek *morphē* = Latin *forma*; apparent metathesis!)
- a morph is a smallest meaningful part of language
- words are all meaningful:
  cut, dish, red; morphologically simple(x) words contain a single morph
- many words contain smaller smaller meaningful parts, they are morphologically complex:
  uncut, dishful, redness
- but in other words the same string may not be meaningful:
  cutlass, radish, credit; these words are morphologically simplex

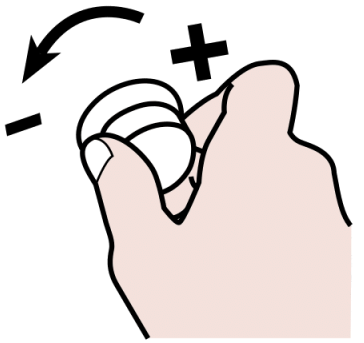
words

- orthographic word: a string separated by spaces in writing; linguistically irrelevant
- phonological word: a string with at least one stressed vowel, which can be used on its own as an utterance
- morphological word: ?
- syntactic word: a string with some degree of independence in the sentence

syntactic word

- my sister left – my little sister left
- John has left – has John left?
- letekeri a hangot ‘down-turn the volume’ – le is tekeri a hangot
- letekeri a hangot – tekeri le a hangot

⇒ *my, has,* and *le* are syntactic words



phonological word

– is it your or my sister that you're meeting?
– *my/mine

– she hasn't left
– *has/she has

– letekeri a hangot?
– le

⇒ *my* and *has* are not phonological words, but *le* is

morphological word?

– what did she buy?
– dogfood

- *food dog, *dog X food

⇒ *dogfood* is one phonological word and one syntactic word

- but clearly *dog* is a word and *food* is another word

morphological complexity: roots and affixes

- morphologically simplex: *tie* is one word and one morph
- morphologically complex: one word but more than one morph
  - *tiebreak*: both parts are words, two roots
  - *untie*: prefix + root
  - *ties*: root + suffix
- rarer types of affixes
  - circumfix: Hu *menő* 'cool' ~ legmenőbb 'cool-SUPERLATIVE'
    (note that this is better analysed as a comparative suffix, *-bb,* and a prefix, *leg-*)
  - infix: La *vīc-it* 'win-PS.3SG' ~ *vinc-it* 'win.IMPF-PS.3SG'
  - simultaneous affix: common in sign languages; En *tórment* noun, *tormént* verb(?)

roots and stems

- a root is one morph (all affixes removed): *cat*-*s*, *un*-*happy*, *weak*-*ness*-*es*, *un*-*law*-*ful*
- a stem *may* consist of several morphs (lastly added affix removed): *weak*-*ness*-*es*, *un*-*law*-*ful*
- a root is always a stem
- a stem may or may not be a root

free and bound morphs

- a free morph can occur on its own: it is a word (in the syntactic sense)
- affixes are bound morphs by definition
- stems can be
  - free: *dogs, whiteness, antimissile, replay*
  - bound: *wives, dogmatic;* Hu *lova* ‘horse-POSS.3SG’, *álmodik* ‘dream-PRES.NDF.3SG’

note, we consider the pronunciation, not the spelling

- *houses* /**haw**z**ɪ**z/ has a bound stem (cf *house* /haws/)
- *hurried* /**hə**ri**d**/ has a free stem (cf *hurry* /həri/)

morphology without affixes

internal change

- **umlaut:** *mouse ~ mice, goose ~ geese, fall ~ fell, full ~ fill*
- **ablaut:** *sing ~ sang ~ sung, write ~ wrote*
- **other:** *hide ~ hid, bite ~ bit*

no change

hit PRES ~ *hit* PAST, *sheep* SG ~ *sheep* PL, *dream* noun ~ *dream* verb

suppletion

- *go* ~ *went* (*went* used to be the past form of †*wend*)
- Fr *vais* ‘go-PRES.1SG’ ~ *allons* ‘go-PRES.1PL’ ~ *irai* ‘go-FUT.1SG’ (cf Latin *vadere, ambulare, ire*)
- Hu *jön* ‘come’ ~ †*jöjj, gyere* ‘come-IMP.2SG’
- *be ~ am ~ is ~ was; good ~ better; bad ~ worse; she ~ her*
- Hu *van* ‘be’ & *lesz* ‘become’

PRESENT	PAST	CONDITIONAL	INFINITIVE	SUBJUNCTIVE
van	volt	volna	★	★
lesz	lett	lenne	lenni ↑	legyen ↑

reduplication

- Maori *wera* ‘hot’ ~ *werawera* ‘warm’
- Swahili *piga* ‘strike’ ~ *pigapiga* ‘strike repeatedly’
- Hu *lassan* ‘slowly’ ~ *lassan-lassan* ‘very slowly’, It *piano* ‘softly’ ~ *piano piano* ‘very softly’
- En *table schmable*; Hu *csiga* ‘snail’ ~ *csigabiga*, *icipici* ‘very small’ (cf *pici* ‘small’ and En *itsy-bitsy*)
- La *mord-et* ‘bite-3SG’ ~ *mo-mord-it* ‘PERF-bite-3SG’,
  curr-it ‘run-3SG’ ~ *cu-curr-it* ‘PERF-run-3SG’
- AGk *lu-ē* ‘loosen-3SG’ ~ *le-lu-k-e* ‘PERF-loosen-PERF-3SG’,
  graph-ē ‘write-3SG’ ~ *ge-graph-e* ‘PERF-write-3SG’

what looks like a morph becomes a morph

- *hamburger* etymologically is *Hamburg+er*, but in it *ham* looks like a morph, so *-burger* becomes a suffix: *cheeseburger, chickenburger, fishburger,* etc
- Fr *cherise* > En *cherries,* looked like a plural form, so the singular must be *cherry*
- Hu *vakondok* 'mole' looks like a plural form, so the singular must be *vakond*
- Hu *szolidaritás* 'solidarity' (from La *solidaritas*) looks like *szolidarít+ás*, so the verb *szolidarít* emerges

⇒ this process (using a morph from the reanalysis of a word) is called backformation

morphological language types

- isolating or analytic: morph = word (no/little morphology): eg Mandarin
- synthetic: most words are morphologically complex
  - agglutinative: clearly separable morphs, eg Hu *bátyá-i-m-ék-é-i-n* ‘older_brother-PL.POSS-1SG-FAMPL-POSR-PL-SUP’ = ‘on those of my older brothers and their families’
  - fusional: meaning merged in morphs, eg Latin

some forms of the word for ‘oak’ in Hungarian and Latin

	SG NOM	SG ACC	SG DAT	PL NOM	PL ACC	PL DAT
Hu	tölgy	tölgy-et	tölgy-nek	tölgy-ek	tölgy-ek-et	tölgy-ek-nek
La	quercus	quercum	quercuī	quercūs	quercūs	quercubus

allomorphs

- the free morph *wife* and the bound morph *wive-* are in complementary distribution: *wife* occurs in singular forms (eg *wife's*), *wive-* in plural forms (eg *wives, wives'*)
- they are allomorphs of the same morpheme (recall allophones and phonemes)
- plural allomorphs in English: `s`, `z`, `ɪz`, `ən`, ∅, etc
  `kæts`, `dɔgz`, `fɔksɪz`, `ɔksən`, `ʃip` (~ `ʃip`), `gis` (~ `gus`)
- phonologically conditioned allomorph selection governed by the sound shape of the stem: after voiceless C: `-s`, after sibilant: `-ɪz`, elsewhere: `-z`
- lexically conditioned allomorph selection: plural is unique to a given stem, independent of its sound shape: `ɔks` is like `fɔks`, yet the plural is `ɔksən`, not *`ɔksɪz`

lexeme

- are *like, likes, liked, liking* the same word or different words?
- these are four word forms but one lexeme
- all the word forms that belong to a lexeme are called its (inflectional) paradigm
- the paradigm of LIKE contains *like, likes, liked, liking*
- the paradigm of MAN contains *man, man's, men, men's*
- the largest paradigm in modern English is that of BE: *be, am, is, are, was, were, been, being*
- the word form used to refer to a lexeme is its lemma or citation form
- misleadingly the word *word* is used for both 'lexeme' and any of the word forms in its paradigm

types of lexemes

- content (or lexical) categories/words
  - nouns (N): *dog, happiness, unlawfulness*
  - verbs (V): *like, run, satisfy, compensate*
  - adjectives (A): *good, fast, green, catlike*
  - adverbs (Adv): *well, fast, quickly, satisfactorily*
- function (or grammatical) categories/words
  - determiners: *a/an, the, this, those*
  - pronouns: *you, her, mine, what, who, nothing*
  - prepositions: *to, in, from, between, with*
  - auxiliaries: *can, must, will, have*
  - conjunctions: *and, or, but, although*

how to identify lexical categories (aka “parts of speech”)?

- by their meaning: “a noun is a person, place, or thing”, “a verb means an action”, etc
  but the word *action* means an action, yet it is not a verb, this is not a good way
- by their distribution
  - a noun, adjective, or adverb can follow a determier, a verb cannot
  - a noun can be modified by an adjective
  - a verb or an adjective can be modified by an adverb, etc, etc
- by their paradigms
  - (many) nouns have a plural form, adjectives do not
  - (many) adjectives have a comparative and a superlative form
  - verbs have a past tense form, etc, etc

derivation creates a new lexeme

- as a result potentially changes category
  - *cheese* noun → *cheesy* adjective
  - *pure* adjective → *purify* verb
  - *run* verb → *runner* noun
  - *govern* verb → *government* noun
- but not necessarily
  - *mother* noun → *motherhood* noun
  - *green* adjective → *greenish* adjective
  - *tie* verb → *untie* verb

inflection creates other word forms within the paradigm

- nouns (also often adjectives)
  - number: *dog* ~ *dogs, child* ~ *children*
  - case: Hu *kapu* ‘gate’ ~ *kaput* ‘gate-ACC’, La *arbor* ‘tree’ ~ *arboris* ‘tree-GEN’
  - possessives: Hu *kapu* ~ *kapum* ‘gate-POSS.1SG, ie my gate’
- verbs
  - tense (present, past): *learn* ~ *learned, hide* ~ *hid*
  - aspect (imperfective, perfective): La *scribit* ‘s/he writes’ ~ *scripsit* ‘s/he has written’
  - mood (indicative, conditional, subjunctive): Hu *rak* ‘put’ ~ *rakna* ‘-CND’ ~ *rakjon* ‘-SBJV’
  - voice (active, passive): La *scribit* ~ *scribitur* ‘-PASS, ie is written’
  - person/number (of subject, object): *put* ~ *puts,* Hu *lát* ‘see.NDF.3SG’ ~ *látlak* ‘-1SG»2’

compounds

- NN: doghouse, website, basketball, sunflower, moonlight, beekeper, heartburn, spaceship
- AN: greenhouse, bluebird, goodwill, lazybones
- VN: breakwater, cheatsheet
- NA: trustworthy, watertight
- AA: purebred, kind-hearted, blue-green
- NV: browbeat, manhandle, sidestep, babysit
- AV: blacklist

compound vs phrase

- spelling is inconsistent
- stress is usually telling:
  - *bláckboard* vs *black bóard*
  - *gréenhouse* vs *green hóuse*
  - *blúebird* vs *blue bírd*
- meaning may be a clue: a *black bóard* is black, a *bláckboard* may not be
- or syntax: *car factory* cannot be a phrase, because two nouns cannot be adjacent in a phrase

endocentric (headed) and exocentric (unheaded) compounds

- compounds are typically right headed: meaning and category are decided by the right-hand side
  - *darkblue* is a kind of blue and an adjective
  - *pitchdark* is a kind of dark and an adjective
  - *blackboard* is a kind of board and a noun
  - *babysit* is a kind of “sitting” and a verb
- some compounds do not have a head (= are exocentric)
  - *sabretooth* (NN) is not a kind of tooth (but a kind of animal) but it is a N
  - *redhead* (AN) is not a kind of head (but a kind of person) but it is a N
  - *spoilsport* (VN) is not a kind of sport (but a kind of person) but it is a N
  - *outcome* (PV) is not a kind of coming and is not even a V (but a N)
  - *cover-up* (VP) is not a kind of up (whatever that would be) and is not even a P (but a N)

dvandva compounds

two heads

- *zigzag, flip flop, riff raff, hocus pocus*
- *bittersweet (symphony), secretary-treasurer, parent-teacher (meeting), blue-green*
- Hu *jön-megy, lót-fut, izegmozog*

blocking

- the inflected plural form **childs* is blocked by the existence of the plural form *children*
- the comparative form **gooder* is blocked by the existence of the comparative form *better*
- the derived agentive form ?*stealer* is blocked by the existence of *thief*

paradigm gap

- *used to* has no present tense forms
- Hu *sínyl-* ‘suffer’ lacks some forms (where the suffix does not begin with a V: **sínyl*, **sínyljük*)

PRES	NONDEFINITE	DEFINITE
1SG	sínylek	sínylem
2SG	sínylesz	sínyled
3SG	★	sínyli
1PL	sínylünk	★
2PL	sínyletek	sínylitek
3PL	sínylenek	sínylik

ambiguity

- *untieable* can mean
  - ‘can be untied’, [[untie] able]
  - ‘cannot be tied’, [un [tieable]]
- *river boat race* can mean
  - ‘a race of river boats’, [[river boat] race]
  - ‘a boat race on the river’, [river [boat race]]
- *toy car factory* can mean
  - ‘a factory for toy cars‘, [[toy car] factory]
  - ‘a car factory that is a toy’, [toy [car factory]]