Carnegie Mellon University
Language
Technologies
Institute

11-324/11-624/11-724 Human Language for AI

Construction Morphology

David R. Mortensen

October 12, 2022

Language Technologies Institute
Carnegie Mellon University

Introduction

After this lecture, students will have a clear understanding of the following concepts:

- Lexicalization
- Grammaticalization
- Morphologization

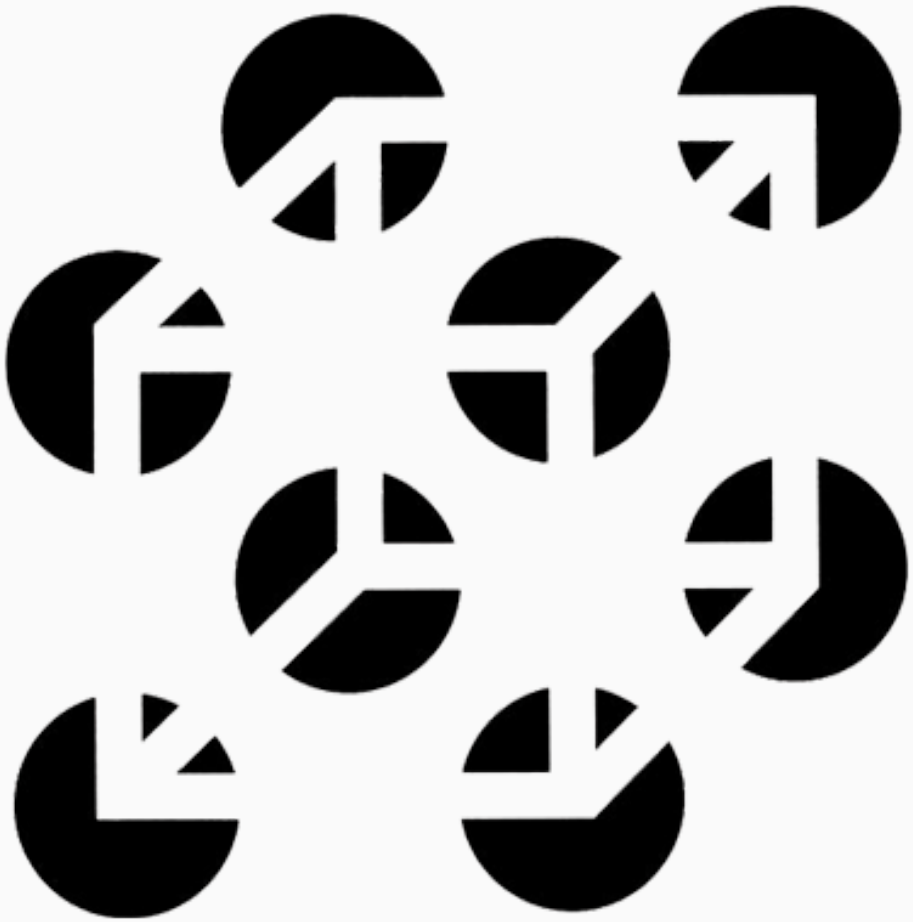
Students will also have a working understanding of Construction Morphology (CxM):

- What a construction and a constructional gestalt are
- How CxM relates to IA, IP, and WP theories of morphology
- What kinds of phenomena it captures well
- How Geert Booij's CxM notation works

Constructional Gestalts

Describe this Picture

Gestalt



Top-Down Reasoning

We often think of reasoning about meaning in language as a bottom-up process:

- We identify a set of atoms (like morphemes)
- We parse them into larger structures
- We use general principles to predict the meaning of the whole from the meaning of the parts
- Compositionality

However, there is a lot of evidence that human reasoning about the meaning of language is partially top-down:

- The probable meaning of an utterance is predicted from its context, both linguistic and extralinguistic
- The meanings of the parts are inferred, partially, from the meaning of the whole
- The meanings of the parts which have been observed in other utterances also act as constraints on the interpretation of the whole

The Correlative Construction

Consider the following sentences:

(1) a. The bigger the better!
b. The smaller the cheaper!
c. The smarter the more interesting!
d. The more convoluted the more highly valued!

The word *the* is being used in a very unique way in these sentences. Note that you cannot substitute another determiner:

(2) a. * A bigger a better!
b. * A smaller a cheaper!
c. * This smarter this more interesting!
d. * This more convoluted this more highly valued!

The meaning comes from the construction as a whole, not a compositional combination of its parts.

The Result

The meaning of a word/phrase/sentence can be predictable without being compositional.

Construction Grammar allows you to make such predictions.

Subordinate Compounds in Chinese

田鼠	*tianshu*	field-mouse	field mouse
唇膏	*chungao*	lip-ointment	lipstick
炮弹	*paodan*	artillery-bullet	artillery shell
书包	*shubao*	book-container	satchel

Coordinate Compounds in Chinese

父母	*fumu*	father-mother	‘parents’
花木	*huamu*	flower-tree	‘vegetation’
天地	*tiandi*	heaven-earth	‘universe’
国家	*guojia*	country-home	‘nation’
風水	*fengshui*	wind-water	‘geomancy’

Schemas for Chinese N-N Compounds

It is possible to give a general schema for each of these types of compounds that states how they are the same and how they are different.

(3) a. $[[x]_{Ni}[y]_{Nj}]_N \leftrightarrow$ 'a sort of SEM_j with properties of SEM_i'

b. $[[x]_{Ni}[y]_{Nj}]_N \leftrightarrow$ 'the union of SEM_i and SEM_j'

The two schemas are the same in that they involve the concatenation of two strings, x and y, both of which are nouns (N). In form (what is on the left side of the arrow), they are identical.

What differs is the semantics (what is on the right side of the arrow). While in (3a), constituent *i* modifies constituent *j*, in (3b), the semantics of the whole are an equal function of both of the parts.

Grammaticalization and Lexicalization

Grammaticalization in English

GRAMMATICALIZATION is the process by which **lexical words** become **grammatical markers**. Consider *going to* in the following example:

(4) I'm going to get married.

If *go* means 'travel along a path' then this means, 'I'm traveling along a path in order to get married.' This was the historical meaning of expressions like this. However, currently the normal reading of this example is an implication of the earlier reading, namely 'I will get married.' The expression *going to* even has a reduced form, *gonna*, in many varieties of English:

(5) I'm gonna get married.

Note that it is not the case that every instance of *going to* can be replaced by *gonna*:

(6) a. He's going to the church.
b. * He's gonna the church.

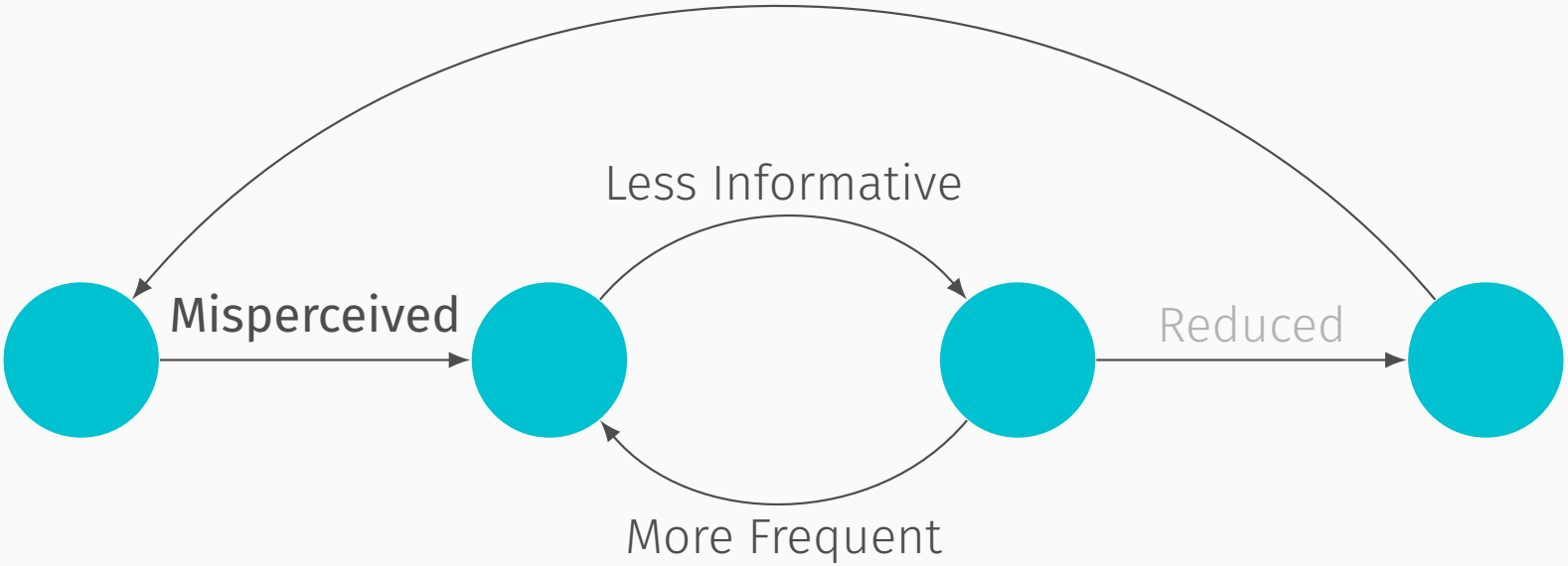
The expression *going to* has gone from being a verb plus a particle to being a grammatical marker of tense.

Information and Grammaticalization

One idea of how and why grammaticalization occurs:

- Linguistic expressions have implications that are more abstract than their most concrete, literal meaning
  - Traveling to be married implies being married in the future
  - Saying *one girl* can imply that the girl is being introduced for the first time
- From conversational and linguistic context, language learners often pick up on the more abstract meaning; a feedback loop ensues
  - As expressions occur in a greater number of contexts, their informativeness goes down
  - As the informativeness of expressions goes down, they are able to occur in a greater number of contexts
- Since less attention is paid by learners to less informative expressions, they are able to atrophy phonologically (phonological reduction)

Information and Grammaticalization



Morphologization

Hmong N-V Compounding

tsuv-dlub	tiger-be black	‘panther’
taum-mog	bean-be soft	‘pea’
toj-sab	hill-be tall	‘highlands’
nteeg-tuag	funeral-die	‘funeral’
khoom-muag	goods-sell	‘goods for sale’

A Schema for N-V Compounding

$$[[x]_{Ni}[y]_{Vj}]_N \leftrightarrow \text{`SEM}_i \text{ modified by SEM}_j\text{'}$$

Hmong Nominalizations

kev-noj	way-eat	‘eating’
kev-haus	way-drink	‘drinking’
kev-kaaj	way-bright	‘brightness’
kev-zoo	way-good	‘goodness’
kev-phem	way-bad	‘evil’
kev-kawm	way-study	‘studying’
tub-txib	son-send	‘messenger’
tub-khaiv	son-send	‘servant’
tub-nyag	son-steal	‘thief’
tub-ncig	son-be_around	‘funeral helpers’

Cline of Morphologization

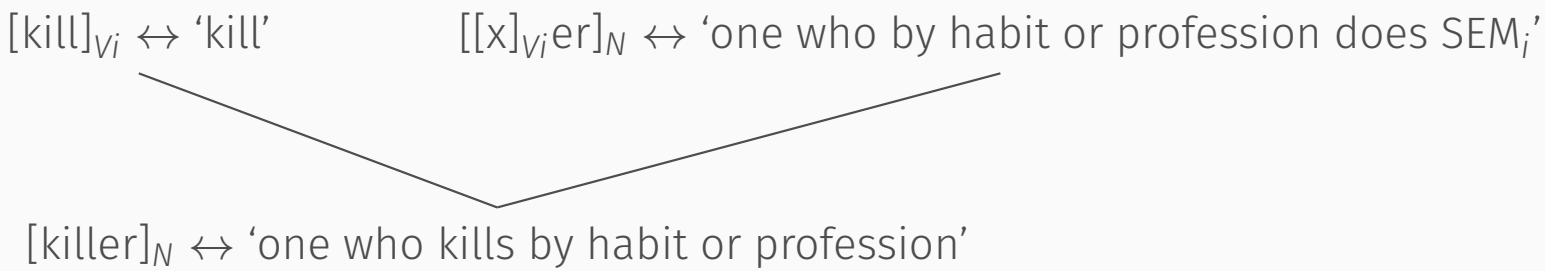
Relative Clause	Compounding	Affixation
tub (kws) nyag	⟶ *tub-nyag*	⟶ *tub-nyag*
$[[x]_{Ni}$ *kws* $[y]_{Vj}]_{NP} \leftrightarrow$ 'SEM$_i$ relating to SEM$_j$'	$[[x]_{Ni}[y]_{Vj}]_N \leftrightarrow$ 'SEM$_i$ relating to SEM$_j$'	$[tub[y]_{Vi}]_N \leftrightarrow$ 'person relating to SEM$_i$'

Constructions and Inheritance

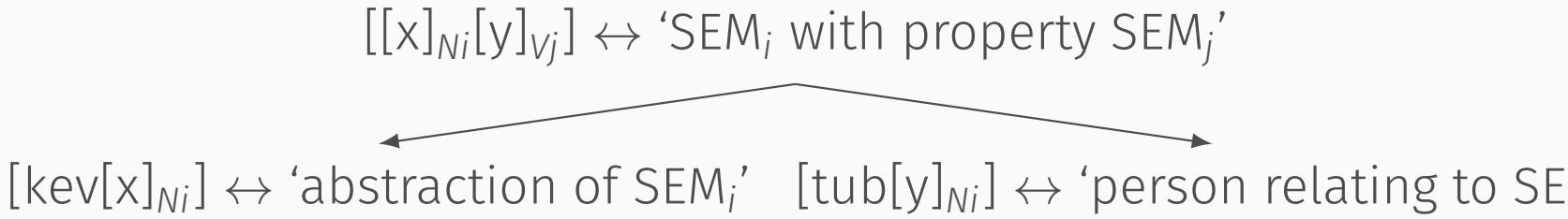
Inheritance

Constructions do not exist in a vacuum. One common way of capturing their relationships with one another is through inheritance.

- Instances inherit properties from more abstract schemas
- More concrete schemas inherit properties from still more abstract schemas
- Multiple inheritance is possible



Another Example of Inheritance



Questions?