

Lexicalization

How meanings are packed into words

January 23, 2017

Reading and homework

- Read for this week:
 - Pages 7-26 of Language Construction Kit
 - Browse pages 190-240 of Language Construction Kit
- Read for Next week:
 - Pages 27-56 of Language Construction Kit
 - Pages 7-47 of Conlanger's Lexipedia
 - Some of it is a list of frequent words from sci fi and fantasy
- Homework due Wednesday, January 25:
 - Write 1-3 pages about your world.
 - Include three consequences for the language spoken in your world

Put and Take Videos

- Created by linguists for field work
- Illustrate known parameters of variation in what words are used for putting and taking
- <http://fieldmanuals.mpi.nl/volumes/2004/put-project/>
- Bowerman, Melissa, Marianne Gullberg, Asifa Majid & Bhuvana Narasimhan. 2004. Put project: the cross-linguistic encoding of placement events. In Asifa Majid (ed.), Field Manual Volume 9, 10-24. Nijmegen: Max Planck Institute for Psycholinguistics.
- Events of Putting and Taking, 2012, Anetta Kopecka and Bhuvana Narasimhan (eds.), John Benjamins Publishing Company
 - Jahai, Mandarin, Tzeltal, Khoisan click language whose name I don't have the characters for, Moroccan Arabic, Basque, Spanish, Romansh, Swedish, Hungarian, Tamil, Hindi, Thaayorre, Japanese, Yeli Dnye, Lowland Chontal, Polish, Kalasha

The purpose of this exercise

- Awareness of what meanings are being expressed
- Awareness of different ways that languages can package those meanings into words and
- So that you can create a conlang that is not just English (or other language) written in a code

Describing Motion Events

- An *Agent* causes a *Figure* to move with respect to a *Ground*
- *Source*: where the figure comes from
- *Goal*: where the figure ends up
- *Path*: from the source to the goal
- *Instrument*: may be used by the agent to move the figure

What to pay attention to in the video

- Comparing video clips, which differences are illustrated in the agent, what the agent does, the figure, and the ground?
- Think about which English verbs and prepositions can be used to describe each video.
- If you speak another language, how does that language differ from English in the description of each event.
 - Pay attention to:
 - How many verbs are used: e.g., unscrew the lid from the jar vs take/remove the lid, unscrewing

Beware if you speak another language

In principle, yes.

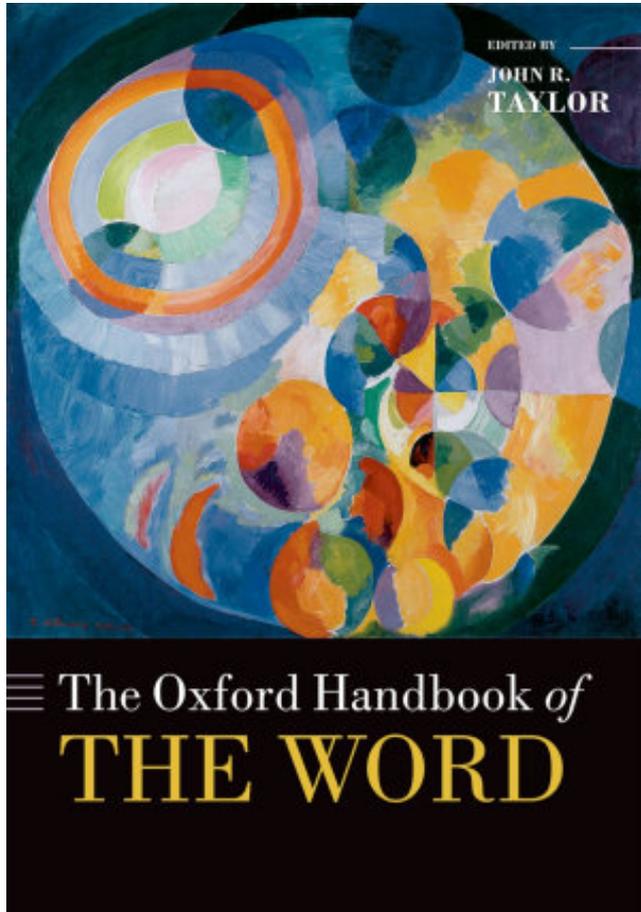
- <https://bratislavaguide.com/archive/radio-yerevan-jokes>
- Question to Radio Yerevan: "Is it correct that Grigori Grigorievich Grigoriev won a luxury car at the All-Union Championship in Moscow?"
Radio Yerevan answered: "In principle, yes. But first of all it was not Grigori Grigorievich Grigoriev, but Vassili Vassilievich Vassiliev; second, it was not at the All-Union Championship in Moscow, but at a Collective Farm Sports Festival in Smolensk; third, it was not a car, but a bicycle; and fourth he didn't win it, but rather it was stolen from him."

There is a tendency to think that because you translate "I don't have a book" into Hebrew as "en li sefer" that they are exactly the same. They are not the same at all. "en" is a negative existential particle. "Li" is a first person dative case pronoun.

So in principle, yes. But instead of a main verb plus an auxiliary verb and a negative word (don't have) there is a negative existence particle and instead of a nominative pronoun there is a dative pronoun. And instead of an indefinite article "a" there is nothing. And "en" doesn't really *mean* "doesn't have": "have" is used for possession and part-whole relationships (e.g., the table has legs" whereas "en" is used for existence and location ("en sefer" There isn't a book), ("en sefer ba shulxan" There isn't a book on the table).

What is a word?

- From Crystal, Cambridge Encyclopedia of Language (Page 91)
 - Potential pause: people usually don't pause in the middle of a word.
 - Indivisibility: except for expletive insertion “fan-freakin-tastic” words are usually not inserted in the middle of words.
 - Minimal free forms (from Bloomfield):
 - But some words never stand on their own: e.g., “the” and “of”
 - Phonetic boundaries: some phonetic processes happen within words:
 - Stress the nth or nth to last syllable
 - Vowel harmony



About 1000 pages. \$139.99

You don't have to read it.

The point is that it takes 1000 pages just to survey the issues related to what words are.

Words and orthography

- We will return to orthography later in the semester.
- Some languages use spaces between words.
 - But with some arbitrariness or strange sub-regularities:
 - I have a baseball, basketball, football, golf ball.
 - Spanish clitics: *él me lo dio, dámelo*
 - Note that the clitics don't cause the stress to move to the penultimate syllable.
 - Sotho uses spaces where other Bantu languages don't:
 - Faass et al. (2009)

a ba tsene

hortative particle + subject concord cl. 2 + verb stem

let + subj-cl2 + come in

”let them come in”

Other types of sub-sentential units

- Japanese bunsetsu (example from Wikipedia)
 - A phonological phrase in which “the pitch can have at most one fall” (Wikipedia).
 - But Japanese also has a word for “word” (tango).

太陽が | 東の | 空に | 昇る。

taiyō ga | higashi no | sora ni | noboru

sun SUBJECT | east POSSESSIVE | sky LOCATIVE | rise

The sun rises in the eastern sky.

Warnings about Rosenfelder's word frequency list in Conlanger's Lexipedia

- All the words on page 36 are grammar words.
- When you make up your conlang, you will think about what meanings these words express and how you want to express that in your language.

Some of my favorite things about words

- What is a word?
- They have a word for it
- They don't have a word for it
 - Chinese in-out-door
- They have n words for it
- Historical depth: lord, lady, snore, snot, sneeze, gleam, glisten, glow, glitter
- Compounds vs mono-morphemic
- Languages with very few verbs (have, be, do) or very few nouns (grass)
- Lexical conflation (figure, manner, ground)
 - Conflation vs serialization (manner of motion, resultative secondary predication)
- What it's like to be an Eskimo: operationalization, happinesslessness
- Words from different sources: multiple rounds of Scandinavian and romance (French and Latin)
- Backformations, folk etymologies

Linguistic Relativism

- Also known as the Sapir-Whorf hypothesis
- Language influences thought
- Take a pragmatic view of linguistic relativism in the real world.
 - Of course language and thought are related
 - But people can have thoughts that transcend their language
- In the world of conlangs?
 - Have fun with it

Words and Culture

- They have a word for it:

- http://www.slate.com/blogs/lexicon_valley/2013/12/27/finnish_the_language_has_a_word_for_the_distance_that_a_reindeer_can_travel.html

Reindeer can't walk too far without answering the call of nature. In fact, they are unable to walk and pee at the same time, so they have to take a bathroom break roughly every 6 miles. In Finnish, this distance is known as "poronkusema" or "reindeer's piss" and was an old-fashioned description of distances in the countryside.

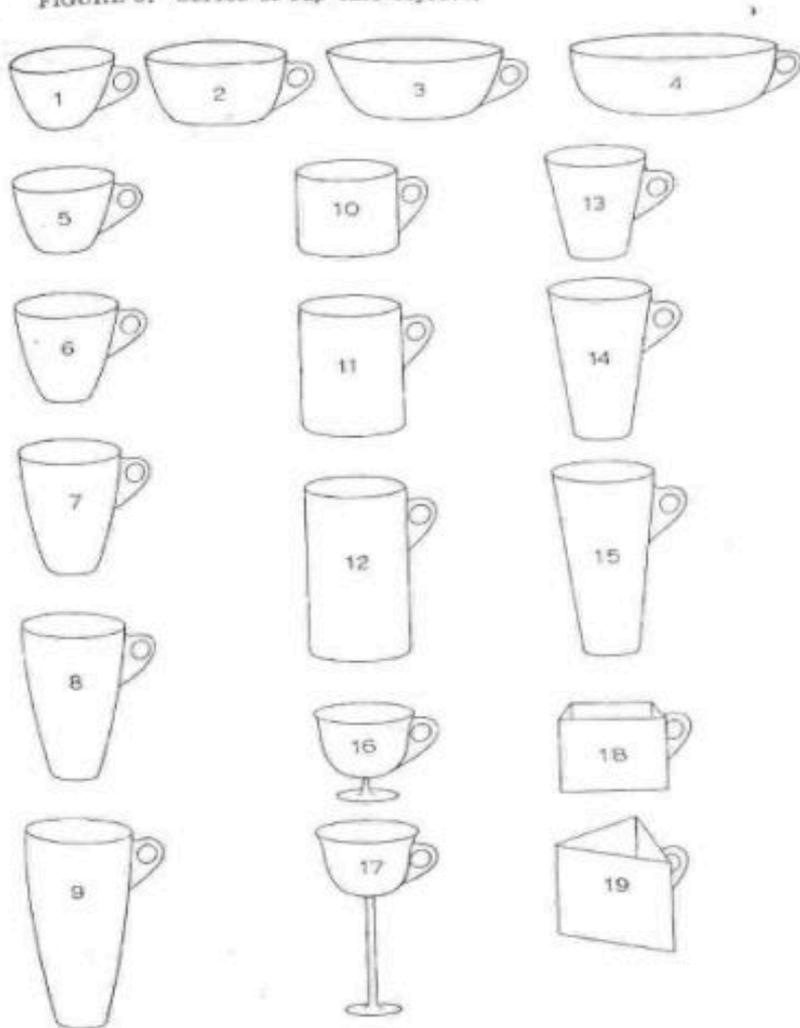
Words and Culture

- They don't have a word for X:
- You can follow this on Language Log:
 - No word for “rape” in Urdu
 - No word for “sorry” in Tagalog
 - No word for “please” in Icelandic
 - No word for “looting” in Japanese
 - No word for dyslexia in languages with good spelling

Words and Culture

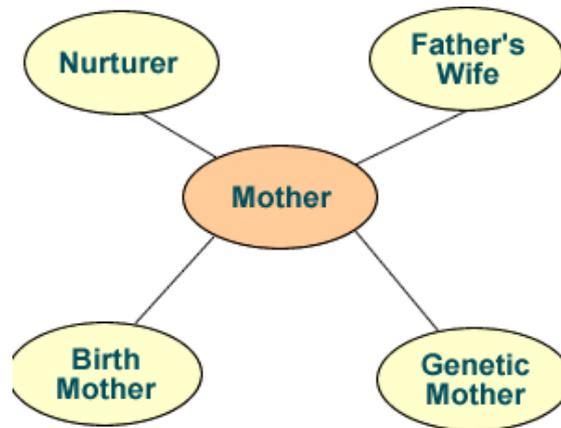
- They have N words for it:
- You can follow this on Language Log too:
 - Words for “snow”, etc.

FIGURE 5. Series of cup-like objects.



Prototype Theory

Rosch E.H. (1973): "Natural categories", *Cognitive Psychology* 4, 328-350.



Prototype and radial categories

Cognitive Linguistics: An Introduction, Vyvyan Evans and Melanie Green, Edinburgh University Press, 2006.

Women, Fire, and Dangerous Things, George Lakoff, University of Chicago Press, 1987.

http://www.linguistics.ucsb.edu/faculty/cumming/ling50/radial_categories.htm

Copied on August 23, 2010

A prototypical mother provides genetic material, provides gestation, is the father's wife, and nurtures the child.

- A birth mother could be a surrogate: she gestates the baby but does not provide genetic material, nurture the child, or marry the father.
- The father's wife might be a step mother and nurture the child, but she does not provide genetic material or gestate the child.
- A nurturer could be an adoptive mother: she doesn't provide genetic material or gestation and may not marry the genetic father.
- A genetic mother could be an egg donor and may fill none of the other functions of "mother."

Polinsky, “Headedness, again*”

- http://scholar.harvard.edu/files/mpolinsky/files/keenon_paper.online.pdf

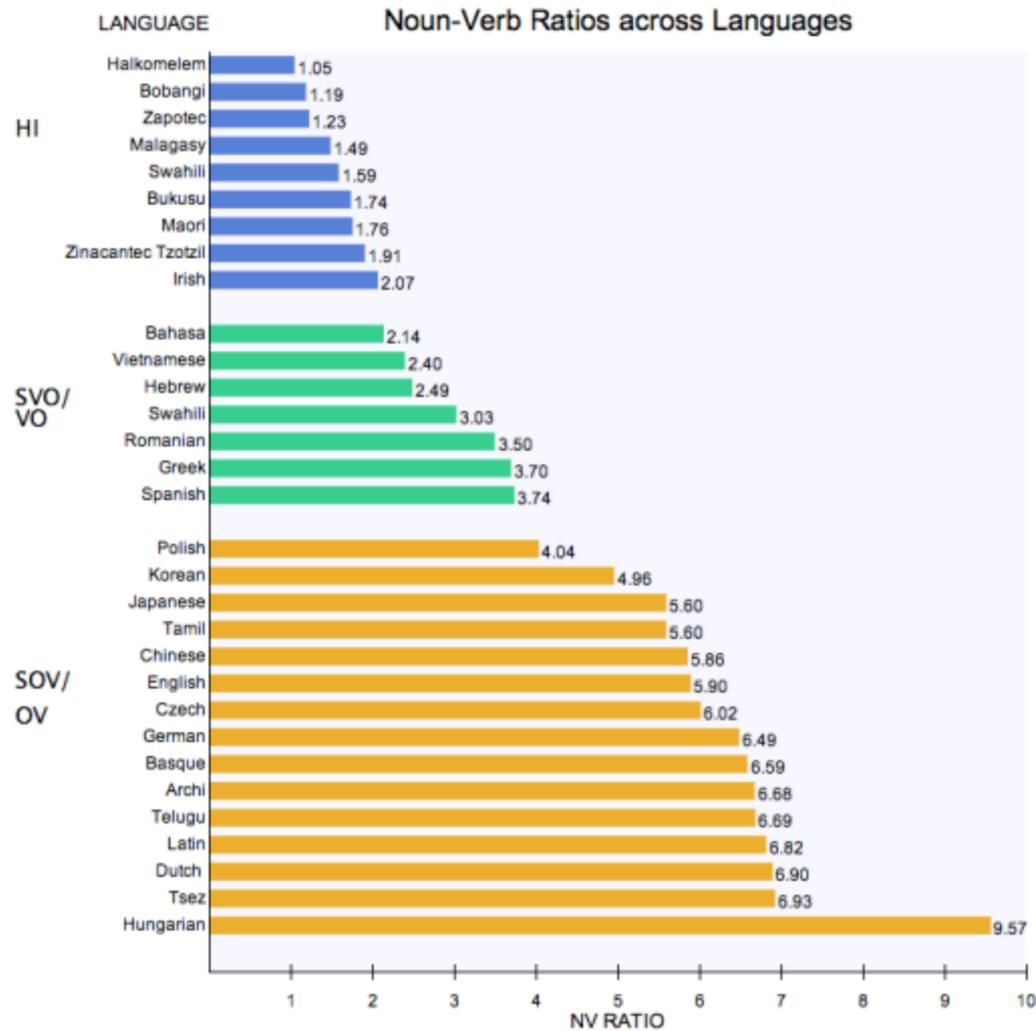
(4) Nouns and verbs across languages: Numerical comparison of lemmas⁵

	Nouns	Verbs	Noun-to-verb ratio
Japanese	86028	15346	5.6
Korean	89125	17956	4.96
Tamil	2403	423	5.6
Telugu	3489	521	6.69
Archi* (Kibrik et al. 1977)	2419	362	6.68
Tsez* (Xalilov 1999)	3508	506	6.93
Hungarian	31600	3300	9.57
Basque	23069	3496	6.59
Latin* (Aronoff 1994; Minozzi 2009)	4777	700	6.82
German* (Barbara Stiebels, p.c.)	72785	11201	6.49
Dutch (average of WordNet and CELEX)	59182	8549	6.9
English	82115	13767	5.9
Chinese* (Xu et al. 2008)	78764	13430	5.86
Polish	14131	3497	4.04
Czech	31029	5158	6.02
Greek	29782	7839	3.7
Romanian	56594	16122	3.5
Spanish	48323	12910	3.74

Hebrew	11961	4804	2.49
Vietnamese	6000	2500	2.4
Bahasa (Indonesian/Malay)	12429	5805	2.14
Zinacantec Tzotzil* (Haviland 1994)	1629	850	1.91
Halkomelem* (Galloway 2009)	967	916	1.05
Zapotec* (Long and Cruz 1999)	542	439	1.23
Irish (Modern, from 1800)	1850	890	2.07
Swahili	6150	3853	1.59
Bukusu* (Larry Hyman, p.c.)	2879	1653	1.74
Bobangi* (Larry Hyman, p.c.)	3973	3324	1.19
Malagasy* (Diksionera 1973)	5436	3643	1.49
Maori* (Williams 1957)	2920	1656	1.76

⁴ For languages marked with an asterisk, the data come from dictionaries or published sources indicated in parentheses; all other numbers are from WordNet, CELEX, and/or corpora. I am grateful to Eneko Agirre, Francis Bond, Verena Hinrichs, Katia Kravtchenko, Sun-Hee Lee, Dan Tufis, and Shuly Wintner for help with the counts.

⁵ Where it is relevant, the counts exclude compound verbs formed using a light verb, as in the Japanese examples above.



Polinsky

Udi: A language with few verbs

- <http://wschulze.userweb.mwn.de/sampletxt.htm>
- **sa pasc^'ag/-en sa pasc^'ag/-ax c^'ax-p-i**
- one king-erg one king-dat2 catch-**lv**:trans-aor
- A king [having] caught a king (and)
-
- **yesir-re-aq'-sa e-ne-sc^a ic^ öлки-n-a ic^ k'ua e-ne-f-sa**
- **prisoner-3sg:a-take-pres** carry-3sg:a- $\text{\$}$:pres refl land-sa-dat1 refl home:Dat1 keep-3sg:a- $\text{\$}$ -pres.
- IMPRISONes (him), CARRIEs (him) to his own land, KEEPs (him) in his own house.
-
- **s^e-t'-a pasc^'ag/lug/-ax-al zaft-t'e-b-sa.**
- dist-sa:obl-gen kingdom-dat2-foc rule-3sg:a-**lv**-pres
- He RULEs over that kingdom, too.

Udi: A language with few verbs

- <http://wschulze.userweb.mwn.de/sampletxt.htm>
- pasc[^]ag/-en xabar-re-aq'-sa me-t'-uxo te
- king-erg **news-3sg:a-take-pres** prox-sa:obl-abl quote
- The king ASK-s this one:

A language with very few nouns

- I heard a talk about it once, but I don't remember what it was.
- A lot of plants were kinds of grass.
- Maybe the language was in Kansas.
- Or the person who gave the talk was from Kansas.
- Maybe the language is extinct. I remember something about all the speakers being old and then there was a flu epidemic. And then there were even fewer speakers (around 4). And that was 20 years ago.

Lexicalization of motion events

- Leonard Talmy (1985) Lexicalization Patterns: Semantic Structure in Lexical Form
 - Figure
 - Ground
 - Manner
 - Path
 - Cause
- The bottle went floating on the river into the cave.
- figure move manner ground path

Talmy (1985) Lexicalization

- The bottle went into the cave.
 - “went” expresses move
- The bottle entered the cave.
 - “enter” expresses Move + path
- Deplane, debark, detrain
 - Move+ground (rare)
- The paper blew off the table
 - “blow” expresses move+cause
- The bottle floated into the cave
 - “float” expresses move+manner
 - Many languages do not easily lexicalize move+manner
 - Vs. The bottle went into the cave floating

Talmy, 13C, page 67

- **Give** me the flour + **move** it down from the shelf + **reach** it with your free hand →
- Could you **reach** me the flour down off the shelf with your free hand.

Talmy (pages 66-68)

Many languages cannot lexicalize these meanings

- Example 13A: Conflation of motion with something other than a manner of motion:
 - Wear+go: wear a dress to the party.
 - Read+go: read comics all the way to NY
 - Induce+go: lure/scare someone out of a hiding place
 - Aim to induce to go: urge/wave/beckon him away from the building
 - “wave” includes a manner too.

13D, page 67-68

- Change of state (become) plus manner:
 - Choke to death
 - Flap dry in the wind
 - Rust stiff
 - Wear thin
 - Freeze stuck
- Many languages do not lexicalize this easily
 - Die choking; become stiff by rusting
- Come into existence (form) + manner
 - a hole burned in the table

Put Project

- Bowerman, Melissa, Marianne Gullberg, Asifa Majid & Bhuvana Narasimhan. 2004. Put project: the cross-linguistic encoding of placement events. In Asifa Majid (ed.), Field Manual Volume 9, 10-24. Nijmegen: Max Planck Institute for Psycholinguistics.
- Stable URL
 - <http://fieldmanuals.mpi.nl/volumes/2004/put-project/>

Putting and Taking

Kopecka and Narasimhan (2012)

- The figure:
 - Rigid or flexible
 - Body part of the agent or not
 - Granular vs liquid vs solid
 - Clothing item
 - Big vs small
 - Shape (e.g., long, round)
- The ground
 - Animate vs inanimate
 - Body part of agent or not
 - Horizontal vs vertical
 - Floor or higher
 - Container vs supporting surface
- Spatial relation
 - Containment vs support
 - Tightly or loosely fitted
 - Figure is vertical or horizontal wrt ground
 - Figure suspended from a point
- Instrument
 - Agent's hand, mouth, vs other instrument
- Manner
 - Dropping, throwing, placing, pouring
 - Agent maintains control until figure reaches destination or not
 - Agent moves with the figure or just moves his/her hand

Motion+ground+path (page 77)

- Shelve books (onto shelves)
- Box the apples (into boxes)
- Quarry granite (from the quarry)
- Mine bauxite (from the mine)

Talmy (1985) Move+Figure

- (18) Atsugewi verb roots of Motion with conflated Figure
- lup- 'for a small shiny spherical object (e.g. a round candy, an eyeball, a hailstone) to move/be-located'
 - t̥- 'for a smallish planar object that can be functionally affixed (e.g. a stamp, a clothing patch, a button, a shingle, a cradle's sunshade) to move/be-located'
 - caq- 'for a slimy lumpish object (e.g. a toad, a cow dropping) to move/be-located'
 - swal- 'for a limp linear object suspended by one end (e.g. a shirt on a clothesline, a hanging dead rabbit, a flaccid penis) to move/be-located'
 - qput- 'for loose dry dirt to move/be-located'
 - staq̥- 'for runny icky material (e.g. mud, manure, rotten tomatoes, guts, chewed gum) to move/be-located'

Talmy (1985) Move+figure in English

- Spit into the cup.
- Pee into the cup.
- Rain on the lawn.
- Ooze under the door.