Words

February 3, 2015
How to do things with words (but not as Austin intended)

- 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, happinesslessnessless
- Words from different sources: multiple rounds of Scandinavian and romance (French and Latin)
- Backformations, folk etymologies
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
“If a word without an accent mark ends in a vowel, *n* or *s*, the stress is on the penultimate (next to last) syllable. For example, *toro*, *computadora*, *joven* and *zapatos* all have their accent on the next-to-last syllable. Most words fit this category.

A word without an accent mark that ends in other letters has the stress on the last syllable. For example, *hotel*, *hablar*, *madador* and *virtud* all have the accent on the final syllable.

If a word isn't pronounced according to the above two rules, an accent is placed over the vowel of the syllable that gets the stress. For example, *común*, *lápiz*, *médico*, *inglés*, and *ojalá* all have the stress on the indicated syllable.”
Spanish stress rules
(Warning: this page may contain errors.)

• Contar (tell.inf) (to tell)
  – last syllable (does not end in vowel, n, or s)

• Cuento (tell.1.sg.pres(subj)) (I tell)
  – second to last syllable (ends in a vowel)
  – “o” changes to “ue” when stressed (?)

• Contamos (tell.1.pl.pres(subj)) (We tell)
  – second to last syllable (ends in s)
  – Note that the stress is not on the first stem vowel.
  – The suffix made the word longer and the stress moved to follow the stress rule.

• Cuenta (tell.imperative) (Tell!)
  – second to last syllable (ends in a vowel)

• Cuéntalo (tell.imperative.3.sg.m(obj)) (Tell it!)
  – The stress is where the accent is, not on the second to last syllable as expected. Why? The clitic
    pronoun “lo” is not a tightly affixed affix. It does not cause the stress to move. In other words, the
    stress is on the second to last syllable. “Lo” is not strictly part of the word.

• Cuente la historia. (Tell the story)
  – Stress is on the “e” (second to last syllable of the first word)
  – The word “la” does not cause the stress to move on the previous word.
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)
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。
\textit{taiyō ga higashi no sora ni noboru}
sun SUBJECT | east POSSESSIVE | sky LOCATIVE | rise
The sun rises in the eastern sky.
Categories of Words
Based on

• Payne, Analyzing Morphosyntax, Chapter 3
• Schachter, Parts-of-speech systems, in Shopen (ed.) Language Typology and Syntactic Description
Categories of Words: *Parts of Speech*

- Noun
- Verb
- Adjective
- Adverb
- Adposition (preposition, postposition)
- Determiner (Article)
- Etc.
Problems with the semantic definitions of parts of speech

- The same concept can function in several parts of speech.
  - Pinker, page 98
- Her *interest* in fungi (noun)
- Fungi are starting to *interest* her more and more. (verb)
- She seems *interested* in fungi. (adjective)
- *Interestingly,* the fungi grew an inch in an hour. (adverb)
The distributional theory of parts of speech

• “A part of speech, then, is not a kind of meaning; it is a kind of token that obeys certain formal rules, like a chess piece or a poker chip.”
  – Pinker, page 98

• Testable and falsifiable

• Assumes discrete categories
Scientific method in linguistics

- Theories (hypotheses) must be testable and falsifiable.

- If I say that “tree” is a noun, I have to provide my criteria for saying so.
  - It can be the subject or object of a verb
    - The tree fell.
    - I saw the tree.
  - it can be modified by an adjective or a possessive pronoun
    - My tree
    - The green tree
Criteria for identifying parts of speech

• Distribution
  – The contexts where the word can appear
  – The functions it can have

• Morphology
  – Prefixes, suffixes, and other changes to the structure of the word.
English Nouns

• Can be subjects and objects
• Can have plurals (*book*, *books*)
• Can have determiners (articles)
• Can be modified by adjectives
• Can have possessors (*my book*)
Scientific method in linguistics

• Results must be reproducible.

• If two people are trying to decide whether “walking” is a noun, they should apply the criteria and come up with the same answer.
  – Walking is fun. (subject)
  – I like walking. (object)
  – His walking is awkward
    • Possessive pronoun
  – One should avoid fast walking.
    • Modified by adjective
However....

• Some members of a category do not pass all the tests.
  – Payne points out that if you are doing fieldwork and trying to determine the typical properties of nouns, you should not start with the word “fist” because it is not stable over time (like “tree” or “house”), and may not be expressed as a simple noun.
  – Nouns derived from verbs (e.g., “walking”) may also be less typical.
    • In English, it is hard to get a plural for “walking”
      – ?? His many night time walkings
    • In English, “walking” can be modified by an adjective (like a noun) or by an adverb (like a verb)
      – Walking is fun.
      – Slow walking is fun.
      – Slowly walking is fun/Walking slowly is fun.
Special types of English nouns

• Pronouns
  – no determiners -- * the she

• Mass nouns
  – no plurals -- *two informations

• Proper nouns
  – no determiners -- *the Mt. Rushmore
  – plurals under special circumstances
    • two Kaylas, a couple of Bens

• Gerunds
  – deverbal (derived from verbs) nouns with -ing
A few things that are different from English

- Language doesn’t distinguish nouns and verbs
- Language doesn’t distinguish nouns and adjectives
- Language doesn’t distinguish verbs and adjectives
- Same concept expressed by different parts of speech in different languages
Interlinear glosses
(aka three line glosses)

• Schachter, page 11

Mamu.k-ma qu.?as-?i
working-PRES(INDIC) man-DEF
‘The man is working.’

• Meaning
  – Both languages have the meaning of work and man.
  – Both sentences express present time, indicative mood.
  – “Man” is definite (specific and identifiable) in both languages.

• Form
  – English has an auxiliary verb
  – English has a present participle
  – English has a determiner (the)
Nootka nouns and verbs

• Schachter, page 11

Mamu.k-ma qu.?as-?i
working-PRES(INDIC) man-DEF
‘The man is working.’

Qu.?as-ma mamu.k-?i
man-PRES(INDIC) working-DEF
‘The working one is a man.’

• Any word can take a definiteness suffix and be an argument of a “verb.”

• Any word can take a tense marker and be a predicate.
Hausa adjectives and nouns

• Schachter, page 15

mutum mai alheri
person havng kindness
“a kind person.”

mutum mai doki
person having horse
“a person having a horse”
Mandarin adjectives and verbs

(38) Neige nühaizi piaoliang, that girl beautiful 'That girl is beautiful'

(39) Neige nühaizi liaojie that girl understand 'That girl understands'

(40) piaoliang de nühaizi beautiful REL girl 'a girl who is beautiful, a beautiful girl'

(41) liaojie de nühaizi understand REL girl 'a girl who understands, an understanding girl'
Mojave Adjectives and Verbs

(42) ʔi:pa-č  homi:-k (iðu:m)
man-subj tall-pres (aux)
‘The man is tall’

(43) ʔi:pa-č  su:pa:w-k (iðu:m)
man-subj know-pres (aux)
‘The man knows’
Mojave adjectives and verbs

(44) ʔi:pa kʷ-su:paw-n⁷-č iva:k
man REL-know-DEM-SUBJ is here
'The man who knows is here'

(45) ʔi:pa (kʷ-)homi:-n⁷-č iva:k
man (REL-)tall-DEM-SUBJ is here
'The tall man is here'
Different from English

(48) a. Ya fi ni hankali
he(PERF) surpass me intelligence
‘He is more intelligent than I am’

b. Ya fi su duka hankali
he(PERF) surpass them all intelligence
‘He is the most intelligent of them all’

Schachter, page 22, Hausa

“Than” is a preposition in
English.

“fi” is a verb in Hausa.
Different from English

Schachter, page 38,

Akan

(107) Wënoa a aduan a edëso
they cooked food REL it is enough
‘They cooked enough/a lot of food’

Hausa

Mutane da yawa
People with abundance
‘many people’
Polinsky, Headedness, again*


<table>
<thead>
<tr>
<th>Language</th>
<th>Nouns</th>
<th>Verbs</th>
<th>Noun-to-verb ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japanese</td>
<td>86028</td>
<td>15346</td>
<td>5.6</td>
</tr>
<tr>
<td>Korean</td>
<td>89125</td>
<td>17956</td>
<td>4.96</td>
</tr>
<tr>
<td>Tamil</td>
<td>2403</td>
<td>423</td>
<td>5.6</td>
</tr>
<tr>
<td>Telugu</td>
<td>3489</td>
<td>521</td>
<td>6.69</td>
</tr>
<tr>
<td>Arche* (Kibrik et al. 1977)</td>
<td>2419</td>
<td>362</td>
<td>6.68</td>
</tr>
<tr>
<td>Tsez* (Xalilov 1999)</td>
<td>3508</td>
<td>506</td>
<td>6.93</td>
</tr>
<tr>
<td>Hungarian</td>
<td>31600</td>
<td>3300</td>
<td>9.57</td>
</tr>
<tr>
<td>Basque</td>
<td>23069</td>
<td>3496</td>
<td>6.59</td>
</tr>
<tr>
<td>Latin* (Aronoff 1994; Minozzi 2009)</td>
<td>4777</td>
<td>700</td>
<td>6.82</td>
</tr>
<tr>
<td>German* (Barbara Stiehls, p.c.)</td>
<td>72785</td>
<td>11201</td>
<td>6.49</td>
</tr>
<tr>
<td>Dutch (average of WordNet and CELEX)</td>
<td>59182</td>
<td>8549</td>
<td>6.9</td>
</tr>
<tr>
<td>English</td>
<td>82115</td>
<td>13767</td>
<td>5.9</td>
</tr>
<tr>
<td>Chinese* (Xu et al. 2008)</td>
<td>75764</td>
<td>13430</td>
<td>5.86</td>
</tr>
<tr>
<td>Polish</td>
<td>14131</td>
<td>3497</td>
<td>4.04</td>
</tr>
<tr>
<td>Czech</td>
<td>31029</td>
<td>5158</td>
<td>6.02</td>
</tr>
<tr>
<td>Greek</td>
<td>29782</td>
<td>7839</td>
<td>3.7</td>
</tr>
<tr>
<td>Romanian</td>
<td>56594</td>
<td>16122</td>
<td>3.5</td>
</tr>
<tr>
<td>Spanish</td>
<td>48323</td>
<td>12910</td>
<td>3.74</td>
</tr>
</tbody>
</table>

Hebrew: 11961 4804 2.49
Vietnamese: 6000 2500 2.4
Bahasa (Indonesian/Malay): 12429 5505 2.14
Zinacantec Tzotzi* (Haviland 1994): 1629 850 1.91
Halcomelem* (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* (Diksoniera 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 Endo Aigrite, Francis Bead, Verena Hinrichs, Katia Krivchikov, Sun-Hye Lee, Dan Tufis, and Shuya Wataner 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.
Noun-Verb Ratios across Languages

<table>
<thead>
<tr>
<th>Language</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hlakolome</td>
<td>1.05</td>
</tr>
<tr>
<td>Bobangi</td>
<td>1.19</td>
</tr>
<tr>
<td>Zapotec</td>
<td>1.23</td>
</tr>
<tr>
<td>Malagasy</td>
<td>1.49</td>
</tr>
<tr>
<td>Swahili</td>
<td>1.59</td>
</tr>
<tr>
<td>Bukusu</td>
<td>1.74</td>
</tr>
<tr>
<td>Maori</td>
<td>1.76</td>
</tr>
<tr>
<td>Zinacantec Tzotzil</td>
<td>1.91</td>
</tr>
<tr>
<td>Irish</td>
<td>2.07</td>
</tr>
<tr>
<td>Bahasa</td>
<td>2.14</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>2.40</td>
</tr>
<tr>
<td>Hebrew</td>
<td>2.49</td>
</tr>
<tr>
<td>Swahili</td>
<td>3.03</td>
</tr>
<tr>
<td>Romanian</td>
<td>3.50</td>
</tr>
<tr>
<td>Greek</td>
<td>3.70</td>
</tr>
<tr>
<td>Spanish</td>
<td>3.74</td>
</tr>
<tr>
<td>Polish</td>
<td>4.04</td>
</tr>
<tr>
<td>Korean</td>
<td>4.96</td>
</tr>
<tr>
<td>Japanese</td>
<td>5.60</td>
</tr>
<tr>
<td>Tamil</td>
<td>5.60</td>
</tr>
<tr>
<td>Chinese</td>
<td>5.66</td>
</tr>
<tr>
<td>English</td>
<td>5.90</td>
</tr>
<tr>
<td>Czech</td>
<td>6.02</td>
</tr>
<tr>
<td>German</td>
<td>6.49</td>
</tr>
<tr>
<td>Basque</td>
<td>6.59</td>
</tr>
<tr>
<td>Archi</td>
<td>6.68</td>
</tr>
<tr>
<td>Telugu</td>
<td>6.69</td>
</tr>
<tr>
<td>Latin</td>
<td>6.82</td>
</tr>
<tr>
<td>Dutch</td>
<td>6.90</td>
</tr>
<tr>
<td>Tsez</td>
<td>6.93</td>
</tr>
<tr>
<td>Hungarian</td>
<td>9.57</td>
</tr>
</tbody>
</table>
Udi: A language with few verbs

- \url{http://wschulze.userweb.mwn.de/sampletxt.htm}

- \text{sa pasc^\'ag/-en sa pasc^\'ag/-ax c^\'ax-p-i}
  - one king-erg one king-dat2 catch-\text{lv:trans-aor}
  - A king [having] caught a king (and)

- \text{yesir-re-aq\'-sa e-ne-sc^a ic^\ ölki-n-a ic^ k\'ua e-ne-f-sa}
  - IMPRISONes (him), CARRIEs (him) to his own land, KEEPs (him) in his own house.

- \text{s^e-t\'-a pasc^\'ag/lug/-ax-al zaft-t\'e-b-sa.}
  - dist-sa:obl-gen kingdom-dat2-foc rule-3sg:a-\text{lv-pres}
  - He RULEs over that kingdom, too.
Udi: A language with few verbs

- [http://wschulze.userweb.mwn.de/sampletxt.htm](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.
What can words mean?
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

• What is the significance or non-significance of this?

  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
• What is the significance or non-significance of this?
Words and Culture

• They have N words for it:
• You can follow this on Language Log too:
  – Words for “snow”, etc.
• What is the significance or non-significance of this?
A well-studied example

• Verbs that describe motion
The Universe

Human communication puts a tiny piece of the universe in the common ground with the intent of getting the pictures in your heads to be the same.

Common Ground

I got this picture by searching Google Images for “common ground”.
Some things you might say about the two pictures:
A car is passing a tree.
A car drove past a tree.
The tree is in front of the car.
The tree is behind the car.
The car is behind the tree.

Some things you are less likely to say:
The tree is passing the car.
The car and the tree are moving past each other.

Starting from the whole universe, or at least as much as you can focus on, you choose to talk about a small part of it. Call this “conceptual framing”.

You decide whether to talk about motion or static location. You choose a figure and a ground (Talmy 1977). You talk about the figure being located or moving with respect to the ground (as in background).

The car (figure) is passing the tree (ground).
The tree (figure) is in front of the car (ground).

Your conceptual frames are a lossy version of the universe.

Then you have to put your conceptual frame into grammar, which is even lossier and doesn’t give you a whole lot of choices. You get to encode your conceptual frame in grammar with subject, verb, object, and prepositional phrases. Not much to work with.
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
• **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 (1985) Lexicalization

- The bottle went into the cave.
  - “went” expresses move
- The bottle entered the cave.
  - “enter” expresses Move + path
- Deplane, debark, detraining
  - 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
Move + X in English

• bounce across the room
• roll down the hill
• swim across the lake
• cartwheel across the room
• rumble down the street
• rustle across the floor
• flop onto the bed
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.
• 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
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’

-t- ‘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’

-staq- ‘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.
Put Project


- Stable URL
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