Word Order

Carl Pollard

Department of Linguistics
Ohio State University

February 7, 2012

"English is an SVO Language" (1/2)

- English is often described as having SVO as its basic (or 'canonical', 'unmarked', or 'preferred') word order

  Chiquita (S) kicked (V) Pedro (O).

  (other examples: Chinese, French, Spanish, Bulgarian)

  as compared with:
- SOV, e.g. Japanese (here '=' indicates cliticization of case markers):

  John=ga tegami=o yon-da
  John=GA letter=O read.PST
  'John read the letter.'

  (other examples: Korean, Basque, Turkish, Uzbek)
- VSO, e.g. Welsh:

  Dywedodd Gwyn y [gwelodd ef y bechgyn].
  Said Gwyn that saw he the boys
  'Gwyn said that he saw the boys.'

  (other examples: Irish, Hawaian, Tongan, Chamorro)

"English is an SVO Language" (2/2)

- VOS, e.g. Malagasy (an Austronesian language of Madagascar):

  Nahita ny mpianatra ny vehivavy.
  saw NY student NY woman
  'The woman saw the student'.

  (other examples: Fijian)

- OSV, e.g. Nadëb (a Nadahup language of Brazil):

  awad kalapéé hapúh
  jaguar child see.IND
  'The child sees the jaguar.'

  (other examples: Xavante (Brazil), Warao (Venezuela))

- OVS, e.g. Hixkaryana (a Carib language of Brazil):

  toto y-ahosi-ye kamara
  man 3:3.grab.distant-past jaguar
  'The jaguar grabbed the man.'

What is Meant by 'Basic' Word Order? (1/2)

1. Is Kim a vegan? (V-S-O, main clause polar interrogative)
2. What are they? (O-V-S, main clause constituent interrogative)
3. (I wonder) what they are. (O-S-V, embedded constituent interrogative)
4. BAGELS, I like. (O-S-V, contrastive topicalization)
   L+H* L-H%
5. She bought the Ford . . . no,
   the CHEVY she bought. (O-S-V, corrective focus)
   H* L-
6. The bigger the dog, the louder the bark. (comparative correlative)
7. No fool he!

What is Meant by 'Basic' Word Order? (2/2)

- In a DECLARATIVE, TRANSITIVE, PRAGMATICALLY UNMARKED, MAIN CLAUSE of English, the subject precedes the verb, and the verb precedes the object.
- For some languages, e.g. French, what counts as 'basic' is further circumscribed by requiring that the arguments be 'full noun phrases' as opposed to pronouns:

  1. Marie voit Jean.
     Marie sees Jean
     'Marie sees Jean.'
  2. Marie le voit.
     Marie him sees
     'Marie sees him.'

- But how do you know which is the 'subject' and which is the 'object'?

What's a 'Subject'?

- In some syntactic frameworks, notions of 'grammatical function' or 'grammatical relation' are taken as undefined theoretical primitives.
  - LFG distinguishes (inter alia) SUBJ, OBJ1, OBJ2, OBL (oblique), COMP (complement), and XCOMP (controlled complement)
  - HPSG distinguishes (inter alia) SUBJ, COMPS, SPR (specifier), MOD (modifier), and FILL (filler)
- In GB, SPEC, COMP, and ADJ were *configurationally defined*—we will come back to this.
- In contemporary categorial frameworks, there are no (primitive or defined) notions of grammatical function.

What's a 'Subject' in our LG English Fragment?

- Consider the following lexical entry:

$$\vdash \lambda_{st}.s \cdot \text{beats} \cdot t; \text{Nom3s} \multimap \text{Acc} \multimap \text{S}; \mathsf{beat}$$

- The argument we *call* the 'subject' can be identified as:
  - the one that comes to the left of the verb
  - the one that must be nominative (if it is a pronoun)
  - the one that the verb agrees with
  - the one corresponding to a certain semantic argument of the beat function (the 'agent' as opposed to the 'patient'—see below)
  - Once more of the grammar is known, we can also identify this argument as the one which can be 'raised', 'controlled', or 'passivized'.
  - In other languages, other properties are sometimes identified as 'subject' properties, e.g. ability to relativize, or to antecede a reflexive pronoun.
- But across languages, these properties may not all line up, or may not even exist.

Semantic (or Thematic) Roles (1/3)

- **Semantic roles** are *ways of participating* in the actions, states, or events described by predicates (usually but not always verbs).
- From one point of view, each such relation has its own set of semantic roles, in the sense that e.g. being the beater in a beating is different from being the feeder in a feeding.
- But semantic roles across different verb meanings with shared properties are often classified as instances of a single role (or role-type) in the interest of accounting for putative linguistic generalizations.
- Roles in this sense include, e.g. agent, patient (or theme), goal (or recipient), instrument, beneficiary, etc.
- A related notion in mainstream generative grammar is **θ-roles**, which are taken to be *syntactic* elements that 'assigned' to arguments by the syntactic entities that 'take' the arguments.

Semantic (or Thematic) Roles (2/3)

- Dowty (1991) introduced the notion of **proto-roles** as *prototypes* characterized by sets of *semantic properties* (or equivalently, entailments).
- Protypical Agent Properties ('Proto-Agent'):
  a. volitional involvement
  b. sentience or perceptivity
  c. causer of the event, or of a change of state of another participant
  d. movement (possibly relative to the position of another participant)
  e. existence independent of the event described by the verb
- Prototypical Patient Properties ('Proto-Patient')
  a. undergoes a state change
  b. incremental theme
  c. causally affected by another participant
  d. stationary relative to motion of another participant
  e. existence dependent on the action described by the verb.

Semantic (or Thematic) Roles (3/3)

Standardly assumed semantic roles can be *defined* as presence or absence of different proto-role properties, e.g.

- Agent $=_{\text{def}}$ volition + causation (+ sentience + movement)
- Experiencer $=_{\text{def}}$ sentience/perceptivity, without volition or causation
- Theme $=_{\text{def}}$ change of state (+ incremental theme + dependent existence + causally affected)
- Instrument $=_{\text{def}}$ causation + movement, without volition or sentience

Case (1/5)

- Roughly, **case** is the *morphological* expression of the grammatical relationship of an argument/modifier to the predicate that it is an argument of (or that it modifies)
- Some languages (e.g. Chinese) lack case altogether:

  Ta xihuan ta.
  s/he like him/her
  'S/He likes him/her.'
- Some (e.g. English) distinguish case only for pronouns.
- Some (e.g. K'iche') express case not on the argument but by **cross-reference** markers on the verb:

  1. x-∅-a-to' ri achi
     CMP-A3-E2-help the man
     'You helped the man.'
  2. x-at-u-to' ri achi
     CMP-A2-E3-help the man
     'The man helped you.'

Case (2/5)

- Some languages express case via **inflection** of the argument itself (e.g. Russian, Serbo-Croatian, Turkish, German):
  dom 'house' (nom.), dom-a (gen.), dom (acc.), dom-u (dat.), dom-e (loc.), dom-om (instr.)
- Some express case via **phrasal affixation**, i.e. clitics that attach after the entire argument phrase (e.g. Japanese, Korean), or before it (e.g. Tagalog):

  Bumili ang=lalake ng=isda sa=tindahan.
  Bought DIR=man IND=fish OBL=store
  'The man bought fish at the store.'
- Languages differ with respect to the number of cases, e.g. 2 (Rumanian), 3 (Tagalog), 4 (German), 6 (Russian), 15 (Finnish)

Case (3/5)

- Dixon (1979) classifies case systems based on proposed universal syntactic-semantic primitives:

  S: the single argument of an intransitive verb
  A: the more agent-like argument of a transitive verb
  O: the more patient-like argument of a transitive verb

  Unfortunately 'S' and 'O' here don't mean quite the same thing as in locutions like 'V-S-O' and 'S-V-O'!
- In **nominative/accusative** case systems (e.g. Latin, German, Russian), S patterns with A (nominative) and against O (accusative).
- In **ergative/absolutive** case systems (e.g. Basque, Tibetan, K'iche', West Greenlandic) , S patterns with O (absolutive) against A (ergative).

Case (4/5)

- In **split ergative** systems, ergativity is *conditioned*, e.g.:
  - in Hindi, the ergative pattern is followed if the verb is perfective, the accusative pattern if it is imperfective.
  - in 'split S' languages (e.g. Dakota), S of an *active* intransitive patterns with A, S of *stative* intransitive patterns with O.

Case (5/5)

- In **Austronesian** case systems, which argument of the transitive patterns like the only argument of the intransitive depends on the **voice** of the verb (here, for Tagalog, AV = agentive voice, OV = objective voice, DV = dative voice):

  1. Bumili ang=lalake ng=isda sa=tindahan.
     PERF.AV.buy DIR=man IND=fish OBL=store
     'The man bought fish at the store.'
  2. Binili ng=lalake ang=isda sa=tindahan
     PERF.buy.OV IND=man DIR=fish OBL=store
     'The man bought the fish at the store.'
  3. Binilhan ng=lalake ng=isda ang=tindahan.
     PERF.buy.DV IND=man IND=fish DIR=store
     'The man bought fish at the store.'

Beyond V, S, and 0

Languages also vary with respect to relative position of:

- pre/postposition and its object
- complementizer and clause
- verb and adverb
- clause and sentence modifier
- noun and attributive adjective
- noun and relative clause or complement clause
- noun and determiner
- noun and possessor
- noun and classfier (if any)

Word Order Freedom (1/4)

Many languages are often claimed not to belong to any of the six typed defined by relative position of S, V, and 0, e.g.

- Korean and Japanese, and subordinate clauses of German and Dutch, are often characterized as **V-final** rather than S-O-V.
- Tagalog and K'iche' are usually characterized as **V-initial**
- Some languages are often said to have **free** word order, such as (poetic) Latin, Romanian, Finnish, Serbo-Croation, and Warlpiri.

Word Order Freedom (2/4)

In reality, order in so-called free-word-order languages is subject to many different kinds of constraints involving a wide range of factors including:

- prosodic properties of the argument (e.g. ability to bear a pitch accent)
- whether the argument is a 'full NP', 'independent pronoun', or clitic
- semantic role of the argument
- inherent properties of the argument such as humanness or animacy
- person of the argument
- pragmatic properties of the argument, such as (in)definiteness, being a (continued, or contrastive) topic, or being a (answer, or corrective, or other) focus

Word Order Freedom (3/4)

- Even in languages where the order of 'major constituents' (arguments and modifiers) within a clause is relatively free (e.g. Finnish), the order of the words *within* each major constituent may be as rigid as in English.
- Much rarer is the "splitting" of NPs (e.g. separation of determiners or adjectives from nouns, as in these Jiwarli (central Australia) examples:

1. Kutharra-rru ngunaha ngurtnta-inha jiluru.
   two.nom-now that.nom lie-pres egg.nom
   'Now those two eggs are lying (there).'
2. Karla wantha-nma-rni jarnpa juma.
   fire.acc give-imper-hence light.acc small.acc
   'Give me a small fire light.'

Word Order Freedom (3/4)

- In some languages, certain nonfinite verbs allow (the appearance of) 'scrambling' of an argument or modifier into the next clause (or VP) up, as in German so-called 'coherent' constructions:

  1. dass ihm der Mann zu helfen versucht
     that him.DAT the man to help tries
     'that the man tries to help him'

- In many languages, e.g. Czech, clitics can scramble out of their 'home' VP (often to "second position"):

  1. Opravit *jsem se mu to* včera snažil marně.
     repair aux refl to-him it yesterday tried fruitlessly
     'I tried to repair it for him yesterday without success.'

- Scrambling out of a finite clause into a higher clause is much rarer.