Linguistic Diversity
Typologies and Constraints

Dan I. Slobin
slobin@berkeley.edu

Departments of Psychology and Linguistics
Institute of Brain and Cognitive Science
Institute of Human Development

Max Planck Institute for Psycholinguistics
Nijmegen, Netherlands

Learning in Retirement, “The Origins and Nature of Language,” UC Berkeley, October 23, 2018
How different are languages?

universal and particular
recurrent patterns: linguistic typology
implicational universals
categories: objects, events, grammatical forms
constraints of processing and acquisition
THE WORLD OF BERKELEY CHILD LANGUAGE RESEARCH
Slobin’s research sites

USA (English, ASL)
Netherlands (Dutch, Sign Language of the Netherlands)
Germany (German)
Russia (Russian)
Spain (Spanish)
Italy (Italian)
(former) Yugoslavia (Serbo-Croatian)
Turkey (Turkish)
Israel (Hebrew)
How are languages put together?

• duality of patterning:
  • meaningless elements (e.g., sounds, handshapes)
  • meaningful elements (e.g., words, signs)

_for example:_ -s

<table>
<thead>
<tr>
<th>PLURAL NOUN</th>
<th>his work-s are interesting</th>
</tr>
</thead>
<tbody>
<tr>
<td>SINGULAR VERB</td>
<td>he work-s</td>
</tr>
<tr>
<td>POSSESSIVE</td>
<td>the work-s theme is significant</td>
</tr>
</tbody>
</table>

_for example:_ ASL handshape

+ orientation
+ movement
+ location
a few of many examples of signs in ASL formed with the V-handshape

LOOK / SEE

STAND

https://www.lifeprint.com/

https://www.handspeak.com/word/search/index.php?id=2104

ELEVATOR

STUCK

https://nl.pinterest.com/pin/319896379766932006/?lp=true
How are languages put together?

• duality of patterning:
  • meaningless elements (e.g., sounds, handshapes)
  • meaningful elements (e.g., words, signs)

• order of elements
  • words (e.g., dog bit man – man bit dog)
  • morphemes (e.g., use-less-ness)

• location in space (signed languages only)
  • placement
  • direction

• combinations of elements
  • word + morpheme(s) (e.g., cat-s, play-ed)
  • word + word (e.g., screwdriver, bus driver)

• modification of elements
  • sounds (e.g., draw – drew)
  • prosody (stress, intonation) (e.g., She can’t come – She can’t come?)
  • tone (pitch, change of pitch – in a syllable)
Examples of tone (from Larry Hyman, Dept. of Linguistics)

http://linguistics.berkeley.edu/~hyman/

**Standard Mandarin Tones**

Tone I  (high level)  \( ma^{55} \)  ‘mother’
Tone II  (rising)  \( ma^{35} \)  ‘hemp’
Tone III  (falling-rising)  \( ma^{214} \)  ‘horse’
Tone IV  (falling)  \( ma^{51} \)  ‘scold’

5 = highest pitch, 1 = lowest
## Iau

**(Lakes Plain; Papúa, Indonesia)**

<table>
<thead>
<tr>
<th>Verbs</th>
<th>Inflectional meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bá ‘came’</td>
<td><strong>totality of action punctual</strong></td>
</tr>
<tr>
<td>bà ‘has come’</td>
<td><strong>resultative durative</strong></td>
</tr>
<tr>
<td>bă” ‘might come’</td>
<td><strong>totality of action incomplete</strong></td>
</tr>
<tr>
<td>bà ‘came to get’</td>
<td><strong>resultative punctual</strong></td>
</tr>
<tr>
<td>bà ‘came to end point’</td>
<td><strong>telic punctual</strong></td>
</tr>
<tr>
<td>bâ ‘still not at endpoint’</td>
<td><strong>telic incomplete</strong></td>
</tr>
<tr>
<td>bà ‘come (process)’</td>
<td><strong>totality of action durative</strong></td>
</tr>
<tr>
<td>bă ‘sticking, attached to’</td>
<td><strong>telic durative</strong></td>
</tr>
</tbody>
</table>
About 40% of the world’s languages are tonal.

WALS (World Atlas of Language Structures) [https://wals.info/](https://wals.info/)
a simple example of universals: WORD ORDER

- **SUBJECT – VERB – OBJECT**
  
  the chicken saw the rat  [English]

- **SUBJECT – OBJECT – VERB**
  
  tavuk fare-yi gör-dü  [Turkish]
  chicken rat-the see-past

- **VERB – OBJECT - SUBJECT**
  
  n-ahita ny voalavo ny akoho  [Malagasy]
  past-see the rat the chicken
dominant order of subject, object, and verb  (1,362 languages)

- SOV 565
- SVO 488
- VSO 120
- no dominant order 189
**PREPOSITIONS AND POSTPOSITIONS**

- **on** the table  
  [English - SVO]  
  **PREPOSITION**

- **masa – da**  
  table – on  
  [Turkish - SOV]  
  **POSTPOSITION**
IMPLICATIONAL UNIVERSALS
(sample of 984 languages)

<table>
<thead>
<tr>
<th></th>
<th>OBJECT BEFORE VERB (OV)</th>
<th>OBJECT AFTER VERB (VO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PREPOSITIONS</td>
<td>14</td>
<td>456</td>
</tr>
<tr>
<td>POSTPOSITIONS</td>
<td>472</td>
<td>42</td>
</tr>
</tbody>
</table>

If you speak a language in which the verb tends to come before the object, relational information will come before the noun – i.e. in the form of **prepositions**.

If you speak a language in which the object tends to come before the verb, relational information will come after the noun – i.e. in the form of **postpositions**.
habits of information processing: learning Turkish – a verb-final language

ENGLISH: Gather referential and relational information as you go.

TURKISH: Gather referential information and wait for relation between referents.

Al put...
Al put a book...
Al put a book on the table...
Al put a book on the table in his study.
Ease of processing: consistent habits across constructions

<table>
<thead>
<tr>
<th>ENGLISH</th>
<th>TURKISH</th>
</tr>
</thead>
<tbody>
<tr>
<td>preposed</td>
<td>postposed</td>
</tr>
<tr>
<td>don’t read</td>
<td>read don’t (oku-ma)</td>
</tr>
<tr>
<td>will read</td>
<td>read will (oku-yacak)</td>
</tr>
<tr>
<td>can read</td>
<td>read can (oku-yabilir)</td>
</tr>
<tr>
<td>will not read</td>
<td>read not will (oku-ma-yacak)</td>
</tr>
</tbody>
</table>
How languages are (apparently) *not* put together – i.e., constraints on universals \textit{(Chomsky)}

• change meaning by reversing order
  • She flew to Chicago yesterday.
  • QUESTION: *Yesterday Chicago to flew she.

• place a grammatical element by counting rather than by linguistic structure
  • POSSIBLE: negation after first auxiliary verb
    • I think they have *not* gone yet.
    • I think they should *not* have gone.
    • Yesterday he told me that they had *not* gone.
  • IMPOSSIBLE: negation after third element, whatever it is
    • *I think they *not* have gone yet.
    • *I think they *not* should have gone.
    • *Yesterday he told *not* me that they had gone.
typological consistency

Interim conclusions:

• Languages select from a universal collection of building blocks and patterns.

• Languages differ considerably in their selection, but they fall into a small set of types.

• Habits of using a language function to maintain consistency across constructions with that language.

• (At the same time, languages are constantly changing. Next week Eve Sweetser will tell us something about this.)
BUT:

So far we’ve followed standard assumptions about the nature of language.

Crosslinguistic studies of categorization in different types of languages challenge assumptions about the **semantic** bases of words and grammatical elements.
some common assumptions

• WORDS
  • There are words that make reference: **NOUNS**
  • There are words that indicate relations between referents: **VERBS**

• MEANING
  • Nouns refer to **SALIENT ENTITIES** in the world (things that stand out in perception, can be manipulated, etc.)
  • Verbs refer to **SALIENT EVENT TYPES**.

*What does it look like from a child’s point of view?*
How can the child make sense of language?

Turkish mother:
yemekbirşeyistiyormusun?

English-speaking mother:
doyouwantsomethingtoeat?

What is “eat”?

What is in the bowl?
What is “eating”? 

The dog is **eating** a steak. 
Der Hund **fisst** ein Steak. 

The child is **eating** a banana. 
Das Kind **isst** eine Banane.
Degree of *granularity* of coding

- English: *eat*

- German:
  - *essen* [human]
  - *fressen* [animal]
Degree of *granularity* of coding

- **English**: *eat*

- **German**:
  - *essen* [human]
  - *fressen* [animal]

- **Tzeltal Mayan**
  - *bik* [things that are swallowed whole]
  - *k’ux* [crunchy solids, beans]
  - *lo’* [soft solids, fruits]
  - *ti’* [meat]
  - *we’* [tortillas, bread]
  - *...and more...*

different bases of categorization

- eat
  - eater
    - human
    - animal
  - eaten
    - swallowed
    - crunchy
    - soft
    - meat
    - bread
    ...
typological consistency:
granularity of coding across domains in Tzeltal

• verbs of carrying / holding
  *pet* in both arms
  *kuch* weight on head/back
  *k’ech* weight across shoulders
  *lik* in hand, supported from top
  *tuch’* vertically extending from hand
  *lut* in mouth
  ...

• verbs of motion
  *tal* toward self
  *ba* away from self, or neutral direction
  *och* inwards
  *lok’* outwards
  *jelaw* crossways
  *jajch* getting up from sitting/lying position
  ...

incompatible bases of categorization across languages


Suppose a child hears a verb used in the context of ‘Someone carrying something’

(Melissa Bowerman 2002)
some situations in **conceptual** space
ENGLISH SEMANTIC SPACE

CARRY
TZELTAL MAYAN SEMANTIC SPACE

- IN ARMS
- ON BACK
- ON SHOULDER
- IN HAND, SUPPORT FROM ABOVE
- ON HEAD
NAVAHO SEMANTIC SPACE

LIVING

CONTAINER WITH CONTENTS

LONG

BULKY
verbs of placement in English and Korean

(Melissa Bowerman & Soonja Choi)

a conclusion with regard to categorization and acquisition

- The appropriate level of granularity varies with conceptual domain and with language.
- Child speech is attuned to the semantic analysis of the input language.
- Children form categories at the appropriate level of granularity.
- They attend to the relevant words as guides (verbs, prepositions, postpositions, word-endings, etc.).
- That is: processes of learning serve to maintain typological consistency.
Another challenge:
What is a thing?
In some languages nouns don’t refer to things, but to more general substances or materials. In Yucatec Maya there is no word for “banana,” but rather a general word for “banana stuff” (há’as) that refers to entity types by addition of a grammatical element called a **CLASSIFIER**.

Two-dimensional banana-stuff = banana leaf

Three-dimensional banana-stuff = a banana

Planted banana-stuff = banana tree

One-load banana-stuff = a bunch of bananas

back to the typology of prepositions/postpositions and word order

• There are languages that don’t have separate prepositions or postpositions, but rather have suffixes on nouns (cases) that perform those functions.

• Many languages (not English!) allow for considerable variability of word order.

• In many languages verbs with many parts carry the meanings of sentences, without prepositions/postpositions, nouns, or pronouns.
locative case endings (noun suffixes indicating locations)

• Turkish
  • NOUN-*da ‘located at’
  • NOUN-*dan ‘moving toward’

• Hindi
  • NOUN-*par ‘supported by surface
  • NOUN-*ko ‘contained’

• Finnish
  • NOUN-*lla ‘located at or on’
  • NOUN-*lle ‘moving onto’
  • NOUN-*ssa ‘located in’
  • NOUN-*an ‘moving into’
coding of locative relations by case inflection on noun


CONTAINMENT

STATIC

DYNAMIC

SUPPORT
CONTAINMENT

STATIC

DYNAMIC

Inessive

Hindi

SUPPORT

Adessive
CONTAINMENT

STATIC
- Inessive

DYNAMIC
- Illative

SUPPORT

- Adessive
- Allative

Finnish
another interim conclusion

• Languages differ in many ways in the categories they encode and how they encode them.

• But they all draw on a very large, common stock of the ways in which human beings are able to analyze their perceptions, experience, thoughts, interactions.
  • containment, support...
  • located at, moving toward, moving away...
  • manners of action (carrying, eating...)
  • and many, many more
back to the typology of prepositions/postpositions and word order

• There are languages that don’t have separate prepositions or postpositions, but rather have suffixes on nouns (cases) that perform those functions.

• Many languages (not English!) allow for considerable variability of word order.

• In many languages verbs with many parts carry the meanings of sentences, without prepositions/postpositions, nouns, or pronouns.
early child speech in a highly inflected language (Inuktitut)

A 3-year-old girl saw her friend bring a puppy to the porch of the house. The girl is in the house and can hear the puppy. She wants to take it in.

This is all done in one verb:

\[ \text{itiq- guma- jara} \]

\begin{align*}
\text{ENTER- WANT-} & \quad \text{FIRST-PERSON-SINGULAR-} \\
& \quad \text{AFFECTS-THIRD-PERSON-SINGULAR} \\
\end{align*}

‘I want to take him/her/it in.’
Atsugewi – a Hokan language once spoken near Mt. Lassen (last speaker died in 1988, studied in depth by UCB linguist Leonard Talmy)


Verb roots refer to categories of objects. For example (part of a very long list):

- **-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’
- **-stäq-** ‘for runny icky material (e.g. mud, manure, rotten tomatoes, guts, chewed gum) to move/be-located’
R1: -swal  
'for limp (not stiff/resilient) material to move/be-located'

PC: ca- (19)  
'from the wind blowing on the FIGURE'

DG: -mič  
'down onto the surface of the ground'

Ax: '- w- -a  
'3, (3), factual' --causative

\[- w- ca- swal -mič -a/ \Rightarrow [\text{̃weswaliːmɨכם}]

Literally: 'limp-material moved down to the surface of the ground from the wind blowing on it'

Casually: 'the clothes blew down from the clothesline'

(b)

BC: ra- (128a)  
'from a linear object moving laterally [with one end] along a surface with the FIGURE'

DG: -im  
'thither'

Ax: '- -a:  
'you, (3), imperative' --effective

\[- ra- swal -im -a:/ \Rightarrow [\text{̃aswaliːmɨ-}]

Literally: 'you effect limp-material to move thither by moving a linear object laterally [with one end] along a surface with it!'

Casually: 'push that dead snake away with this stick!'
Some examples of “exotic” meanings expressed by grammatical markers on verbs

Native America languages

• -cis ‘into a fire’ (Atsugewi)
• ca- ‘from the wind blowing on it’ (Atsugewi)
• siwi- ‘swimming-on-surface’ (Nez Perce)
• tukwe- ‘swimming-within-liquid’ (Nez Perce)
• qisim- ‘in anger’ (Nez Perce)
• čh- ‘by gambling’ (Pomo)
• ersi + BODY PART ‘feel cold in one’s BODY PART’ (Eskimo)
English

PRESENT PERFECT ‘still alive’

Barack Obama has lived in the White House.
*Abraham Lincoln has lived in the White House.

Donald Trump (February 2017)
“Frederick Douglass is an example of somebody who’s done an amazing job and is getting recognized more and more, I notice.”
another typological distinction: head-marking and dependent-marking (Johanna Nichols, UCB)

• The verb is the head of the clause and the nouns and pronouns are dependents.
• In languages like English, relations between a verb and its dependents are marked on the dependents – e.g., We see them versus They see us.
• The verb doesn’t indicate who did what to whom.
• English is a dependent-marking language.
• In Atsugewi and Inuktitut all of the information about who did what to whom, how, and why is marked on the verb.
• Atsugewi and Inuktitut are head-marking languages.
another example: English versus Yucatec Mayan
English (dependent-marking)
Yucatec Mayan (head-marking)

he sees me

k -uy -il -ik -3SG see -ONGOING -1SG

HABITUAL -3SG ACTOR

- en -1SG UNDERGOER
Both types of language are widespread and must be considered when generalizing about universals.

<table>
<thead>
<tr>
<th>Dependent-marking</th>
<th>Head-marking</th>
</tr>
</thead>
<tbody>
<tr>
<td>• English</td>
<td>• ASL and probably all sign languages</td>
</tr>
<tr>
<td>• German, Dutch</td>
<td>• Nootka (Arctic)</td>
</tr>
<tr>
<td>• French, Spanish</td>
<td>• Navajo (North America)</td>
</tr>
<tr>
<td>• Hebrew, Arabic</td>
<td>• Mayan (Central America)</td>
</tr>
<tr>
<td>• Turkish</td>
<td>• Pirahã (South America)</td>
</tr>
<tr>
<td>• Japanese</td>
<td>• Abkhaz (Caucasus)</td>
</tr>
<tr>
<td>• Chinese</td>
<td>• Asmat (New Guinea)</td>
</tr>
<tr>
<td>• ...</td>
<td>• Konua (Oceania)</td>
</tr>
<tr>
<td>• ...</td>
<td>• ...</td>
</tr>
</tbody>
</table>
“The head-marking type is well attested in the Americas and Melanesia but vanishingly rare elsewhere; the dependent-marking type is strongest in Africa, Eurasia, and perhaps Australia-New Guinea and infrequent (though not rare) elsewhere.”

a caveat

All claims about linguistic universals and the nature of human language must take account of linguistic typology – and there is a growing science of typology.

• head-marking vs. dependent-marking
• analytic, synthetic, agglutinative morphology
• word order pattern
• subject-focus vs. topic-focus
• verb-framed vs. satellite-framed
• ...and many more...
What can we conclude about the nature of human language in the light of limited evidence?

• Archaeology: Language leaves no fossils or artifacts.
• What are the limits of human language?
  • At least 500,000 languages have ever been spoken.
  • About 7,000 languages are still spoken.
  • So what we have now is a non-random sample of less than 2% of the full range of human linguistic diversity.
• And we have good information for about 10% of existing languages (i.e., 10% of 2%!)
“The myth of language universals”
(Nicholas Evans & Stephen Levinson, 2009)

“At this stage of linguistic inquiry, almost every new language that comes under the microscope reveals unanticipated new features.”

“Languages differ so fundamentally from one another at every level of description (sound, grammar, lexicon, meaning) that it is very hard to find any single structural property they share.”

in conclusion: paraphrasing Johannes Kepler

“The diversity of the phenomena of nature is so great and the treasures hidden in the heavens human languages so rich precisely in order that the human mind shall never be lacking in fresh nourishment.”

Mysterium Cosmographicum - 1596
some basic reading

• Edward Sapir (1921). *Language: An introduction to the study of speech.* When I read this little book as a student it fascinated me and fired up my imagination. I think it’s still a wonderfully lucid and generally accurate overview of how languages work and change and how language relates to thought and culture – all written with a stylistic grace that is hard to find today.

• John McWhorter (2001). *The power of Babel: A natural history of language.* This is an accessible and readable overview of language history and change, with insightful discussion of how dialects become languages plus serious issues of ongoing language extinction.

• Daniel L. Everett (2012). *Language: The cultural tool.* This controversial research argues strongly for the social and cultural bases of human language. Everett is famous for his widely-discussed claims about the language of the Amazonian Pirahã which challenge Chomskyan assumptions about the nature of language.

• Nicholas Evans (2010). *Dying words: Endangered languages and what they have to tell us.* An insightful and beautifully written book. The subtitle tells it all.
to learn more  (all papers available from slobin@berkeley.edu )


