





# OVERVIEW

## INTRODUCTION

The Basic Order of Subject, Object, and Verb

## DATA AVAILABLE

Inventory-Based Accounts of Word Order

Distribution-Based Accounts of Word Order

## RECENT RESEARCH

Information-Theoretic Considerations

Word Order in “Proto-World?”

Basic Order of Agent, Patient, Action

Word Order Case-Marking Trade-Off

Phylogenetics



# THE BASIC ORDER OF SUBJECT, OBJECT AND VERB

In a prototypical transitive sentence the

## ► **Subject**

shows *agreement* with the predicate,  
directly *precedes* the predicate,  
is the *topic* of the sentence,  
and represents the *semantic role of agent*

## ► **Object**

does not agree with the predicate,  
represents the *patient*

## ► **Verb** represents the *action*

e.g. [s The dog] [v chased] [o the cat]

Note that there are counterexamples for any of these criteria: e.g. in passive sentences the subject might be the patient of an action.

[s The dog] [v was chased] [o by the cat]



# WORD ORDER AT DIFFERENT LEVELS OF ABSTRACTION

Syntactic role level [S][V][O]  
[<sub>S</sub> colorless green ideas] [<sub>V</sub> escape] [<sub>O</sub> meaning]

phrase level [NP][VP][NP]  
[<sub>NP</sub> colorless green ideas] [<sub>VP</sub> escape] [<sub>NP</sub> meaning]

word level [colorless] [green] [ideas] [escape] [meaning]

“colorless green ideas escape meaning”

Actually: typologists talk about “**syntactic role order**” or “**phrase order**” rather than “**word order**”



# GREENBERG'S WORD ORDER UNIVERSALS

Greenberg (1963) Gave a set of statistical universals (i.e. trends rather than absolute statements) of word order. He used a relatively small sample of geographically and genetically diverse languages

Problems with this approach:

- ▶ it is somewhat **hard to define “basic word order”**  
(most notorious case: German, with different word order in different clause types)
- ▶ there arguably are **languages without fixed word order**  
(Slavic languages, Latin, ...), for which we can only talk about preferred word order

[Greenberg, Joseph (1963). Some universals of grammar with particular reference to the order of meaningful elements.]



# THE SIX POSSIBLE ORDERS

## SOV

- (1) Ainu (Isolate: Japan)

kamuy aynu rayke  
bear person kill

“The bear killed the person”

## SVO

- (2) Matuumbi (Niger-Congo)

abunwaási aachéngite ñụúmba  
PN he.built house

“Abumwas built a house”

[Velupillai (2012) *An Introduction to Linguistic Typology*. John Benjamins, Amsterdam, p. 285]



# THE SIX POSSIBLE ORDERS

## VSO

- (3) Irish (Indo-European)

tógann Máire an cat  
lift.PRES PN ART cat

“Mary lifts the cat”

## VOS

- (4) Cèmuhî (Austronesian)

ε ālī-hĩ ā-li mwà ɔ pā-li āpūlīp  
3SG see-TR ART:NEUT-DEF house SUBJ ART:NF-DEF man

“The man saw the house”

[Velupillai (2012) *An Introduction to Linguistic Typology*. John Benjamins, Amsterdam, p. 285]



# THE SIX POSSIBLE ORDERS

## OVS

- (5) Pāri (Nilo-Saharan)

dháagò á-yàan                  ùbúrr-ì  
 woman COMPL-insult PN-ERG

“Ubur insulted the woman”

## OSV

- (6) Warao (Isolate: Venezuela)

erike hube abun-ae  
 PN snake bite-PAST

“A snake bit Enrique”

[Velupillai (2012) *An Introduction to Linguistic Typology*. John Benjamins, Amsterdam, p. 285]



# FLEXIBLE WORD ORDER

## Nhanda (Pama-Nyungan)

abarla-lu	wumba-yi	wur'a-tha
child-ERG	steal-PPERF	money-1SGOBL
S	V	O
abarla-lu	wur'a-tha	wumba-yi
S	O	V
wumba-yi	wur'a-tha	abarla-lu
V	O	S
wumba-yi	abarla-lu	wur'a-tha
V	S	O
wur'a-tha	wumba-yi	abarla-lu
O	V	S
wur'a-tha	abarla-lu	wumba-yi
O	S	V

“The child stole my money”

[Velupillai (2012) *An Introduction to Linguistic Typology*. John Benjamins, Amsterdam, p. 282]



# TWO BASIC ACCOUNTS OF WORD ORDER

## Inventory-based

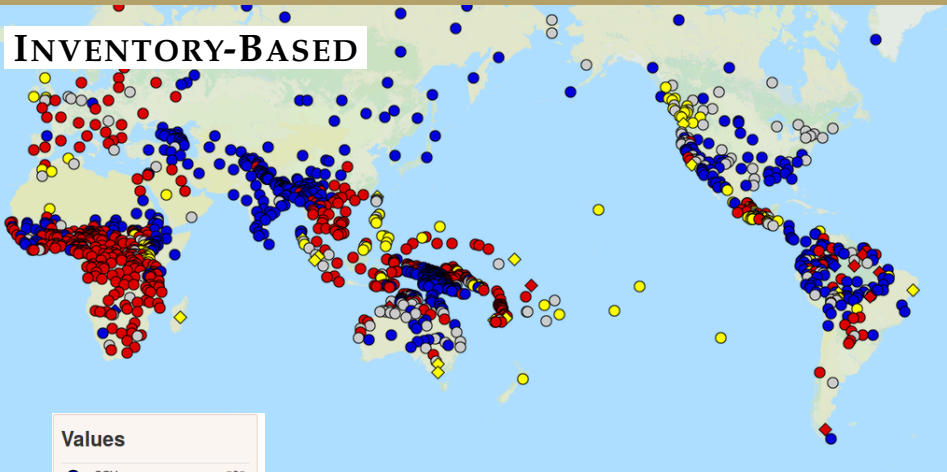
- ▶ Assign one of 7 categories (SOV, SVO, VSO, VOS, OSV, OVS, non dominant) based on a grammatical description of the language or a small “corpus” of collected sentences

## Distribution-based

- ▶ Give the frequencies of sentences with the 7 word order options in a corpus of the language



## INVENTORY-BASED



## Values

SOV	565
SVO	488
VSO	95
VOS	25
OVS	11
OSV	4
No dominant order	189

# WALS Chapter 81

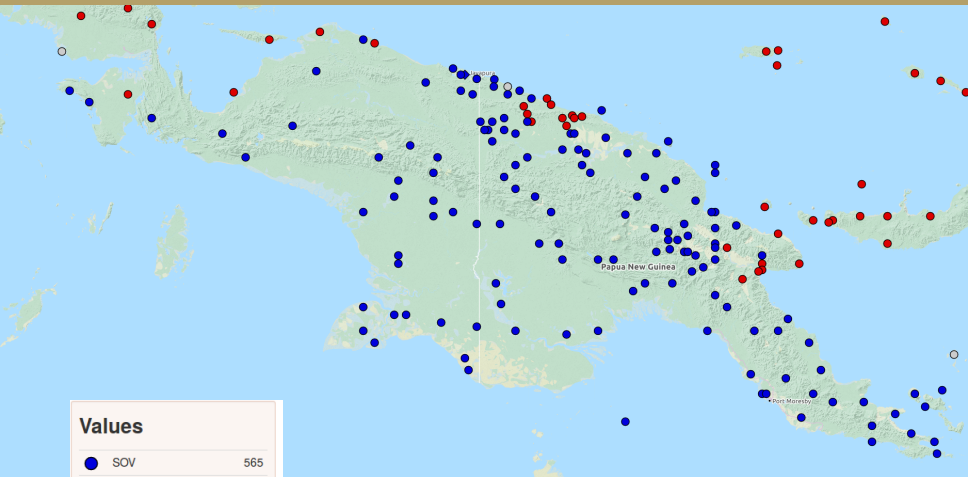
# Basic order of Subject, Object and Verb for 1377 languages



INTRODUCTION  
○○○○○○○

DATA AVAILABLE  
○○●○○○○○○○

RECENT RESEARCH  
○○○○○○○○○○○○○○○○○○



### Values

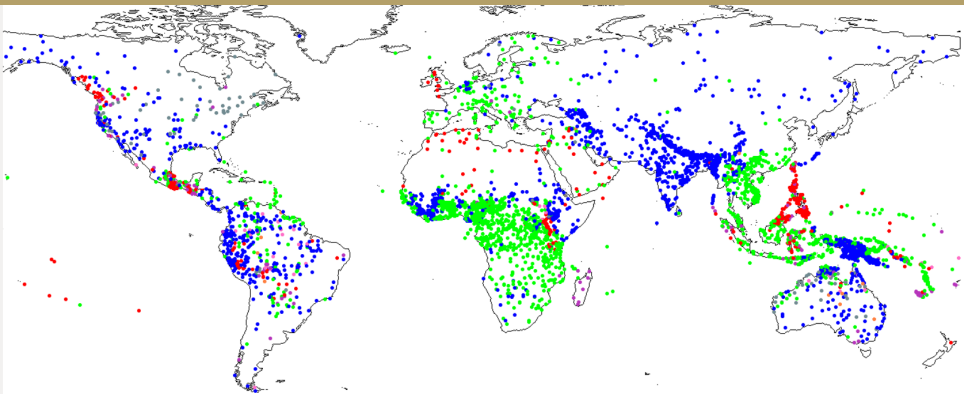
●	SOV	565
●	SVO	488
●	VSO	95
◆	VOS	25
◆	OVS	11
◆	OSV	4
○	No dominant order	189

## WALS Chapter 81 Papua New Guinea









SOV

SVO

VSO

VOS

OVS

OSV

## No Dominant Order

## Hammarstöm (2013, 2016)

Basic order of Subject, Object and Verb for more than 5000 languages (not yet available online)



# DISTRIBUTION-BASED

Account by Östling (2015) based on 987 parallel texts of the Bible Corpus. Collects the frequencies of word orders occurring in the texts.

## Examples

(7) German

Im Anfang schuf Gott Himmel und Erde  
in.the beginning [create.PERF]V [god]S [heaven and earth]O

“In the beginning god created the heavens and the earth”

Finsternis lag über der Tiefe  
[darkness]S [lay.PERF]V over [the deep]O

“Darkness was on the face of the deep”

[Östling, Robert (2015). Word order typology through multilingual word alignment. *Proceedings of ACL, Beijing.*]



# DISTRIBUTION-BASED

ISO	SOV	SVO	OSV	OVS	VSO	VOS	Total	SOV_perc	MAX	WALS
<u>aai</u>	64	8	12	6	1	2	93	0.69	SOV	—
<u>aak</u>	81	17	20	33	5	9	165	0.49	SOV	—
<u>aa</u>	37	1	7	0	1	0	46	0.80	SOV	SOV
<u>abt</u>	29	7	5	8	1	1	51	0.57	SOV	SOV
<u>aby</u>	53	11	10	11	6	5	96	0.55	SOV	—
<u>acc</u>	4	27	5	1	2	2	41	0.10	SVO	—
<u>acd</u>	5	30	3	1	1	2	42	0.12	SVO	—
<u>ace</u>	3	70	10	6	6	6	101	0.03	SVO	No dominant order
<u>acf</u>	9	89	13	1	0	1	113	0.08	SVO	—
<u>acn</u>	103	2	26	12	2	1	146	0.71	SOV	—
<u>acr</u>	4	22	3	2	2	2	35	0.11	SVO	—
<u>acu</u>	21	5	3	0	0	2	31	0.68	SOV	—
<u>ade</u>	10	121	9	2	9	10	161	0.06	SVO	—
<u>adh</u>	16	94	4	1	5	10	130	0.12	SVO	—
<u>adj</u>	8	39	10	2	6	5	70	0.11	SVO	SVO
<u>adz</u>	48	15	11	9	7	5	95	0.51	SOV	SVO
<u>ae</u>	72	9	5	5	5	1	97	0.74	SOV	SOV
<u>afr</u>	523	142	52	13	44	2	776	0.67	SOV	—
<u>agd</u>	69	18	18	27	13	6	151	0.46	SOV	SOV
<u>agg</u>	39	9	8	3	3	0	62	0.63	SOV	SOV

[Östling, Robert (2015). Word order typology through multilingual word alignment. *Proceedings of ACL, Beijing.*]



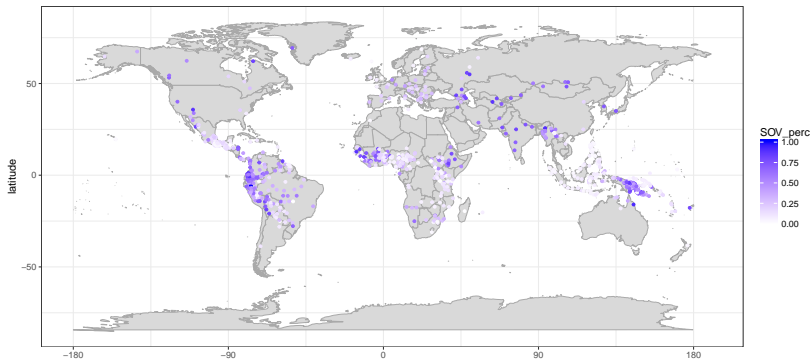
# DISTRIBUTION-BASED

ISO	Name	SOV	SVO	OSV	OVS	VSO	VOS	Total	SOV_perc	SVO_perc	MAX	WALS
hin	Hindi	194	14	18	14	2	1	243	0.8	0.06	SOV	SOV
eng	English	47	4576	291	23	74	14	5025	0.01	0.91	SVO	SVO
haw	Hawaiian	3	31	2	2	70	3	111	0.03	0.28	VSO	VSO
mav	Sateré-Mawé (Tupian)	3	9	4	7	8	11	42	0.07	0.21	VOS	—
czt	Zotung Chin (Sino-Tibetan)	34	8	35	25	4	9	115	0.3	0.07	OSV	—
cnw	Ngawn Chin (Sino-Tibetan)	29	6	14	33	4	4	90	0.32	0.07	OVS	—
deu	German	1571	1391	251	53	529	105	3900	0.4	0.36	SOV	No dominant order

[Östling, Robert (2015). Word order typology through multilingual word alignment. *Proceedings of ACL, Beijing.*]



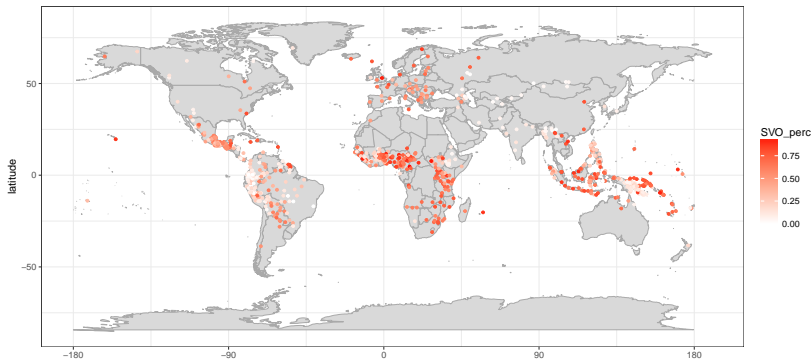
# DISTRIBUTION-BASED



[Östling, Robert (2015). Word order typology through multilingual word alignment. *Proceedings of ACL, Beijing.*]



# DISTRIBUTION-BASED



[Östling, Robert (2015). Word order typology through multilingual word alignment. *Proceedings of ACL*, Beijing.]

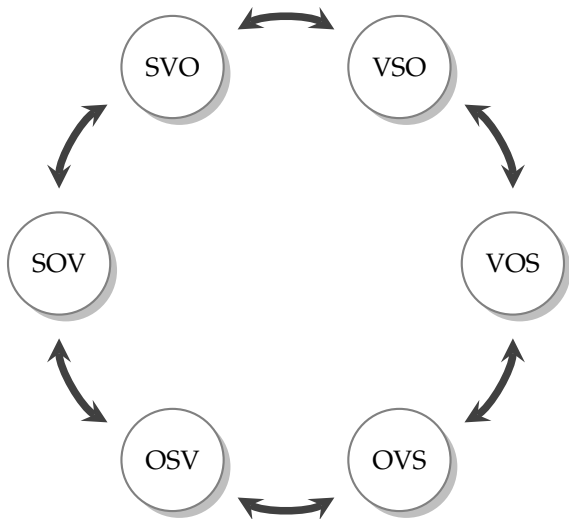


# RECENT RESEARCH

1. Information-Theoretic Considerations
2. Word Order in “Proto-World?”
3. Basic Order of Agent, Patient, Action
4. Word Order and Case-Marking Trade-Off
5. Phylogenetics



## WORD ORDER CHANGE AND EVOLUTION: THE PERMUTATION RING



[Ferrer-i-Cancho (2017). The placement of the head that maximizes predictability. An information-theoretic approach.]



## TWO COMPETING PRESSURES THAT SHAPE WORD ORDER

### Dependency-length minimization

The head of a sentence/phrase (e.g. the verb) should be placed in a way that **minimizes** dependency lengths

### Predictability maximization

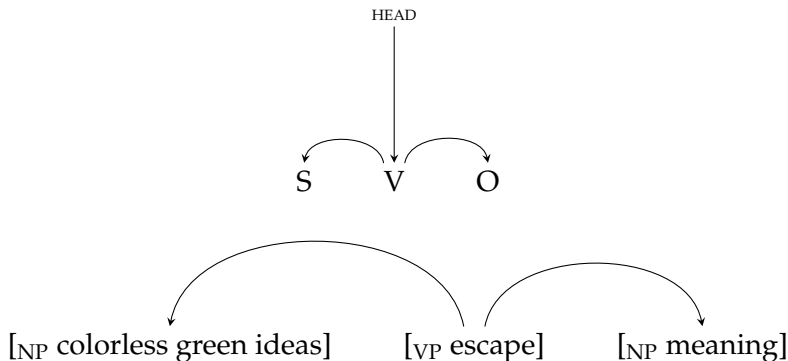
The head of a sentence/phrase should be placed in a way that **maximizes** its predictability

[Ferrer-i-Cancho (2017). The placement of the head that maximizes predictability. An information-theoretic approach.]



# DEPENDENCY-LENGTH MINIMIZATION

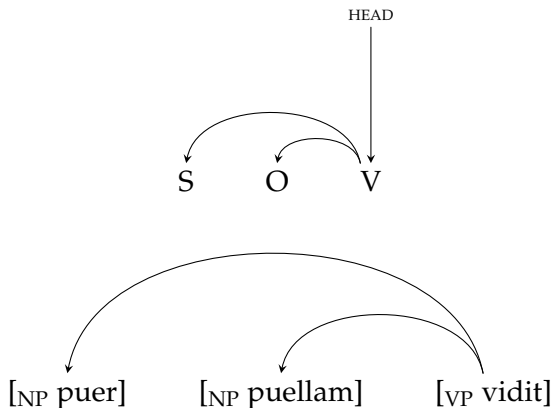
Placing the verb (head) in the **middle position** minimizes dependency lengths.





# PREDICTABILITY MAXIMIZATION

Placing the verb (head) in the **last position** maximizes its predictability





## PNAS

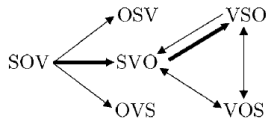
Murray Gell-Mann<sup>a,1</sup> and Merritt Ruhlén<sup>b,1</sup>

<sup>a</sup>Santa Fe Institute, Santa Fe, NM 87501; and <sup>b</sup>Department of Anthropology, Stanford University, Stanford, CA 94305

Contributed by Murray Gell-Mann, August 26, 2011 (sent for review August 19, 2011)

Recent work in comparative linguistics suggests that all, or almost all, attested human languages may derive from a single earlier language. If that is so, then this language—like nearly all extant languages—most likely had a basic ordering of the subject (S), verb (V), and object (O) in a declarative sentence of the type “the man (S) killed (V) the bear (O).” When one compares the

man") and uses prepositions. (Nowadays, these correlations are described in terms of head-first and head-last constructions.) In light of such correlations it is often possible to discern relic traits, such as GN order in a language that has already changed its basic word order from SOV to SVO. Later work (7) has shown that diachronic pathways of grammaticalization often reveal relic



**Fig. 1.** Evolution of word order.

[Gell-Mann, Murray & Merritt Ruhlen (2011). The origin and evolution of word order. *PNAS*.]



**Table 2. Distribution of word order types in the world's languages**

World	1008-770-164-[40-16-13]
Khoisan	22-11-1-[0-0-0]
Congo-Saharan	61-318-16-[0-1-0]
Niger-Kordofanian	39-279-1-[0-0-0]
Nilo-Saharan	22-39-15-[0-1-0]
Indo-Pacific	223-25-1-[0-0-2]
Australian	59-20-1-[3-1-1]
Austriac	30-220-67-[16-0-2]
Austroasiatic	8-34-0-[1-0-0]
Miao-Yao	0-4-0-[0-0-0]
Daic	1-19-0-[0-0-0]
Austronesian	21-163-67-[15-0-2]
Dene-Caucasian	157-13-0-[0-0-0]
Basque	1-0-0-[0-0-0]
Caucasian	29-0-0-[0-0-0]
Burushaski	1-0-0-[0-0-0]
Sino-Tibetan	84-13-0-[0-0-0]
Ket	1-0-0-[0-0-0]
Na-Dene	41-0-0-[0-0-0]
Nostratic-Amerind	456-163-78-[21-14-8]
Afro-Asiatic	58-37-14-[0-0-0]
Nostratic	182-59-6-[0-0-0]
Kartvelian	4-0-0-[0-0-0]
Dravidian	28-0-0-[0-0-0]
Eurasian	149-59-6-[0-0-0]
Indo-Hittite	79-47-6-[0-0-0]
Uralic	10-10-0-[0-0-0]
Altaic	50-1-0-[0-0-0]
Ainu	1-0-0-[0-0-0]
Gilyak	1-0-0-[0-0-0]
Chukchi-Kamchatkan	2-1-0-[0-0-0]
Eskimo-Aleut	6-0-0-[0-0-0]
Amerind	216-67-58-[21-14-8]

The numbers after each family represent the number of languages with SOV, SVO, VSO, VOS, OVS, and OSV orders, given in that order, with the final three word orders in brackets. Note that we have chosen one of the several definitions of Nostratic.

## Word order in “Proto-World”

[Gell-Mann, Murray & Merritt Ruhlen (2011). The origin and evolution of word order. *PNAS*.]



# BASIC ORDER OF EVENTS



## The natural order of events: How speakers of different languages represent events nonverbally

Susan Goldin-Meadow<sup>\*,†</sup>, Wing Chee So<sup>‡</sup>, Aslı Özyürek<sup>§||</sup>, and Carolyn Mylander<sup>\*</sup>

<sup>\*</sup>Department of Psychology, University of Chicago, 5730 South Woodlawn Avenue, Chicago, IL 60637; <sup>†</sup>Department of Psychology, National University of Singapore, Republic of Singapore 119077; <sup>‡</sup>Department of Linguistics, Radboud University Nijmegen, 6525 HT Nijmegen, The Netherlands; <sup>§</sup>Max Planck Institute for Psycholinguistics, 6500 AH Nijmegen, The Netherlands; and <sup>||</sup>Department of Psychology, Koc University, 34450 Istanbul, Turkey

Edited by Rochel Gelman, Rutgers, The State University of New Jersey, Piscataway, NJ, and approved May 8, 2008 (received for review November 12, 2007)

To test whether the language we speak influences our behavior **Gesture Task (6).** Forty adults [10 English speakers, 10 Turkish



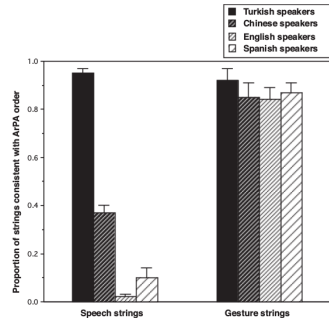


Fig. 2. Proportion of speech (Left) and gesture (Right) strings produced by speakers of Turkish, Chinese, English, and Spanish to describe transitive actions that were consistent with the ArPA order. Included are both in-place and crossing-space transitive actions.

[Goldin-Meadow et al. (2008) The natural order of events: How speakers of different languages represent events nonverbally. *PNAS*.]



# SEMANTICS AND WORD ORDER: EXTENSIONAL AND INTENSIONAL EVENTS

extensional (above), intensional (below)



Fig. 1. Example item: intensional event 'Pirate throws guitar'.



Fig. 2. Example item: extensional event, 'Cook thinks of sock'.

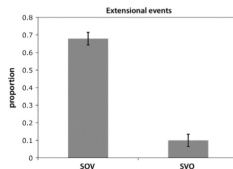


Fig. 3. Results: extensional events. Mean proportions of SOV and SVO gesturing orders for extensional events. Error bars indicate standard error of the mean.

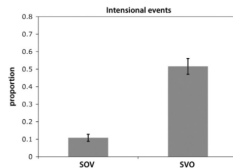


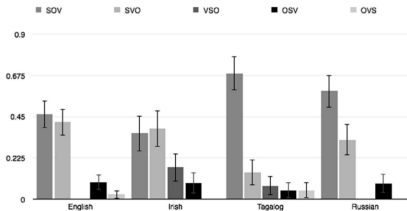
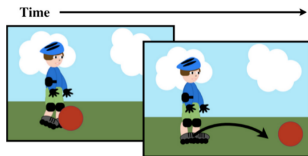
Fig. 4. Results: intensional events. Mean proportions of SOV and SVO gesturing orders for intensional events. Error bars indicate standard error of the mean.

[Schouwstra, Marieke, and De Swart, Henriette (2014) The semantic origins of word order. *Cognition* 131.]

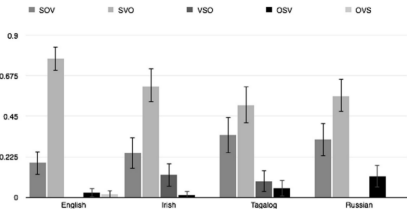


## SEMANTICS AND WORD ORDER: REVERSIBLE AND NONREVERSIBLE EVENTS

Nonreversible event



**Fig. 2.** Proportions of responses in basic three-word orders for nonreversible events (inanimate objects).

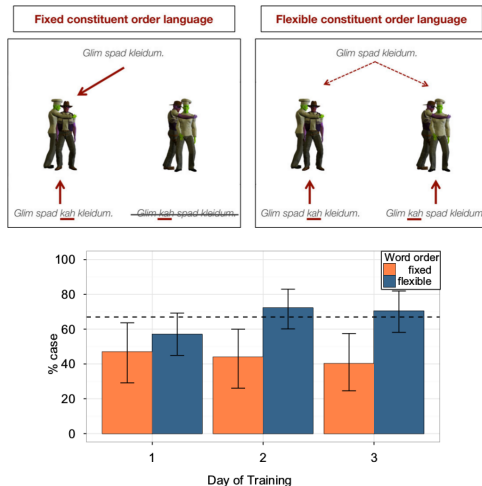


**Fig. 3.** Proportions of responses in basic three-word orders for reversible events (human objects).

[Futrell et al. (2015) Cross-linguistic gestures reflect typological universals: A subject-initial, verb-final bias in speakers of diverse languages. *Cognition* 136.]



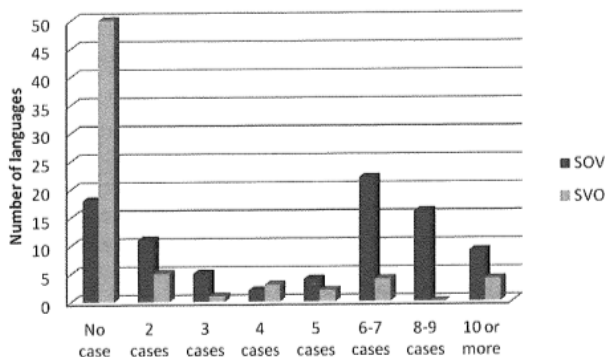
# WORD ORDER AND CASE-MARKING: EXPERIMENTAL EVIDENCE



[Fedzechkina et al. (2016) Balancing effort and information transmission during language acquisition: Evidence from word order and case marking. *Cognitive Science*.]



# WORD ORDER AND CASE-MARKING: THE CROSS-LINGUISTIC EVIDENCE



[Bentz & Christiansen (2013). Linguistic Adaptation: The trade-off between case marking and fixed word orders in Germanic and Romance languages . In: *Eastward flows the great river*. Festschrift in honor of Prof. William S-Y. Wang on his 80th birthday, ed. Feng Shi and Gang Peng. City University of Hong Kong Press, 48-56. ]



## WORD ORDER AND CASE MARKING: STATISTICAL ANALYSIS

## Typological Analysis

- ▶ balanced sample of 50 languages
- ▶ data from WALS

	M-	M+
WO-	3	22
WO+	8	17
$\chi^2 = 2.91, p = 0.088$		

WO-: word order *does not* disambiguate S and O

WO+: word order *does* disambiguate S and O

M-: no morphological marking

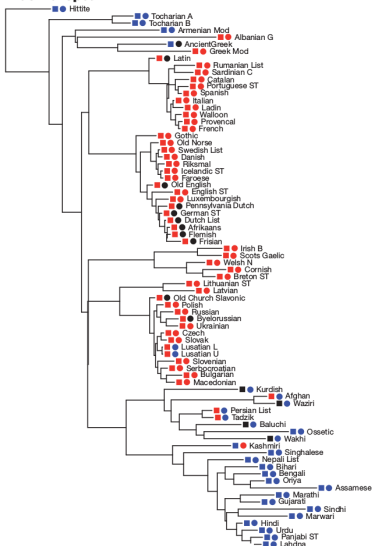
M+: morphological marking

[Sinnemäki (2008). Complexity trade-offs in core argument marking. In: Miestamo, Sinnemäki and Karlsson (ed.) Language complexity: Typology, contact, change. pp. 67-89.]



## WORD ORDER CHANGE AND EVOLUTION: PHYLOGENETICS

## Indo-European



● : VO order

● : OV order

● : no dominant order

■ : AdN order (prepositions)

■: NAd order (postpositions)

■ : no dominant order

[Dunn et al. (2011). Evolved structure of languages shows lineage-specific trends in word-order universals. *Nature*.]



THANKS. SEE YOU NEXT WEEK!

