

Language Diversity and Typology: LIN-1050

Word Order

Marina Pantcheva

September 25, 2007

1 Basic Word Order

1.1 How do we determine the basic word order

Basic word order is to be found in *stylistically neutral, independent indicative clauses with full noun phrase participants, where the subject is definite and human, the object is a semantic patient and the verb represents an action.*

Other criteria are *formal markedness* and *textual frequency*.

Assigning languages to some basic word order can be difficult and for some free word order languages the concept of basic word order is irrelevant.

- Splits (German)
- Free word order (Dyrbal)

The concept of basic word order is irrelevant to languages with a flexible word order. (Comrie 1989).

1.2 S, V and O

3 major constituents (**S**ubject, **O**bject and **V**erb) → 6 logical permutations.

- | | | |
|-----|---|---------------|
| (1) | John ate the cake | SVO (English) |
| (2) | Mæriæm ketab mixune
<i>Mary book reads</i>
'Mary reads books.' | SOV (Persian) |
| (3) | Lladdodd y ddraig y dyn
<i>killed the dragon the man</i>
'The dragon killed the man.' | VSO (Welsh) |

- (4) Sũy yi qawùh
howlermonkey people eat
'People eat howler monkeys.' OSV (Nadëb)
- (5) Nahita ny mpianatra ny vehivavy
saw the student the woman
'The woman saw the student' VOS (Malagasy)
- (6) Toto yahosie kamara
man it-grabbed-him jaguar
'The jaguar grabbed the man.' OVS (Hixkaryana)

1.3 Basic word order at non-clausal level

1.3.1 Adpositions

1. Preposition and Noun

- (7) behind the house PrN (English)

2. Noun and Postposition

- (8) isu-ni
chair-on
'on the chair' NPo (Japanese)

1.3.2 Adjective and Noun

1. Adjective and Noun

- (9) big house AN (English)

However,

- (10) people fluent in three languages

2. Noun and Adjective

- (11) cette maison blanche
that house white
'that white house' Dem N A (French)

- (12) rumah besar itu
house big that

‘that big house’

N A Dem (Malay)

1.3.3 Genitive

1. Genitive and Noun

- (13) ob da-guś
father his-tent
 ‘father’s tent’ GN (Ket)

2. Noun and Genitive

- (14) el coche de la mujer
the car of the woman
 ‘the woman’s car’ NG (Spanish)

It is difficult to specify which one of the English possessive constructions is the basic one.

- (15) a. Mary’s dog (Saxon genitive)
 b. the roof of the house (Norman genitive)

1.3.4 Relative clauses

1. Relative clause and Noun

- (16) gizon-a-k liburu-a eman dio-n emakume-a
man-the-SBJ book-the-give has-REL woman-the
 ‘the woman that the man has given the book to’ RelN (Basque)

2. Noun and Relative clause

- (17) the beer that I drink is cold NRel (English)

2 Word Order Correlations

At first glance, there is no reason for there to be a connection between basic clausal word order and the relative order of elements of the constituent of the clause. However, there exist many significant correlations.

2.1 Greenberg (1963)

Joseph Greenberg was the first to establish the importance of universal statements. The universals in his pioneering paper are mainly absolute and implicational.

- (18) UNIVERSAL 1
In a declarative sentence with nominal subject and object, the dominant order is almost always one in which the subject precedes the object.

All of Greenberg's implicational universals are *unilateral*, i.e. irreversible.

- (19) UNIVERSAL 25
If the pronominal object follows the verb, so does the nominal object.

Implicational chains:

- (20) UNIVERSAL 3
Languages with dominant VSO order are always prepositional.
- (21) UNIVERSAL 4
With overwhelmingly greater than chance frequency, languages with normal SOV order are postpositional.
- (22) UNIVERSAL 2
In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.

Universals involving complex correlation of conditions:

- (23) UNIVERSAL 24
If the relative expression precedes the noun either as the only construction or as an alternative construction, either the language is postpositional or the adjective precedes the noun, or both.

The first to decode that one wins a beer!

Greenberg also reveals the connection between word order and morpheme order.

- (24) UNIVERSAL 27
If a language is exclusively suffixing, it is postpositional; if it is exclusively prefixing, it is prepositional.

Dominance: prepositions are *dominant* over postpositions because the former have unrestricted distribution, that is, they occur regardless of word order (both in VSO

and SOV). Postpositions are *recessive* because they are not to be found in VSO languages.

Cross-linguistically the dominant word orders are the most common permutation.

Harmony: Postpositions are *harmonic* with SV and OV and **disharmonic** with VS and VO.

A dominant order may occur always, but a recessive order may occur only when there is a harmonic construction present.

2.2 OV-VO Typology

All word order rules are dependent on the relative position of the **V**erb and its **O**bject in such a way, that it is possible to predict all functional word order relationships only from the relative order of V and O.

2.2.1 Lehmann (1973)

Lehmann (1973) puts forth the *Fundamental Principle of Placement* (FPP) and reduces Greenberg's tripartite typology to *two basic word orders*: OV and VO.

The FPP stipulates that modifiers be placed at the **opposite** side of a basic constituent from its primary concomitant.

- OV languages:
 - Verbal (V) modifiers: negation, causation, reflexives are to the right of the verb.
 - Nominal (O) modifiers: adjectives, genitive and relative expressions are to the left of the noun.
- VO languages
 - Verbal (V) modifiers: negation, causation, reflexives are to the left of the verb.
 - Nominal (O) modifiers: adjectives, genitive and relative expressions are to the right of the noun.

$$(25) \quad \begin{array}{cc} \text{OV} & \text{VO} \\ \hline \text{AN} & \text{NA} \\ \text{V-vm} & \text{vm-V} \end{array}$$

Modifiers, whether verbal or nominal, will be placed on either sides of the complex of V and O, but never *inside it*.

2.2.2 Vennemann (1974)

The Principle of Natural Serialization (PNS): the order of operators and operand tends to be serialized in **one** direction, either operators before operands, or operators after operands.

	OPERATOR	OPERAND
	modifiers	modified elements
	dependents	heads
	object	verb
(26)	main verb	auxiliary
	adjective	noun
	genitive	noun
	relative clause	noun
	noun phrase	adposition

In other words, in OV languages the operand is placed consistently after the operator, whereas in VO languages the operand precedes the operator.

Greenberg lists the 24 logically possible combinations of the four word order parameters

1. SOV/SVO/SOV (Universal 1)
2. PrN/NPo
3. NG/GN
4. NA/AN

Vennemann's PNS positively sanctions only the operand-operator and the operator-operand sequensec, that is:

1. VSO/PrN/NG/NA
2. SVO/PrN/NG/NA
3. SOV/NPo/Gn/AN

3 Hawkins (1983)

Basically Hawkins abandons the verb-based word order typology and goes for a preposition-based one, claiming that adpositions are much better *type indicators*. Thus we have two basic word order types: prepositional and postpositional languages.

Hawkins proposes unilateral absolute implications, but they involve complex implicational relationships using three or more parameters.

- (27) a. $Po \rightarrow (AN \rightarrow GN)$
 b. $Pr \rightarrow (NA \rightarrow NG)$

Hawking introduces the *Heaviness Serialization Principle* (HSP) according to which “heavy” modifiers tend to occur to the right of lighter ones.

- (28) Heaviness Serialization Principle (HSP)
 $Rel \geq_R G \geq_R A \geq_R \{Dem, Num\}$
 “ \geq_R ” means “exhibits more or equal rightward positioning relative to the head noun across languages”

- (29) $Po \ \& \ NA \ \& \ RelN$

In order to explain (29) he invokes the *Mobility Principle* (MP) according to which Dem, A and Num are more mobile than G and Rel and can move more freely around their heads.

3.1 But why HSP?

There is a general linguistic preference for light modifiers to be positioned to the left and heavy modifiers to the right of the head. This preference is motivated by the **ease of processing**: early head recognition is crucial for the recognition of the whole basic clause structure.

3.2 ... and why MP?

The more mobile constituents have a less syntactically complex structure.

4 Explaining the basic clausal word order

In general, we can represent the frequency of the six basic word order types as follows:

$SOV > SVO > VSO > VOS > OVS > OSV$

4.1 Tomlin (1986)

Tomlin proposes three functional principles to account for the relative frequencies of the six basic word orders.

4.1.1 The Theme First Principle

- (30) *Information that is relatively more thematic precedes information that is less so.*
- a. Definite NPs precede indefinite
 - b. Pronouns precede full NPs

4.1.2 The Animated First Principle

- (31) *In simple transitive clauses the NP which is more animated will precede NPs which are less animated*
- a. Animacy Hierarchy: human > other animate > inanimate
 - b. Semantic Roles Hierarchy: Agent > Instrumental > Benefactive/Dative > Patient

When there is a conflict in the animacy, the semantic role hierarchy kicks in.

4.1.3 The Verb-Object Bonding Principle

- (32) *The object of a transitive verb is more tightly bonded to the verb than its subject.*

The more of the three principles is realized in a given word order, the more frequent this word order will be among languages.

	WO	TFP	AFP	VOB	Score
	SOV	+	+	+	3
	SVO	+	+	+	3
(33)	VSO	+	+	–	2
	VOS	–	–	+	1
	OVS	–	–	+	1
	OSV	–	–	–	0

5 The Branching direction Theory: The return of OV-VO Typology

Dryer (1992) takes the word order of the verb relative to its object to be predictor of various word order properties.

- identification the correlations between the word order of certain elements with that of the verb-object complex.
- explanation of the aforementioned correlations

Dryer rejects the *Head Dependent Theory* according to which dependents are placed consistently on one side of the heads.

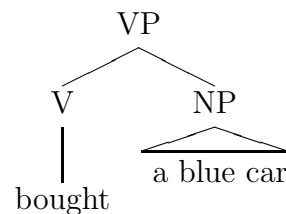
If we have the order $\langle X, Y \rangle$ in an OV language, X is said to be an *verb patternner* (head), Y is said to be a *object patternner* (dependent).

- (34) *Branching Direction Theory* (BDT)
Verb patternners are non-phrasal (non-branching, lexical) categories whereas object patternners are phrasal (branching) categories.

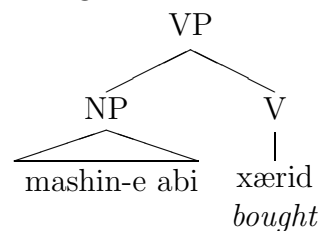
Languages can be divided in two types:

1. **Right-branching languages** - phrasal categories are to the right of (follow) the non-phrasal.
2. **Left-branching languages** - phrasal categories are to the left of (precede) the non-phrasal.

- (35) a. English (VO, left-branching)



- b. Persian (OV, right-branching)



The pair A and N is not a correlation pair with respect to V and O, since both A and N are non-phrasal (non-branching, lexical).

- (36) BDT differentiates between adjectives and adjective phrases in terms of positioning:
- a. *a rich country* (AN)
 - b. *a country rich in natural resources* (NA)

Still, BDT cannot explain the AN order in (36a).

Dryer explains the BDT with the ease of parsing. Human parsers are quicker when the branching is consistent in one direction. Mixed-branching is hard to process.

6 Early Immediate Constituents Theory

Hawkins (1994) proposes the Principle of Early Immediate Constituent, motivated by processing ease. According to this principle *word order in general reflects the way languages respond to the demands of quick and efficient processing in real time.*

Different permutations of constituents/words give rise to different level of structural complexity. Basic word order is the grammaticalization of the most optimal order in performance.

- (37) Different degree of structural complexity
- a. Mary rang up the boy.
 - b. Mary rang the boy up.
 - c. Mary rang the boy in the class up.
 - d. Mary rang the boy whom she met yesterday up.

Constituent Recognition Domain (CRD): The CRD for a phrasal node M consists of those terminal and non-terminal nodes that must be parsed in order to recognize M and all Immediate Constituents of M.

The shorter the CDR the more rapid the parse.

Hawkins invents the IC-to-non-IC ratio, which basically equals the number of immediate constituents in a CRD divided by the total number of the words in the CRD.

Human parsers prefer linear orders with maximal IC-to-non-IC ratios.

Word orders with higher IC-to-non-IC ratios are the most commonly attested.

7 From word order to morpheme order

Languages generally have preference for suffixes over prefixes.

	PREFIXES	SUFFIXES
PrN/VO	yes	yes
NPo/OV	(almost) no	yes

If a language has (exclusive) prefixing, it almost always has PrN and VO order.

If a language has NPo and/or VO word order, it has (exclusive) suffixing.

7.1 The Head Ordering Principle

Hawkins and Gilligan (1988)

- (38) *The Head Ordering Principle*
 The affixal head of a word is ordered on the same side of its subcategorized modifier as P is ordered relative to NP within PP, and as V is ordered relative to a direct object NP (O).

7.2 The diachronic approach

- (39) Affixes develop in whatever position they happened to be in when their grammaticalization occurred.

7.3 The psycholinguistic approach

Hawkins (1983): Humans achieve maximal processing efficiency, if lexico-semantic information has to be processed as early as possible.

In other words, if stems come before affixes people process the conveyed information better.

Furthermore, humans generally tend to listen carefully for a lamentably short time. Hence, the average listener pays maximum attention only to the beginning of words and it is wise to have the more informative part in the beginning.

References

- Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. University of Chicago Press, Chicago.
- Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68 1: 81–138.
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language. Report of a Conference Held at Dobbs Ferry, NY, April 13-15, 1961*, edited by Joseph H. Greenberg, pp. 73–113. MIT Press, Cambridge, Ma., 2nd edn.
- Hawkins, John A. 1983. *Word Order Universals*. Quantitative analyses of linguistic structure series. Academic Press, New York.
- Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge studies in linguistics; 73. Cambridge University Press, Cambridge; New York.
- Lehmann, Winfred P. 1973. A structural principle of language and its implications. *Language: Journal of the Linguistic Society of America, Los Angeles, CA* pp. 49, 47–66.
- Song, Jae Jung. 2001. *Linguistic Typology: Morphology and Syntax*. Longman, London, New York.
- Tomlin, R. 1986. *Basic Word Order: functional principles*. Croom Helm, London.
- Vennemann, T. 1974. Analogy in generative grammar: the origin of word order. *Proceedings of the Eleventh International Congress of Linguists* .