

9. September, 8.45-10.50

Anton Zimmerling (Moscow Open Pedagogical University/ Institute of Slavic Studies,
Russian Academy of Sciences)

meinmat@yahoo.com

Encoding strategies in Word Order: the evidence of Slavic Languages¹

1. Basic assumptions:

- Word order is predictable in all natural languages. It is possible to establish language-specific or type-specific systems of constraints which allow for grammatical linear orders and ban ungrammatical ones.
- All natural languages including the so called ‘free word order languages’ have some linearization constraints. Languages, where all linear orders in all clauses are equally acceptable, do not exist.
- Presumably, all natural languages have pairs (or sets with n elements, $n \leq 2$) of sentences with one and the same structure, but different linear orders.
- For each pair/set of such sentences it is possible to establish the variant representing the basic order and get the derived orders from it.
- It is possible to get the derived order from the basic one by singling out the element that moves: $\{a + b + c\} \Rightarrow \{b + a + t_b + c\}$.

2. Greenbergian typology: [Greenberg 1963; Hawkins 1979; Hawkins 1983; Dryer 1996].

- Word/constituent order embodies hierarchal relations of meaningful sentence categories.
COMMENT: This point is characteristic of every modern approach to word order studies, not only of Greenbergian typology though it was first proposed by Joseph Greenberg. We accept this point but we reject other claims made by Greenberg and his followers.
- All languages have some underlying rigid word order pattern.
COMMENT: This is an aprioristic postulate, it can be either taken or rejected, but it cannot be proved empirically.
- All world languages may be effectively described in terms of a very small number of presumably universal categories (S, O, V, etc).
COMMENT: This is merely an empirical issue. It can be verified or falsified with linguistic data.
- The clausal type (SOV, VSO, SVO etc) depends on branching conditions, ultimately – on the Head → Complement or Complement → Head order in the binary groups.
COMMENT: This is an empirical issue, too. There is strong empirical evidence in favor of the view that branching is inconsistent in a vast majority of languages.

NB: Greenbergian typology leaves no place for languages with ‘free’ word order, since it is assumed that all world languages many alternating surface orders (cf. Modern Russian, Old Greek, Pashto etc.) have some underlying rigid word order pattern.

¹ The paper was prepared with financial support of Russian Foundation of Humanities (RGNF), project № 06-04-00203a.

3. 'Free' or 'pragmatic' word order: two popular criteria.

- **3.1.** The language does not display any dominant word order (at least with respect to the placement of S, O, V etc.).
- **3.2.** The language lacks purely grammaticalized constraints on linear order (cf. the Verb-Second Constraint in German or Dutch).

4. Paradoxes of 'free word order'.

- The criterion **3.1.** does not imply **3.2.**, and vice versa.

5. Free word order languages must be redefined in terms of inversion, relative ordering of elements, cf. [Zimmerling 1999; 2002].

- Language L has free (=pragmatically determined) word order iff it permits inversion of the relative ordering in binary groups consisting of any two sentence categories X and Y (e.g., Direct Object and Subject, Infinitive Verbs and Postverbs, Infinitive Verbs and Subject etc).
- If language L permits inversion of the relative ordering in all possible groups of the type $[X > Y] \sim [Y > X]$, then word order alternations in L are ruled out by some linear-accent transformations, i.e. by principles both changing the placement of elements and assigning given prosodic status (specific of topic, focus, degraded topic etc) to them.

6. Can Verb-Second Languages and languages with Wackernagel's law be free word order languages?!

- !!! This proves possible, since relative inversion in binary groups is compatible with constraints on the placement of elements attached to sentence 1st, 2nd or 3^d positions.

7. Old Norse as a free word order language:

$[X > Y] \rightarrow [Y > X]$

1) — [V_{aux} + V_{lex}] cf. vildi láta drepa "wanted to let to kill" ~ vildi drepa láta

2) — [V_{inf} + PV].

3) — [V_{inf} + Obj].

4) — [Obj + PV].

5) — [Obj_{adr} + Obj].

6) — [Obj_{caus} + Obj].

7) — [Obj_{prep} + Obj].

8) — [V_{fin} + Subj].

9) — [Subj + Obj].

10*) — [V_{fin} + PV], 11*) — [V_{fin} + V_{inf}] 12) — [NP + DetPoss] 13) — [NP + AdjAttr]

8. The relevant differences between word order systems may be explained in terms of **encoding strategies**. It is necessary to distinguish at least three types as the dominating strategy is concerned:

- F- systems: word order systems encoding formal constituents (syntactic functions).
F = 'Formal' or 'Functional'

- C-systems: word order systems encoding topic-focus articulation.
C = ‘Communicative’
- W-systems: word order systems encoding tact groups (stress patterns).
W = ‘Word’ or ‘Wackernagel’

9. Pure and mixed F, C and W-systems.

- Pure F-systems: Danish, Swedish, English.
- Pure C-systems: Modern Russian, Lithuanian.
- Pure W-systems: Old Novgorod Russian, Serbo-Croat, Colloquial Czech, Warlpiri, Pashto.
- Mixed W/C/F systems: Old Norse, Bulgarian, Tagalog.
- Mixed F/C systems: Modern Icelandic, Faroese, Yiddish.
- Mixed F/W systems: Middle Mainland Scandinavian, Modern Norwegian.
- Mixed C/W systems: Gothic, Old Church Slavonic, Upper Sorbian.
- Degraded W-systems (= *W systems): Late Greek *koine*, Old Spanish, (?) Macedonian.

10. Focalization in F-systems vs C-systems:

Rus. Студенты бьют ПРОФЕССОРА → ПРОФЕССОРА бьют студенты

Eng. The students are beating the PROFESSOR → *The PROFESSOR are beating the students.

- C-systems usually allow for in-situ focus marking, but F-systems often ban movement of focalized elements.

11. Linear-accent transformations in C-systems [Paducheva 1984]:

Modern Russian:

1) Rheme shift: the final position is strongly associated with contrastive focus stress.

"The man who loves her is John"

Иван ↗ её любит ↘² ⇒ Любит ↗ её Иван ↘ ↘

"Ivan indeed loves her (the suspicion that he does not love her, is false)"

Иван ↗ её любит ↘ ⇒ Иван ↗ её любит ↘ ↘.

2) Contrastive beginning: the initial thematic constituent acquires contrastive stress:

"Ivan does love her, but someone else is not obliged to",

Иван ↗ её любит ↘ ⇒ Иван ↗ ↗ её любит ↘.

- The initial position is not iconic. It is somewhere associated with contrastive rhematic constituents, cf. *Только я ↘ ↘ мучился.*³ "Only I worried about it"

3) Emphatic preposition.

Иван ↗ её любит ↘ ⇒ Иван ↗ её любит ↘ ↘ (по правилу 2) ⇒ Любит ↘ ↘ её

Иван.

4) Rheme dislocation The rheme splits itself into two parts, whereby the part not bearing focus accent is moved to the left and placed in the sentence-initial position. This mechanism lowers the thematic constituent in the information hierarchy. The special intonation contour is **falling + even tone**, designed below as "↘→":

² Symbol ↘ denotes focus accent, symbol ↗ – thematic accent, symbol ↘ ↘ – contrastive rhematic accent, symbol ↗ ↗ – contrastive thematic accent.

³ The group *только X* can otherwise be a contrastive theme: *Вася ↗ проиграл ↘, Петя ↗ проиграл ↘, и только Иван ↗ ↗ выиграл ↘.* "Basil lost, Pete lost too, and only John won".

Мой доктор ↗ пустился в разговоры ↘ ⇒ Пустился ↗ мой доктор в разговоры ↘.

"My doctor began talking"

Rheme ₁	Wackernagel theme	Rheme ₂
Приходит ↗	Муж	домой ↘
Пустился ↗	мой доктор	в разговоры ↘

5) Rheme extraction:

Строители | живут по всей трассе канала **R** ⇒ Живут ↗ строители | по всей трассе канала; "The builders| live along the whole canal track"

Меня томила | страшная скука **R** ⇒ Скука ↗ меня томила | страшная; **R**

"I was oppressed by all this bore"

Она | превосходно выглядит **R** ⇒ Выглядит ↗ она | превосходно. **R**

"She looks perfectly well"

Theme	Wackernagel theme	Rheme
Скука ↗	меня томила	страшная ↘
Живут ↗	Строители	по всей трассе канала ↘

12. V 1 declaratives in C-systems: one linear order, different communicative structure:

1) **Thetic sentences:**

Rus. стоял мороз, выходят на площадку игроки "Автодора".

OSwe Liggær (1) lik (2)a wighwalli "(For instance) there lies a corps on the battle field"

2) **Rheme dislocation:**

Rus. | посадил_R | дед_T | пенку_R |.

OIce. | Sendir_R | Þordís_{TWack} | tvá menn í byggð_R | ok stefnir at sér mönnum

Lit. sends Þ. Two men to the country and calls people to join her

ok | skal-tu_{TWack} = 2Sg. Pres | víst vináttu í móti hafa = INF_R |

lit. and shall-thou sure friendship in exchange have

3) **Topicalized verb, rhematic subject:**

Rus. <служил на Кавказе офицером один барин>. | Звали_T | его_T | Жилин_R |.

<En Þiðrandi i kom til Kórekssona j um sumarit. | Tóku_T | þeir=NOM_j við honum_i T_{Wack} | einka vel_R |

lit. <But Þ. came to the sons of K. in sommer> received they him very well

Cf. OIce. | Af öllum hirðmönnum_{TContr} | virði konungr mest_{TWack} | skáld sín_{RContr} |

Lit. from all his followers esteemed the king highest his poets

4) **Rhematic verb, postonic theme:**

Rus. <Пришел больной к врачу> | Осмотрел_R | врач больного_T |.

OIce. | Vita_R | nú hvárir=NOM til annarra_T |

Lit. know now no men about each other

13. W-systems with Wackernagel's law are syntactically uniform, irrespective of the genetic relationship of the languages:

- Ordering of phrasal (sentential) clitics in clitic clusters is always rigid.
- No language puts 100 % clitics in clausal second position. Presumably all W-systems have either first word/first constituent variation or rules placing clitic clusters to the right of second position (called 'Skipping' or 'Barrier' rules in different formal theories).
- Claims made elsewhere in typological literature that some exotic languages with second-position clitics lack syntactic constituents, are erroneous.

- Clitics and clitic clusters, attached to second position in W-systems are always ‘syntactic clitics’, not ‘prosodic clitics’.
- Wackernagel’s law is always a grammaticalized syntactic mechanism and not an iconic reflection of oral speech, though the ‘Wackernagel position’ has a communicative load of its own and plays an important role in linear-accent transformations, characteristic of free word order languages.

14. SPLITTING (SR), BARRIER RULE (BR), COMPLEX CONSTITUENT RULE (CCR):

Serbian/Croat, South Slavic, Indo-European

- a) Dobar *se* čovjek lako prepozna (SR)
- b) Србхрв. [Dobar čovjek]= *se* lako prepozna (CCR)
- c) Србхрв. [Dobar čovjek] || lako *se* prepozna (BR +CCR)
- d) Србхрв. prepozna || *se* lako Dobar čovjek (‘prosodic inversion’)

15. BARRIER RULE, COMPLEX CONSTITUENT RULE (CCR):

Warlpiri, Pama-Nyungan, [Nash 1983]

- Clitics after the first spelled-out constituent)
[Kurdu=ngku (1) wita=ngku (2)] || *ka* (3) maliki wajilipi=nyi
[child= Erg. (1) little=Erg. (2)]|| LINK=PRES. (3) dog chase=not.pst
"a little boy is chasing a dog"

- Clitics move to right if the initial word is emphatic:

Tumaji (1) ngaju(2)=ku=ju (3) wiri ngawarra yali=ji
too much (1) I (2)=DAT. =to me (3) big river that=TOP
"That big river is too much for me"

Mayo, Uto-Aztecan [Freeze 1989]

wo:yi saila-*m-ne* hipu-re
dos/hermanos-*pl-yo/tener-r* ‘Tengo dos hermanos’

Haiki-*m-su-e*’ saila-k
Cuánto-*pl-enf-tu/hermano-r* ‘¿Cuántos hermanos tienes?’

[Haku-ni **entoko** hawe:y]-*su-e* biča-k
Dónde/y/quando-*enf.-tú/me-ver.-r* . ¿Dónde y cuándo me viste?

Cf. Rus. [Где и когда]=же=т(ы)=мя видел

16. “**Skipping**” or Syntactic Operation: do clitics really move (downwards) to the right from clausal-second position?

Slovak

XP	CL	(YP)	VP
Aby	<i>si</i>	človek	[predstavil srazu]

That	<i>for-oneself</i>	a man	[could imagine __at once]
------	--------------------	-------	---------------------------

Slovak

XP	Vfin	CL	YP
Vodič autobusu The bus driver	zapálil lighted	<i>si</i> for himself	cigaretu a cigarette

- Possessive and reflexive clitics in Slovak attract the verb to clausal second position.

17. Bulgarian (South Slavic) and Tagalog (Central Philippine, Austronesian): the same set of microparametres.

- #XP—Clitic —Vf ~ #Vf — Clitic, #*XP — ...—Clitic, #*XP — Clitic ...—Vf, #*Clitic—Vf, *Vf —...Clitic.

kahapon *ka ba nila* nakita
yesterday 2SgSbj Q 3Pl.DO be-seen
“Was it yesterday they saw you?”

Nakita *ka ba nila* (kahapon)
be-seen 2SgSbj Q 3Pl.DO (yesterday)
“Did they see you (yesterday)?”

- Clitic clusters don't move from clausal second position, verbal forms must take contact position if clitics are present, otherwise the position of verbal forms is not fixed.

18. Hypothesis:

- Bulgarian and Tagalog originated as pure W-systems but later developed a formal constraint on verb movement (adjacency parameter): the clitics ‘pied-pipe the verbs to adjacent positions, not vice versa.
- Bulgarian/Tagalog clitics are ‘strong’ with respect to Verb Movement: they don't give up clausal second position to any other sentence categories.
- Slovak clitics are ‘weak’: under certain condition they move out of clausal second position, - which is then filled by verbal forms.
- Germanic Verb-Second constraint might have been triggered by a mechanism of the Modern Slovak type: early Germanic clitics were ‘weak’ and moved out to the post-verbal position in a number of early Germanic idioms, cf. [Zimmerling 2002].
- In W-systems, Clitic Movement has a priority over Verb Movement.
- Cross-linguistically, word order systems with verbal clitics originate from W-systems (cf. Modern Romance languages and Modern Greek). The reversed evolution is not attested.
- Decline of W-systems in the Slavic area and beyond can be linked either with a drift toward F-systems and development of formal constraints or with gradual elimination of phrasal clitics and clitic clusters.

LITERATURE

Anderson S.P. The Wackernagel revenge // *Language*, 1993, 1, 68-93.

Billings L., Konopasky A. (2002) “The role of morphology in ordering verb-adjacent clitics: from syntax to prosody in Bulgarian and Tagalog.” // *Papers from the workshop*

“Language Change from a Generative Perspective,” Thessaloniki, Feb. 2002. Artemis Alexiadou *et al.*, eds. (Linguistics in Potsdam, vol. 19.) Potsdam: Institut für Linguistik, Universität Potsdam, pp. 1-26.

Chomsky N. A Minimalist Program for Linguistic Theory // The view from building 20. /Hale, K. S.L.Keyser (eds). Cambridge, Mass. MIT Press. 1993.

Clitics in the languages of Europe. Eurotyp 20-5. /Ed. by H. van Riemsdijk. Mouton de Gruyter. Berlin- New York. 1999. 1020 S.

Dryer M.S. Focus, pragmatic presupposition, and activated propositions // Journal of Pragmatics 26 (1996), 475 - 523.

Fiengo R. On Trace Theory // Linguistic Inquiry 8, 1977. 35-62.

Freeze R. A. Mayo de los capomos, Sinaloa. México. Archivo de lenguas indígenas de México, 1989.

Gazdar Gr, Klein E., Pullum G., Sag I. Generalized Phrase Structure Grammar. Harvard University Press. Cambridge, Mass, 1985.

Greenberg J.H. Some universals of grammar with particular reference to the order of the meaningful elements // Universals of language, Cambridge, Mass. 1963, p. 73-113.

Hawkins J. Implicational universals as Predictors of Word Order Change // LANGUAGE, 1979, vol. 55, N 3.

Hawkins J. Word Order Universals. N.Y., 1983.

Jakobson R. Les enclitique slaves // Selected Writings II. Word and language. The Hague - Paris: 1971, 16 -22.

Luraghi S. The pragmatics of Verb Initial Sentences in Some Ancient Indo-European Languages // Word Order in Discourse, Downing, Noonan 1995, 335-386.

Nash D. Topics in Warlpiri Grammar. Outstanding dissertations in linguistics. Cambridge, Mass., 1986.

Progovac L. Clitics in Serbian/Croatian: Comp as the Second Position // Approaching Second/ A.Halpern, A. Zwicky (eds.). CSLI Publications. Stanford, California, 1996, 411-428.

Radanović-Kocić V. The Placement of Serbo-Croatian Clitics: a prosodic approach // Approaching Second/ A.Halpern, A. Zwicky (eds.). CSLI Publications. Stanford, California, 1996, 429-445.

Ross 1967 – Ross J.R. Constraints on variables in Syntax. Ph.D Dissertation, MIT. 1967.

Wackernagel J. Über ein Gesetz der indogermanischer Wortstellung // Wackernagel J. Kleine Schriften. Bd. I, Basel, 1953, 1-103.

Zwicky A.M. On Clitics. Indiana University Linguistic Club. 1977.

Гладкий А.В., Мельчук И.А. Элементы математической лингвистики. М., Наука, 1969.

Дыбо В.А. О фразовых модификациях ударения в праславянском // Советское славяноведение, 1971, № 6, 77-84.

Дыбо В.А. Закон Васильева-Долобо в древнерусском языке на материале Чудовского Завета // International journal of linguistics and poetics, vol. 18/ 1, 1975, 7-81.

Дыбо В.А. Славянская акцентология. Опыт реконструкции системы акцентных парадигм в праславянском. I., 1981.

Зализняк А.А. «Слово о Полку Игореве»: взгляд лингвиста. М., Языки Славянской Культуры, 2004.

Падучева Е.В. Коммуникативная структура предложения и понятие коммуникативной парадигмы // НТИ, Сер. 2. 1984. N 10.

Перекрестенко А.А. Об автоматическом синтаксическом анализе в некоторых классах контекстно-зависимых языков //Московский лингвистический журнал, том 6, №2. Москва, 2003.

Циммерлинг А.В. Типологический синтаксис скандинавских языков. М., «Языки славянской культуры», 2002. 896 стр.

Циммерлинг А.В. Порядок слов и синтаксические позиции. // Труды международного семинара "Диалог 1998" по компьютерной лингвистике и её приложениям /А.С. Нариньяни (ред.). Казань, 1999.

Циммерлинг А.В. Отношение свободного порядка слов и модели инверсии. // Труды международного семинара "Диалог 2006" по компьютерной лингвистике и её приложениям. Москва, 2006, 540-544.

Циммерлинг А.В. Порядок слов в славянских, германских и романских языках. Славяно-германские исследования, т.3 Спб, Алетейя, 2006 (in press).