A Rule-Based Approach to Free Word Order Languages

Anton V. Zimmerling[1]
[1]Institute for Modern Linguistic Research, Sholokhov Moscow State University for the Humanities, Moscow, Russian Federation

Abstract – *This paper is addressed the problems of modeling a rule-based approach to free word order languages. Grammar-based parsing for free word languages is problematic. Functionalist models of free word order are based on context-sensitive rules which cannot be parsed automatically. Generative models based on Chomskyan and Stablerian minimalist grammars undergenerate word orders licensed by natural free order languages. I argue that word order alternations in Russian can be predicted by Linear-Accent Transformations (LAT) linking together pairs of sentences with the same numeration but different information structure. LAT theory is compatible with the postulate on basic word order. Each LAT rule amounts to a pair of operations <Active movement; Remnant movement>, which makes it possible to reset LAT rules as mildly context-sensitive. Deaccenting of sentence categories in Russian can be analyzed as remnant movement*.

Keywords: syntax, information structure, movement, scrambling, free word order, mildly-context-sensitive rules

1 Introduction

Natural language processing interacts as a research field with the theory of formal grammars, cf. [7], [16] and formal frameworks in linguistics, cf. [2], [6], [11]. Generative capacity of native speakers, i.e. their ability of generating/recognizing well-formed structures and sorting out ill-formed structures is usually interpreted as proof for the hypothesis that the grammars of all natural languages share a core corresponding to some class of formal grammars. However, it is an open issue, whether languages with free word order, where syntactic trees can be linearized in more than one way, can be effectively recognized by grammar-based parsers. Empirically adequate descriptions of free word order languages include context-sensitive rules reordering already generated trees. Context-sensitive languages cannot be generated and parsed automatically, but in recent decades an attempt of eliminating context-sensitive rules was made. So called tree-adjoining grammars [7] and minimalist grammars [16] generate and parse syntactic trees according to a down-to-top principle and operate with mildly context-sensitive rules. The most consistent and elaborate formal grammars based on mildly context-sensitive rules are Stablerian minimalist grammars, cf. [16], [17], [18]. Stablerian grammars have a number of restrictive constraints on movement and adjunction of sub-trees. They can be extended with a scrambling operator, i.e. operator responsible for word-order alternations which makes them a suitable tool for parsing free word order languages. Still, lifting basic minimalist constraints on movement and adjunction, in a combination with a scrambling operator leads to derivational crash, at least in a sub-class of free-word languages known as languages with unbounded scrambling [8], [4], [14]. Formal linguistic frameworks, such as Chomskyan Minimalist Program [2], [3] operate with locality conditions on movement and adjunction similar or identical to the constraints of Stablerian minimalist grammars. A number of authors have addressed the diversity of scrambling types in natural languages like Russian, cf. [9], [6], [1], [21], [28]. There is general consensus that productive scrambling patterns have a communicative motivation. Generative models of Russian syntax capitalize the idea that word orders licensed by minimalist grammars represent the core of Russian grammar, while those not licensed represent its periphery and are only possible with given prosodic markings and communicative status, e.g. with contrastive focus or in thetic sentences, cf. [1], [6]. Still, these models are too rigid, sort out many well-formed Russian sentences and do not always predict the correct mapping of word order to information structure, so the undergeneration/completeness problem remains unresolved. Functionalist models of Russian syntax capitalize the idea that linearization of syntactic trees and assignment of communicative status are triggered in Russian by the same set of transformational rules called Linear-Accent-Transformations (LAT rules). LAT rules changing both the placement of sentence elements and their communicative status/prosodic markings correctly predict the diversity of word orders associated with one and the same numeration, which has been demonstrated in [9], [11], [21]. However, LAT theory in its current shape has a substantial drawback, since LAT rules are context-sensitive. Therefore, word order calculus based on LAT rules is impossible.

I aim at combining the advantages of the generative and functionalist models and offer a transformational approach to free word order languages like Russian, based on modified LAT rules. The paper has the following structure. In section 2, I render the notions of scrambling, conditions on movement and argue that Russian is a language with unbounded scrambling and direct prosodic marking of communicative status. In section 3, I argue that word order calculus based on LAT rules is feasible, if one adopts a postulate on basic word

order and defines LAT as mildly context-sensitive by resetting each rule as a pair <Active movement; Remnant movement>. Russian can be described by a small set of unidirectional LAT rules linking together pairs of sentences with the same numeration but different information structure. Deaccenting of sentence elements can be analyzed as remnant movement. LAT rules in Russian make use of 5 syntactically relevant prosodic markings, 4 of which belong to a strictly limited inventory of Russian tonemes, 'intonation constructions'.

2 Scrambling and formal grammars

The term 'free word order' is metaphoric, since all natural languages are restrictive: no language allows all possible linear orders or sentence categories in 100 % of sentences. It is reasonable to think that linearization constraints are present in all word order systems. However, there is a general consensus that free word order is a condition, when sentence categories can be linearized in two or more different ways at least in some well-formed sentences of a given language. This condition is also known as scrambling of predicate arguments and/or other sentence categories. It has become customary to classify natural languages into a class of languages with a fixed order of lexical sentence categories and a class of scrambling languages. For instance, an English sentence like *Pete ate a tomato* does not have a linear variant **A tomato ate Pete*, since this language blocks for OVS orders[1]. The class of scrambling languages can be defined in two ways – either as a) languages displaying a number of diagnostic movement patterns responsible for alternations like SVO > VSO, SVO > OSV, SVO > OVS, SVO > SOV; or b) languages completely lacking any fixed order of diagnostic sentence categories, say S and O or S, O and V. Both approaches proceed from the assumption that the same numeration, i.e. tree structure with a given number of positions filled by identical elements, may be linearized differently. A movement approach to scrambling languages capitalizes the idea that there is a unidirectional relation between different linear variants of the same numeration, one of the variants being the source of the other (s), cf. the presumably base-generated order in Rus. […] *Pet'a s''el pomidor* and the derived order [*Pomidor*$_i$] *Pet'a s"el* t_i: the symbol *t* marks the original placement of the moved category before the reordering, and the brackets […] mark the target position of the movement. A non-movement approach to scrambling denies the idea of a fixed order of sentence categories in a scrambling language and treats all linear variants as representing the same level of derivation. The domain, where categories scramble may be called scrambling domain. In the standard case illustrated by the Russian examples above, argument scrambling is bounded with a single clause, while all scrambled arguments S, O..U..W belong to one and the same verbal head v^o:

(i) $[_S\{_{\text{SCRAMBLING DOMAIN}} \ldots S \ldots v^o \ldots O \ldots\}]$.

Fig. 1. Local Scrambling.

Scrambling of the type (i) is called local or bounded; it does not pose big problems for linguistic theory with either non-movement or movement analysis, since all positions available for a scrambled category are in one and the same domain. Natural languages also have unbounded scrambling, where the permuting arguments may belong to different verbal heads v^1, v^2.. v^n. This has been proven in [15] for German, where unbounded argument scrambling takes place in complement clauses in the domain between the complementizer and the verbal complex, cf. (ii). Note that the verbal heads themselves are placed in a rigid order, so that the scrambling domain is more narrow than the complement clause:

(ii) $[_{CP}$ Comp $\{_{\text{SCRAMBLING DOMAIN}}$ A^1 +B^2 + $C^3\}$ $[_{VP}$ $[v^3$, $[v^2, [v^1]]]$ AUX].

Fig. 2 Unbounded Scrambling in German.

Stablerian minimalist grammars [16] and generative grammars based on Chomskyan Minimalist Program [2] generate ordered trees. Grammars of this type are mildly context-sensitive [4]. They can be adjusted for parsing scrambling languages, if their formalism is extended by a special Scrambling operator in addition to standard Merge and Move operators responsible for merging and moving sub-trees [14].

Under movement analysis, the scrambling type (local vs unbounded) is established in the end positions of the scrambled elements, not in their original positions before the reordering. There is a different tradition, where scrambling is understood as a characterictics of original domains. J.Baylin [1] distinguishes 'short' scrambling, when an element moves to a target position in the same clause and 'long-distance scrambling', when an element is extracted into a higher clause. However, extraction entails scrambling in the end domain only if the extracted element has more than one available target position in the higher clause. In the most simple case, under local scrambling, the elements remain in the same clause, so the original and the end domains match. This matching does not hold for unbounded scrambling. Therefore, it would be better to reserve the term 'scrambling' only for the distinction 'local vs unbounded' but replace it by the term 'movement' in Bailyn's opposition of 'short vs long-distance scrambling'.

Minimalist grammars (MGs) fail to parse languages with unbounded scrambling, cf. [8]. Both Chomskyan and Stablerian MGs are tree-adjoining grammars with a Move-operator, Merge-operator, Scramble-operator, Adjoin-operators and a number of locality conditions on movement

[1] A sentence like *A tomato ate Pete* will be proven well-formed if we assume that carnivorous vegetables exist, but this sentence won't get a linear variant *Pete ate a tomato* used in the same bizarre meaning "A human has been eaten by a vegetable". Consequently, the ungrammaticality of the SVO > OVS alternation in English does not depend on ontological assumptions about carnivorous vegetables and hungry humans.

and adjunction, known as Shortest Move Condition, Specifier Island Condition, Adjunct Island Condition etc. [3], [4]. MGs generate mildly-context-sensitive languages, while Chomskyan MGs aim at describing all natural languages as mildly-context-sensitive. The most salient condition affecting computational efficiency of a MG is the Shortest Move Condition (SMC)[2]. It requires that an element moves to the closest available target (Specifier of some P). Unbounded scrambling is a severe violation of the SMC. Blind application of the SMC to linear orders would mean that certain well-formed structures attested in languages like Russian or German would not be recognized/generated by an MG. Lifting the SMC and preserving the Specifier Island Constraint (SPIC) leads to Turing-equivalent grammars, i.e. to a derivational crash [8]. Moreover, MGs with unbounded scrambling (i.e. with scrambling without the SMC) and a single indiscriminating barrier[3] make the recognition problem NP-hard[4] [14].

2.1 Russian as an unbounded scrambling language

In this section I show that Russian is an unbounded scrambling language. Sentences with three scrambled NPs A^1, B^2, C^3 linked with three hierarchically arranged verbal heads are rare. Sentences with two scrambled NPs A^m, B^n, linked with two hierarchically arranged verbal heads v^m, v^n are wide-spread. One of the common cases of long-distance unbounded scrambling is triggered by non-projective embedding of a constituent or its element into a higher clause. Let A° B° C° D° E be the basic word order, A° B° C° D° be lexical heads and each next head be a dependent of the preceding one. It gives a projective structure (1), where blocks DE, CDE, BCDE, ABCDE are embedded constituents:

(1) [A ° [B° [C° [D °E]]]].

Moving the blocks DE, CDE and embedding the heads A°, B° into lower constituents one can get orders like $[CDE]_i$ A°B° t_i, $[[DE]_j$ C° $t_j]_i$ A°B° t_i, $[[DE]_j$… A°$_k$ … C° $t_j]_i$ t_k B° t_i, …A°$_k$ …$[[DE]_j$ C° $t_j]_i$ t_k B° t_i , where $t_{i, j, k}$ – traces of the moved heads or blocks. An illustration is provided in Fig. 3.

(1') Ru. *Arbitry*[1] *ne imeli prava*[1] [$_{IP}$ *fiksirovat'*[2] [*pobedu*[2] *«Triumfa»*]]][5].

'The referees[1] had no right[1] to fix[2] the victory[2] of 'Triumph''.

	Pattern	Word orders
Basic word order	[A° [B° [C° [D° E]]]]	(1a) *Arbitry*[1] *ne imeli prava*[1] [$_{IP}$ *fiksirovat'*[2] [*pobedu*[2] *«Triumfa»*]].
Derived orders	$[CDE]_i$ A°B° t_i,	(1b) ⇒ [$_{IP}$ *Fiksirovat' pobedu «Triumfa»*] $_i$ *arbitry ne imeli prava* t_i.
	$[[DE]_j$ C° $t_j]_i$ A°B° t_i,	(1c) ⇒ [*Победу «Triumfa»*] $_j$ [$_{IP}$ *fiksirovat'* t_j] $_i$ *arbitry ne imeli prava* t_i.
	$[[DE]_j$… A°$_k$ … C° $t_j]_i$ t_k B° t_i	(1d) ⇒ [*Pobedu «Triumfa»*] $_j$ *arbitry*$_k$ [$_{IP}$ *fiksirovat'* t_j] $_i$ t_k *ne imeli prava* t_i..
	…A°$_k$ …$[DE]_j$ C° $t_j]_i$ t_k B° t_i	(1e) ⇒ *Arbitry*$_k$ [*pobedu «Triumfa»*] $_j$ [$_{IP}$ *fiksirovat'* t_j] $_i$ t_k *ne imeli prava* t_i..

Fig 3. Long-Distance Unbounded Scrambling in Russian

2.2 Undergeneration problem

Since it is impossible to get the set of well-formed sentences of an unbounded scrambling language L by combining a scrambling operator with the SMC, all MGs undergenerate languages with unbounded scrambling and do not license some sentences which native speakers treat as well-formed. Bailyn's account of Russian short scrambling crucially relies on the SMC. He assumes that the basic word order in Russian is SVO, but with a wide class of transitive verbs every argument NP is equidistant from the preverbal position. This correctly predicts that OVS orders can be as discourse-neutral as SVO orders [1]. However, the architecture of his model urges him to make a wrong prediction that scrambled orders with a fronted verb and two-three post-verbal arguments should be unacceptable. In fact, sentences like (2) are well-formed in Russian.

(2) Dal učitel' včera knigu mal'čiku V-S-O-IO <a mal'čik eje zabyl posmotret'>

gave teacher-NOM book-ACC boy-DAT yesterday

"The teacher gave a book to the boy yesterday. <but the boy forgot to look it through.> "

T.L.King assumes that the basic word order in Russian is VS(O), while all other orders arise due to topicalization and focalization [6]. That means that Russian VS-sentences are thetic and base-generated, while Russian SVO ~OVS ~ SOV ~ OSV sentences are categorical and derived by overt movement. There are multiple issues with this approach. Russian thetic sentences are compatible both with VS and

[2]. Earlier called the Minimal Link Condition.

[3]. A non-discriminating barrier is a barrier not sensitive to the type of syntactic category the movement of which it blocks.

[4]. An NP-hard recognition can be fulfilled in a polynomial time, only if the language is NP-hard. This possibility cannot be eliminated as such, but practically it means that natural languages with unbounded scrambling are unparsable for MGs.

[5] For the sake of simplicity I treat the predicate *imet' pravo* 'have a right' as a single element.

with SV order [20], [30], while many Russian verb-initial sentences are categorial, non thetic, cf. [21], [28]. MG-based models still can be a useful, if they explain how the unparsable residue of well-formed sentences is derived from the alleged MG-compatible core. In the current generative research this problem is not solved yet.

2.3 Syntax vs information structure

Functionalist models of free word order capitalize the idea that scrambling is triggered by the same mechanisms that assign communicative status to sentence categories. The most elaborate formal model of the syntax-to-information structure interface in Russian was proposed by I.I.Kovtunova [9] and developed in [11], [12], [21], [28]. Kovtunova and her followers claim in Russian well-formed sentences with the same numeration but different constituent order and communicative status of elements are generated by a set of transformational rules called Linear-Accent Transformations. They also raise a claim that communicative categories are intrinsic features of sentences :

> (iii). Topic, focus, contrast and other communicative categories are intrinsic features of sentences, which take the same value by the speaker and the addressee.

A well-formed Russian sentence with topic and focus elements must have characteristic phrasal accents associated with topics and foci [10], [22], [24]. If topic, focus and other kinds of communicative status are intrinsic characteristics of sentences and Russian marks them prosodically, it is natural to assume that there is only one correct way to associate communicative structure and phrasal accents it correlates with.

.

> (iv) Russian is a language with direct prosodic marking of communicative status. Any correct interpretation of a well-formed Russian sentence implies the reader/addressee's ability to reconstruct the path from communicative semantics to phrasal accentuation, and vice versa. Communicative and prosodic structure are mapped to each other in a one-to-one correspondence in Russian.

Note that according to (iv) the information structure of a sentence may be reconstructed either from the verbal context or from phrasal prosody. The claim that Russian has direct prosodic marking of communicative status needs clarification. The two main intonation constructions, IC-1 ('↘') and IC-3 ('↗') are double-loaded: the falling pattern IC-1 marks either the focus in declarative sentences or the non-question constituent in yes-no questions, while the raising pattern IC-3 marks either the topic in declarative sentences or the question component in yes-no questions [5], [21]. Therefore, the claim that communicative and prosodic structure are mapped to each other in a one-to one correspondence in Russian can only mean that a) a sequence of Russian phrasal accents bears all the necessary information for the reconstruction of topic-focus articulation in a sentence, cf. [21], b) each relevant phrasal prosody in Russian can be interpreted as topic, focus marker etc. if we get the whole sequence of phrasal accents or learn the type of a sentence, whether it is a declarative, a question etc. The exact number of Russian phrasal tonemes and the limits of the allophonic variation in the realizations of Russian ICs is still an open issue, cf. different approaches in [24], [10]. I nevertheless argue that one does not need the whole alphabet of Russian tonemes for the purposes of word order calculus and that it is possible to build a LAT grammar using tags for just four phonologically distinct patterns: IC-3, IC-6, IC-1 and IC-2. The contrast in the pairs of two rising accents (IC-3 vs IC-6) and two falling accents (IC-1 vs IC-2) can probably be accounted for in terms of timing [28], but I simply state here that all these four accents are perceptually different.

IC	Tag	Communicative load
IC-3	↗	1) Topic. 2) The question component.
IC-6	↱	1) 2nd, degraded topic. 2) Left edge of a dislocated focus.
IC-1	↘	1) Focus. 2) The non-question component.
IC-2	↘↘	1) Focus. 2) The non-question component.

Fig. 4 Phonologically and Syntactically relevant Tonal Prosodies in Russian.

I will of speak of IC-3, IC-6, IC-1 and IC-2 as *non-zero accentual markings*. In addition I introduce an extra marking which does not correspond to any phonologically relevant prosody: it denotes a deaccenting operation. A tag like '$_0$X' reads 'constituent X got deaccented'[6].

IP	Tag	Prosodic cues
*	$_0$X	- (even tone/ absence of a tonal accent)

Fig. 5 Syntactically relevant 'eliminated accent'

I place the basic accentual tags before the constituent they attach to. This is done on several reasons, one them being the need to distinguish the basic tag for a given IC and the additional tag for its discourse-driven allophones[7].

[6]. Phonetic details, such as absolute pitch level (High, Semi-High, Semi-Low, Low), dynamic range, phonation type, tempo etc. do not bear for syntactic purposes — only the absence of a tonal movement is relevant.

[7] E.g. the two falling focal accents '↘X' or '↘↘X' are sporadically replaced in coherent speech by a rising tone marking the incompleteness of a text fragment: the latter option is tagged 'X↗'.

3 C-paradigms and LAT grammar

This section contains a list of transformational rules that derive accentually marked Russian sentences and form the *C*(ommunicative) paradigm of a sentence. I define a *C*-paradigm as a set of sentences sharing the same numeration i.e. constituent structure and given amount of lexical categories but having different constituent order and/or accentual tags. I define *LA–transformations* as rules, which both change constituent order and accentuation of at least one communicative constituent. Several issues have to be clarified. 1) LAT-rules establish the linkage and derivation vector between well-formed sentences with the same numeration. They do not create new syntactic positions. 2) Contrary to the original claim made in [11], LAT-rules are non-synonymic and can change the boundaries of communicative constituents. 3) The shift of accentual-marking from ↗X ~ ↘X ~↘↘X to $_0$X is a transformation**.** 3) LAT-rules are context-sensitive but can be reset as mildly context-sensitive [28]. 4) LAT-rules must be accounted for in terms of overt movement, not adjunction. 5) LAT-rules do not cover the so called afterthought elements. If LAT-rules are set as context-sensitive and all variants in a C-paradigm are mutually derivable, as [11] and [21] suggest, no word order calculus is possible. However, one can modify LAT-theory, combine it with the postulate on basic word order and reset LAT as mildly-context sensitive unidirectional rules. I am suggesting a procedure based on several interface principles. Topic and Focus are analyzed as communicative phrases headed by Topic Proper and Focus Proper. In the basic LAT-variant of a declarative sentence the boundaries of the TopicP and grammatical subject overlap. The FocusP consists of a Focus Proper phrase, where the main focus accent is located, and a transitional zone (Transition). The transitional elements are analyzed as belonging to FocusP, not to TopicP.

	TopicP		FocusP	
Communicative structure		Topic proper	Transition	Focus proper
Syntactic structure	Grammatical subject		Grammatical Predicate	
	External argument		Verbal head	Complements (Internal arguments & adjuncts)

Fig. 6 Prototypic mapping of information structure and syntactic structure for a verb with internal arguments

The previous research has shown that all previously LAT-rules have communicative motivation, be topicalization, focalization, splitting of FocusP etc. At the same time, all previously described LAT-rules also have predictable side effects, such as deaccenting of elements crossed by other elements undergoing leftward movement. E.g., a derived verb-initial sentence ↱*Posadil*$_i$ $_0$*ded* t$_i$ ↘*repku* 'Gramps planted a turnip' is generated from a basic SVO sentence [$_{TopicP}$ ↗ *Ded*] [$_{FocusP}$ *posadil* [$_{FocusProper}$ ↘*repku*]], where both the topical subject ded 'gramps' and the focal VP have their characteristic prosodic markers. In the derived V1 sentence ↱*Posadil*$_i$ $_0$*ded* t$_i$ ↘*repku* splitting of FocusP has a side effect in the deaccenting of the subject *ded*. The deaccenting of *ded* in such structures is not the primary communicative goal of the LAT-rule, but a consequence of the fact that the dislocated verb *posadil* 'planted', which lacked a tonal prosody in the basic SVO variant, was moved, got a special accent marking '↱' characteristic of dislocated elements, and crossed the node *ded* on its way. I suggest that LAT-rule 'Verb dislocation' examplified by this pair of sentences, can be reset as a pair of operations <Active movement (dislocation of the verb) & Remnant movement (deaccenting of the subject)>. Remnant movement is kind of compensatory effect responding to active movement. Russian data prompt that active movement always correlate with acquiring new non-zero prosodic markings, which is generalized in (v).

> (v) LAT-rules in Russian can be reset as pairs of operations <Active movement; Remnant movement>. Active movement puts sentence elements into target positions, where they get non-zero accent markings: '↗X', or'↘X', or '↘↘X', or'↱X'). Operations, which put sentence elements into positions with zero accent $_0$X, instantiate Remnant movement.

In order to make LAT grammar of Russian feasible, one has to adopt one more further postulate :

> (vi). Russian thetic sentences are derived through LAT-rules from categorial sentences with the same numeration, but non vice versa.

I argue that all Russian thetic sentences irrespective of their surface order (SV, VS, VSO, SVO) are derived from categorial sentences by deaccenting their theme. Topic daccenting results from Left Focus Movement i.e. an operation moving a postverbal complement X which bears the focus accent (schematically - ↘X) to the left for its governing verbal category. The moved element gets a reinforced focus accent (↘↘X): [$_{VP}$ V° ↘ X] ⇒ ↘↘X$_i$ …V° t$_i$. Russian does not allow post-focal accented themes [22]. Therefore, if a focal element moves outside VP and crosses the position of an accented thematic subject marked with IC-3 (↗X), the subject gets deaccented (↘$_0$X): [$_{NP}$ ↗S°] [VP V° ↘X] ⇒ ↘↘X$_i$ **[$_0$S]**…V° t$_i$. This explains why Russian thetic sentences can be realized both with SV and VS-orders since ↘↘S$_0$V structures like *Babuška spit* 'Grandma is asleep' are just inverted variants of $_0$V↘S structures, cf. the derivation in (4). Neither King's nor Bailyn's analysis fits Russian VS-

sentences since VS-orders apart from marking theticity can also mark three types of categorial sentences. The inverted verb can be a) the theme (↗V) b) the rheme (↘V) c) part of the dislocated rheme, schematically marked as (↱V). In all these cases the verb gets different accent markings.

(3) <Pocemu tak malo narodu?> $_0$Direktor $_0$p'at' $_0$sotrudnikov v [$_{FocusProper}$ ↘ ↘ *komandirovku*$_i$] $_0$poslal t_i.
<'Why so few people here?> 'The director has sent five workers to *a business trip*'.
Thetic, S-DO-IO-V.

(4a) ↗ Babuška ↘ spit. ⇒ (5b) $_0$Spit [$_F$ ↘babuška]. ⇒ (5c) [$_F$ ↘ ↘ Babuška]$_i$ spit t_i.
grandma-Nom sleeps sleeps grandma-Nom

I introduce 4 basic symbols for communicative phrases — F (Focus), T (Topic), Tr (Transition), SF (Dislocated Focus component. Symbol '↦' stands for 'movement of X to the right from its base position', 'right movement' symbol '↤' reads 'movement of X to the left from its base position', ' left movement', symbol '↧' reads 'deaccenting of X'. Formula X/F ↦ ' reads 'right movement of X to a clause-final Focus position', formula ↤ X/T' reads 'left movement of X to the clause-initial Topic position', formula '↤ Tr/T' reads 'left movement of X from the base position of Tr(ansition) to the clause-initial Topic position' etc. The symbol '&' is inserted between the formulas of the Principal and Remnant movement (cf. X/T & Y/F) and signals a determinist relation between Active and Remnant movement patterns. A formula like X/F↦ & ↧ T, which describes Right Focus Movement, reads 'Right movement of X to the clause-final Focus position; deaccenting of the node T is due to the fact that X crosses node T'.

	Rule	Operation	Active movement	Remnant movement
1.	Right Focus Movement	X/F ↦ & ↧ T	X/F ↦	↧ T
2.	Left Focus Movement	↘F/↘ ↘F ↤ & Tr	↘F/↘ ↘F ↤	Tr
3.	Verb Topicalization	Tr/T ↤ & ↧T	Tr/T ↤	↧T
4.	Dislocation	Tr/SF ↤& ↧T	Tr/SF ↤	↧T
5.	Verb Focalization	Tr/F ↤ & ↧T , ↧F	Tr/F ↤	↧T, ↧F
6.	Topic-Focus Inversion	F/T ↤& T/F ↦	F/T ↤	T/F ↦

Fig. 7. Linear-Accent Transformations in Russian.

4 Conclusions

Word order in scrambling languages like Russian can be predicted by Linear-Accent Transformations changing both the placement of sentence elements and their communicative status. Each LAT rule amounts to a pair of operations <Active movement; Remnant movement>, which makes it possible to reset LAT as mildly context-sensitive rules.

5 Acknowledgements

The paper is written with financial support from the Russian Foundation of Sciences, project RSCF 14-04-18-03270 'Word order typology, communicative-syntactic interface and information structure in world's languages'.

6 References

[1] John F. Bailyn. "Generalized inversion"; Natural Language and Linguistic Theory, 22, 1-49, 2004.

[2] Noam Chomsky. "A Minimalist Program for Linguistic Theory" ; The view from building 20. /Hale, K. S.L.Keyser (eds). MIT Press, 1993.

[3] Noam Chomsky. "Three factors in language design" ; Linguistic Inquiry, 36: 1–22, 2005.

[4] Hans Martin Gärtner & Jens Michaelis. "Some remarks on the Locality Conditions and Minimalist Grammars" ; Interfaces + Recursion = Language? Chomsky's Minimalism and the View from Syntax and Semantics. Mouton de Gruyter, 162-195, 2007.

[5] Nina Y. Shvedova et alii (eds.). Russkaja grammatika. Moscow: Nauka, 1982.

[6] Tracy H. King. "Configuring Topic and Focus in Russian". CSLI, 1995.

[7] Aravind K. Joshi. "Tree adjoining grammars: How much context-sensitivity is required to provide reasonable structural descriptions?"; D. R. Dowty, L.Karttunen, and A.M. Zwicky (eds.), Natural Language Parsing. Psychological, Computational, and Theoretical Perspectives, 206–250. Cambridge University Press, 1985.

[8] Gregory M. Kobele and Jens Michaelis. "Two type 0-variants of Minimalist Grammars"; Jäger, G., Monachesi, P., Penn, G., Wintner, S., eds. FG-MOL 2005. Proceedings of the 10th conference on Formal Grammar and the 9th Meeting on Mathematics of Language, Edinburgh, Scotland (2005).

[9] Irina I. Kovtunova. "Sovremennyj russkij yazyk. Por'adok slov i aktual'noe členenie predloženija." ['Modern Russian. Word Order and Communicative Perspective' – in Russian]. Moscow: Nauka, 1976.

[10] Cecilia Odé. "Perspektivy opisanija i transkripcii russkoj intonacii v korpusah zvučasčix tekstov" ['Perspectives of describing and transcribing Russian

intonation in Russian sound corpora' – in Russian]; Problemy i metody ekxperimental'no-fonetičeskix issledovanij. K 70-letiju professora L.V.Bondarko. Sankt-Petersburg, 2003.

[11] Elena V. Padučeva. "Kommunikativnaja struktura predloženija i ponjatie kommunikativnoj paradigmy " [= 'Communicative structure and the notion of communicative paradigm' – in Russian]; Naučno-texničeskaja informacija, Series 2. N 10, 25-31, 1984.

[12] Elena V. Padučeva. "Vyskazyvanie i ego sootnesennost' s dejstvitel'nost'ju" ['Utterance and its reference' – in Russian]. Moscow: LKI, 2008.

[13] J.B.Pierrehumbert. "The Phonology of English Intonation". M.I.T. Doctoral dissertation, 1980.

[14] Alexander Perekrestenko. "Minimalist Grammars with unbounded scrambling and non-discriminating barriers are NP-hard"; Martin-Vide, Carlos, Friedrich Otto, Henning Fernau (eds.), Language and Automata Theory and Applications [Lecture Note on Computer Science], Springer, 321-432, 2008.

[15] Owen Rambow. "*Formal and Computational Aspects of Natural Language Syntax"*. PhD thesis. IRCS Technical Report, University of Pennsylvania, 1994.

[16] Edward P. Stabler. "Derivational minimalism"; Christian Retore, ed., Logical Aspects of Computational Linguistics. Springer, p. 68–95, 1997.

[17] Edward P. Stabler. "Acquiring languages with movement"; Syntax, 1: 72–97, 1998.

[18] Edward P. Stabler. "Remnant movement and complexity"; G. Bouma, G.-J. M. Kruijff, E.Hinrichs, and R. T. Oehrle, (eds.), Constraints and Resources in Natural Language Syntax and Semantics, pp. 299–326. CSLI Publications, 1999.

[19] Edward P. Stabler. "Recognizing head movement"; P. De Groote, G. Morrill, C. Retoré (eds.). Logical Aspect of Computational Linguistics, Springer, 245-260, 2001.

[20] Tatyana E. Yanko. "Kommunikativnaja struktura s neingerentnoj temoj" ['Communicative structure with a non-inherent theme' – in Russian]; Naučno-texničeskaja informacija, Series 2., 25-32, 1991.

[21] Tatyana E. Yanko. "*Kommunikativnye strategii russkoj reci v sopostavitel'nom aspekte"* ['Communicative strategies of Russian speech in a comparative aspect' – in Russian]. Moscow, 2001.

[22] Tatyana E. Yanko. (2008) "*Intonacionnye strategii russkoj reci v sopostavitel'nom aspekte*." ['Intonation strategies of Russian speech in a comparative aspect' – in Russian], Moscow, 2008.

[23] Tatyana E. Yanko. "Accent placement principles in Russian"; Computational linguistics and intellectual technologies, vol. 10 (17), 712-724, 2011.

[24] Olga Yokoyama. "Neutral and non-neutral intonation in Russian: A reinterpretation of the IK system"; Die Welt der Slaven. XLVI. 1-26, 2001.

[25] Anton V. Zimmerling. "*Tipologičeskij sintaksis skandinavskix yazykov*." ['Typological Scandinavian Syntax' - in Russian]. Moscow: Yazyki slavyanskoj kul'tury, 2002.

[26] Anton V. Zimmerling. "Topic-Focus Articulation, Verb Movement and the EPP in Russian"; Slavic Linguistic Conference SLS 2. Berlin, 2007.

[27] Anton V. Zimmerling. "Locative Inversion and Right Focus Movement in Russian". Moscow, 2008.

[28] Anton V. Zimmerling. "Local and global rules in syntax" ; Computational linguistics and intellectual technologies, vol. 7 (14), 551-562, 2008.

[29] Anton V. Zimmerling. "Scrambling types in Slavic languages" ; Computational linguistics and intellectual technologies, vol. 10 (17), 750 - 766, 2011.

[30] Anton V. Zimmerling. "Word order, intonation and thetic sentences in Russian" ; Linguistics Association of Great Britain, LAGB 2013 Annual Meeting. Abstracts. University of London, School of Oriental and African Studies, 28-31 August, 2013. P. 67.