

The geometrical structure of linguistic theory: The case of Russian split scrambling^{*†}

THOMAS JOLYON WOOD
UNIVERSITY OF EDINBURGH

1 INTRODUCTION

1.1 ‘Argument Structure’

In section 2 I introduce the frameworks later to be compared with each other against Russian Split Scrambling data: Base Generation (2.2); Simple Movement (2.3); Distributed Deletion (2.4); and Representation Theory (2.5).

Then in section 3 I present the majority of my data, largely along formal lines—with respect to predicate structure (3.2) from in-to ditransitives, finding subject-object and direct-indirect object asymmetries; through the large (in terms of order and position) but not unlimited (with respect to multiplicity) permutational possibilities of 3.3; to, in 3.4, the ‘categories’ derived from traditional parts of speech, with prepositions and adverbs of special interest; to what seem like more or less impermeable complex adjectives, nouns, and sentences in 3.5—but also with an eye to meaning (3.6), from ‘neutrality’ to ‘style’ to ‘salience’ to ‘significance’.

Section 4 is devoted to theoretical issues that emerge from the data and confront the analyst—grouped into three sections, each with the first subsection by way of preface, and at every stage some questions about Chomskyan ontology: 4.2, with material on Phrase Structure and the DP Hypothesis; 4.3, which is typological in orientation and focuses on locality as well as landing sites; and 4.4, which formalises speaker intuitions about meaning into ‘discourse’ or ‘information’, as well as addressing connected questions about ‘interfaces’, ‘optionality’, and ‘gradience’.

Section 5 features the three themes of constituency, movement, and discourse across three established analyses of SS (Base Generation in 5.2, Simple Movement in 5.3, Distributed Deletion in 5.4); as well as the sketch of an account in Representation

* This is a revision of my MPhil paper (submitted to the University from Churchill College in June 2015). Its empirical focus is Split Scrambling in Russian. By comparing established Base Generation (Fanselow 1988), Simple Movement (Bošković 2005), and Distributed Deletion (Fanselow & Ćavar 2002) analyses with the sketch of an account in the Representation Theory framework (Williams 2003) of this phenomenon against data old and new, I seek to make a contribution to a relatively understudied area of Russian linguistics and to point out some structural flaws in the Chomskyan edifice without undermining its foundations. Michelle Sheehan supervised me then and Rob Truswell is now my main supervisor; theirs is the credit for all that is good here, and mine the responsibility for the rest.

† To the memory of John Richard Wood (1960-2013).

Theory in 5.5—as at the least an interesting reflection on the outstanding problems in the other three frameworks.

The paper is closed out by section 6, which summarises the results of my researches (6.1) before considering prospects for future development (6.2).

1.2 ‘Thematic Structure’

I take this opportunity to say a few words about the other layers of structure in the present work. Its title, for one thing, is nothing but a promissory note—in that the reference to *The Logical Structure of Linguistic Theory* (Chomsky 1975/1955) is not justified by the scope of the present work, but rather anticipates a longer-term project (which would/will explore the concept of ‘theoretical space’, of the ‘geometrical’ aspect of theories, in something like the exploratory spirit with which Chomsky laid down the ‘logical’ foundations of his own sixty years ago).

There are several ‘directionalities’ underlying the layout of 1.1. In the process of planning sections 3, 4 and 5 I was guided by a heuristic principle of *parallelism*, founded on Culicover & Jackendoff’s (2005) Base-Derivation-Interfaces trio. Although this intra-section organisation is now most clearly on show in section 4, I hope its presence may still be felt elsewhere. The headings of the sections themselves indicate a *progression* from ‘raw facts’ to ‘arising questions’ to ‘proposed answers’. In a similar way, I have seen the work as a sort of chain between the empiricist and rationalist extremes of ‘pure data’ and ‘pure theory’, with intermediary links including—in rough order from ‘left’ to ‘right’—Culicover & Jackendoff’s Base-Derivation-Interfaces trio, as well as its translation equivalents into Chomskyan and other non-Chomskyan languages; Bailyn’s (2011) isomorphic classification Scrambling analyses; and the register of ‘wh-questions’ conventionally confronting the analyst of movement operations (cf. 4.3.1). In an attempt both to draw the links together and avoid possibly unwanted associations, I have coined my own trinity of terms in parallel with Culicover & Jackendoff’s and on an alliterative principle: ‘segmentation’, ‘stratification’, ‘segregation’.

Progression implies a certain teleology, and the targets which should be met en route from one end of the chain to the other are: a quasi-neutral and -comprehensive overview of Russian SS data; a review of the background issues and existing analyses; and the outline of an analysis different from but incorporating the insights of what Culicover & Jackendoff’s call ‘mainstream’ linguistics.

2 FRAMEWORK

2.1 Introduction

In a slight departure from the inductive scheme of 1.1, I now briefly introduce the four frameworks which will make their return in section 5, after my Russian Split Scrambling data and the theoretical issues arising from that empirical base have been laid out in sections 3 and 4 respectively. In order of appearance they are: Fanselow’s Base Generation (2.2); Bošković’s Simple Movement (2.3); Fanselow & Féry’s Distributed Deletion (2.4); and Williams’s Representation Theory (2.5).

2.2 Base Generation

Although it no longer has much currency the ‘Base Generation’ analysis still serves as a point of reference in contemporary debates on how to account for the facts of Split Scrambling. In the present context it is represented by Fanselow (1988), a paper from the height of the GB era. The essence of BG, as suggested by its name, is that ‘free’ word orders attested notably in SS are not derived by the computation of syntax but generated in its base. Specifically, Fanselow suggests that ‘discontinuous NPs’ are underlyingly *two* NPs, where one is in a non-argument position from which it binds the coindexed *pro*-head of the other in an argument position; they are then united at LF. Cf. (1) (Fanselow 1988: 101):

- (1) [Schöne Frauen_i] hat er [keine beautiful-ACC.PL.F woman-ACC.PL have-3SG.PRS he-NOM none-ACC.PL.F *pro*] eingeladen.
pro invite-PP
 “He has not invited any beautiful women.”

2.3 Simple Movement

Bošković’s variation on the Simple Movement theme is chosen to represent a cluster of accounts, partly in the interests of space, but also because his approach not only shares essential features with other members of the SM group but also has some points of interest all its own. Bošković’s empirical focus is still on nominals but closer to the home of this paper, in West Slavic; the main proposal in Bošković (2005) is to parameterise the structure surrounding a noun so that some languages have DP ergo AP-over-NP ergo no LBE/SS (because Adj is not a constituent), whereas others—Russian included—do not have DP ergo NP-over-AP ergo yes LBE/SS (AP, in which Adj now resides separate from NP, being a constituent). The difference in template between the two types of languages is shown in Figure 1.

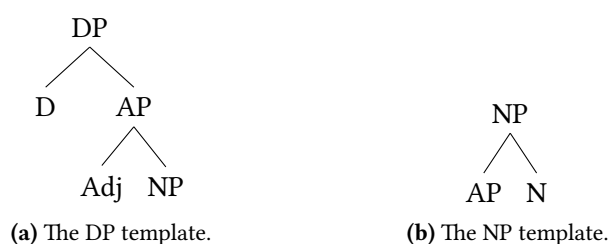


Figure 1 Models of the nominal projection.

2.4 Distributed Deletion

If one wanted to pin the ‘mainstream’ tag on any of the current SS accounts the best choice would probably be ‘Distributed Deletion’. The key reference is Fanselow & Ćavar (2002); note too adaptations to split phrases in Colloquial Russian (Pereltsvaig 2008) as a whole and to preposition doubling in the language more specifically

(Goncharov 2012). Seeking to tow a middle line between Base Generation and Simple Movement, DD conscripts the Copy Theory of movement (Chomsky 1995, Nunes 1995) to allow a single underlying phrase to fulfil two ‘functions’ in the sentence by cloning it from one position into another and then selectively suppressing complementary chunks of phonological material in the two twins. The pre-deletion stage is shown schematically in (43) (Fanselow & Ćavar 2002: 85):

(2) $[[_{XP} a^P [b c]^q] [HP \dots [[_{XP} a^P [b c]^q] [H^q \dots [_{XP} a^P [b c]^q]]]]]$

2.5 Representation Theory

An alternative analysis of Russian SS inherits multiple tensions—directly from the data of section 3 (S-O and IO-DO asymmetries, adjective-preposition-adverb contrasts, ‘inversion’ and ‘height’/distance’, complex constituents and long distances, ‘salience’, ‘significance’, and ‘scalarity’); indirectly from the issues in section 4 (constituency, movement, discourse, with particular reference to landing sites, locality, optionality and gradience); and immediately from the accounts in 2.2-2.4 (mechanical restrictivity vs. permissivity, interdependence of design features, the gap between a framework and analyses constructed in it). What can Representation Theory do to resolve these tensions?

So far the story has been one of back-and-forth between positions describing a fairly narrow circle around the Chomskyan centre of gravity; Representation Theory (Williams 2003) falls beyond that theoretical pale. It is however in close dialogue with MP, and acknowledges loose communion with e.g. Optimality Theory (Prince & Smolensky 2004, Smolensky & Legendre 2006), Lexical-Functional Grammar (Bresnan 2001, Dalrymple 2001), and Marantz’s model of argument structure (Marantz 1984). It is rooted both explicitly and implicitly in its author’s earlier ideas about e.g. autosegmentalism (Williams 1976), rule ordering (Williams 1974), grammatical theory as a whole (van Riemsdijk & Williams 1986), and thematic structure (Williams 1994). And to my mind it also has affinities of ‘letter’ with Tree-Adjoining Grammar (Steedman 1996), Zubizarreta (1998), and Lebeaux (2011); as well as of ‘spirit’ with other monographs such as Collins (1997), Kitahara (1997), Reinhart (1996), Stroik (2009), Richards (2010) concerned like Williams with the desiderata of ‘substantive minimalism’. A diagram of the model, which combines the insights of 5.27 (p. 133) and 9.5 (p. 244) is given in Figure 2.

RT embodies specifically the MP ‘economy imperative’ in an architecture resting on multiple formal levels of representation, or ‘Structures’. Each one is its own sphere (the size of which increases from left to right) of grammatical relations, locality effects, and movement operations. The ‘Representation’ in RT comes from the way in which the Structures are connected, which is by mappings which ‘scramble’ for certain properties, preserving isomorphism between strata as much as possible and standing in parameterisable power relations to each other. Form—incorporating syntax, morphology, and prosody into Case, Predicate, Surface, and Accent Structures—and meaning—Thematic, [unnamed], Quantification, and Focus Structures—are placed in parallel, with (mis)mappings obtaining both left-right and top-bottom as depicted in the figure.

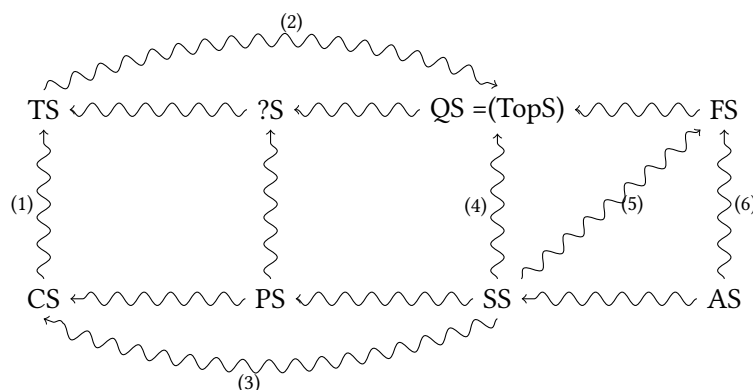


Figure 2 The RT Architecture

(1)=movement for Case; (2)&(5)=scrambling for focus; (3)=scrambling for scope;
(4)=in-situ quantifiers; (6)=in-situ nonfinal focus.

This overview has to be abstracted out of the organisation of the monograph ‘on the ground’, which is driven by the successive coverage of selected topics from Generative Grammar reworked into the system outlined thus far. The exposition is framed on either side by introductory sections on ‘Economy as Shape Representation’, ‘Topic and Focus in Representation Theory’, and ‘Embedding’ on the one hand and on the other by more speculative ones about ‘X-Bar Theory and Clause Structure’, ‘Inflectional Morphology’, and ‘Semantics in Representation Theory’. The core third of the book ‘parameterises’ the notions of ‘Anaphora’, ‘A/A’ relations’, and ‘Superiority’ to the (in principle open) set of RT Structures. RT has not made a particularly noticeable stir in the MP ‘mainstream’, although it evoked some enthusiasm in reviewers (Csirmaz 2004, Hornstein & Nevins 2005).

3 DATA

3.1 Introduction

In this section I present the bulk of the data I have collected on Russian Split Scrambling—though important example islands are located in 4.3.2 and 4.3.3. I endeavour to frame the facts as atheoretically as possible, though both the terminology—largely that of descriptivist grammarians—I bring to the data and the generalisations I formulate on the basis of it inevitably prejudice the neutrality of the exposition. Nevertheless. In Section 3.2 I build up the complexity of the examples from the minimal types of construction which make SS visible, intransitives (3.2.2) by the addition of places in the predicate, through mono- (3.2.1) to ditransitives (3.2.3)—finding subject-object and direct-indirect object asymmetries. Section 3.3 is something of an interlude, in which the permutational possibilities of ‘inversion’ and ‘height’/‘distance’ (3.3.1)—where more or less all logical possibilities are in—as well as ‘multiplicity’—where more or less all logical possibilities are out—are explained and explored using the core Adj-N SS cases of 3.2 as the point of departure. In

Section 3.4 I move to the categorial level of detail in clause structure, starting with prepositions (3.4.1) and adverbs (3.4.2), which have been far from neglected by scholars; both classes show a certain dependency on and preference for Adj/N in their NP complements. I move to adjective-like words which are more or less substitutable for adjectives in 3.4.3, and then other categories (3.4.4) which can inhabit the clause but are by and large not SS-friendly; all of which, probably due to the fact that their behaviour is more specific to Russian, have to my knowledge received fairly little attention. In Section 3.5, finally, the embedding of nominal (3.5.1) and clausal (3.5.2) constituents in each other and the adjectival projections is examined and found to be almost entirely excluded as an SS possibility. Section 3.6 takes a step back to seek some method of meaning in the madness of surface word order options, narrowing down from the neutrality of SVO (3.6.1) to specifics of ‘style’ with S-final and preposition-doubling cases (3.6.2), ‘sociolinguistics’ (3.6.3), ‘salience’ (3.6.4) from left to right, ‘significance’ (3.6.5) at the left and right, and ‘scalarity’ (3.6.6) of both. Section 3.7 concludes the section.

A few notes on notation. From Pereltsvaig (2008) I borrow the device of using Roman letters to refer to parts of the target of SS, going in alphabetical order from the leftmost fragment in the surface string; I replace the letters where useful where useful with abbreviations for parts of speech (N, V, Adj, Adv, P, etc.) Examples are marked ‘SVO’ etc. for the purposes of classification rather than as a claim about derivation, with B appearing in the position indicated by the letter of the ‘target’ (S,V,O ...) and A somewhere to the left. Thus, ‘New president interesting reads book’ is simultaneously SAVB, SAdjVN, SVO. Finally, I used the following standard syntactic symbols: nothing for full acceptability, ? for slight and ?? for greater ‘markedness’, ??? for marginality, * for ungrammaticality; % indicates significant disagreement among speaker judgements, with the poles separated by /. Other concepts will be explained along the way.

3.2 Predication

3.2.1 Monotransitives

Although monotransitives are not the minimal case for predicate places, I begin with them here to continue on from the introductory examples of Scrambling in (3a):

- (3) a. *Ivan čitaet interesnuju knigu.* (SVO)
 Ivan-NOM read-3SG.IPFV interesting-ACC.SG.F book-ACC.SG
 “Ivan is reading a book.”
- b. *Ivan interesnuju knigu čitaet.* (SOV)
- c. *?Čitaet Ivan interesnuju knigu.* (VSO)
- d. *?Čitaet interesnuju knigu Ivan.* (VOS)
- e. *Interesnuju knigu Ivan čitaet.* (OSV)
- f. *Interesnuju knigu čitaet Ivan.* (OVS)

These are canonical because they correspond to the three-part template used in the typological study of word order. First let us see whether it is possible to apply Split Scrambling to S:

- (4) a. *Novyj čitaet prezident interesnuju knigu.
 New-NOM.SG.M president-NOM.SG interesting-ACC.SG.F read-3SG.IPFV
 book-ACC.SG
 “The new president reads an interesting book.”
- b. *Čitaet novyj interesnuju knigu prezident.
- c. *Novyj interesnuju knigu prezident čitaet.
- d. %??/* Interesnuju knigu novyj čitaet prezident.

Only OAdjVN (4d)—and its inverted counterpart, ONVAdj, command a consensus of even marginal acceptability, and then only in stylistically marked contexts—described as e.g. ‘poetic’ or ‘folksy’ (cf. 3.6.2). Its exceptionality would seem to be supported by the fact that (4a), in which Adj is also separated from N by V alone, is unanimously rejected. One might then conclude from (4) that SS from the subject is blocked to all intents and purposes. This result seems to come under the Condition on Extraction Domains (Huang 1982)—bolstering it, at least in the prism of SS, against Stepanov (2007)’s claims about its inefficacy in Russian—and more specifically Uriagereka (1999)’s proposal about extraction from true complements; it is relevant to 4.3.2.

In order to regain acceptability one needs to turn to O as target:

- (5) a. ?Novyj prezident interesnuju čitaet knigu.
 b. ?Interesnuju novyj prezident knigu čitaet.
 c. ??Čitaet interesnuju novyj prezident knigu.
 d. ??Interesnuju čitaet knigu novyj prezident.

All of the sentences in (5) are less neutral than (3a), but the somewhat degraded acceptability of V-initial (5c) and (5d) from S-initial (5a) and (5b) replicates the pattern in the unsplit VSO and VOS word orders of (3).

3.2.2 Intransitives

I now take a step back to the ‘logical’ starting point of this section, i.e. to the somewhat amorphous family of one-place predicates. I begin with the ‘core’ intransitive cases of unaccusatives (6a) and unergatives (6b), which are distinguished by having no logical subject and no logical object respectively.

- (6) a. ?Novyj umer prezident.
 New-NOM.SG.M die-M.PFV.PST president-NOM.SG
 “The new president died.”
- b. %?/* Novyj razgovarivaet prezident.
 New-NOM.SG.M talk-3SG.IPFV president-NOM.SG
 “The new president talks.”

Although there is disagreement between speakers about (6b), in both cases the subject seems more amenable to SS than the transitive (4). Passives (7a,b)—which have received many different analyses in the literature but which in descriptive terms see the ‘logical object’ converted into the ‘grammatical subject’—and raising predicates (7c) are further testing grounds:

- (7) a. ??*Novyj izbran prezident.*
 New-NOM.SG.M elect-PPPP.M president-NOM.SG
 “A new president has been elected.”
- b. ??*Novyj izbralsja prezident.*
 New-NOM.SG.M elect-M.PFV.PST.REFL president-NOM.SG
 “A new president has been elected.”
- c. %???/**Podozritel’nym novyj kažetsja prezident.*
 Suspicious-INSTR.SG.M new-NOM.SG.M seem-3SG.IPFV president-NOM.SG
 “The new president seems suspicious.”

SS seems to apply quite easily to the surface subject of the passive. Less so with raising: in fact, the OVS-like order given in (7c)—along with its inverted variant—is, like its counterpart in (4), more or less the only even marginally acceptable member of the set. Both passives and raising predicates have been analysed as unaccusatives; copulas (8), conversely, bear some resemblance to the raising construction but submit more easily to SS:

- (8) ??*On novym byl prezidentom.*
 He-NOM.SG new-INSTR.SG.M be-M.PFV.PST president-INSTR.SG
 “He was the new president.”

Here again SS seems able to target the subject. My final example in this subsection comes from the sub-family of impersonals:

- (9) ?*Včera sil’nyj šel dožd’.*
 Yesterday strong-NOM.SG.M go-M.IPFV.PST.UD rain-NOM.SG
 “There was heavy rain yesterday.”

The ‘weather impersonal’ in (9) once more confirms the general trend: the ‘non-agentive’ subjects of intransitives are by and large susceptible to SS, whereas agentive monotransitive ones are not. SS thus seems to be sensitive to the distinction between ‘syntactic’ and ‘thematic’ subjecthood.

3.2.3 Ditransitives

Although Russian grammarians are fond of citing e.g. *prodavat’/prodat’* ‘to sell’ as a four-place predicate (involving who did the selling, what was sold, who it was sold to, and for how much) in this section I will not venture beyond the class of ditransitives. Even they are only included to compare the compatibility with SS of the ‘direct’ and ‘indirect’ objects involved. As the example space is exponentially larger than in 3.2.1 (with $4! = 24$ ‘baseline’ orders from SVIODO against the 6 of SVO in (3), only a few ‘simple’, ‘SVIODO’ instances are given here. First, with DO as target:

- (10) a. %??/* *Ivan dal interesnuju Petru knigu.*
 Ivan-NOM.SG give-M.PFV.PST interesting-ACC.SG.F Petr-DAT.SG
 book-ACC.SG
 “Ivan gave Petr an interesting book.”
 b. ?? *Ivan interesnuju dal Petru knigu.*
 c. ? *Interesnuju Ivan dal Petru knigu.*

With the increased complexity of the ditransitive construction in (10) seem to come a general degradation compared to (5) and a positive correlation between acceptability and height/distance. Taking IO as target we get:

- (11) a. * *Ivan novomu dal prezidentu knigu.*
 Ivan-NOM.SG new-DAT.SG.M give-M.PFV.PST president-DAT.SG
 book-ACC.SG
 “Ivan gave the new president a book.”
 b. * *Novomu Ivan dal prezidentu knigu.*

The contrast between the (gradient) acceptability in (10) and the (categorical) exclusion of (11) indicates that there is an asymmetry, at least with ditransitives like *davat’/dat’*, between the susceptibility of DO and IO to SS—in parallel to that between the subjects of (4) and (6a)–(9). Neither of these patterns has received much concentrated coverage in the SS literature; cf. Fillmore (1965) for a pioneering analysis, Emonds & Whitney (2006) for review, and 5.5.2.

3.3 Permutation

3.3.1 Inversion and Distance/Height

Having in 3.2 imposed the layer of variation from predicate structure on top of the Scrambling options, I begin this section by introducing some more parameters. Again from Pereltsvaig (2008) I take the distinction between ‘simple’ (or ‘pull’, as per Fanselow & Ćavar (2002)) and ‘inverted’ SS, in which A and B surface in canonical (e.g. Adj[...]N) and counter-canonical (e.g. N[...]Adj) order, respectively. ‘Height’ (L/M/H) and ‘distance’ (x ‘steps’) refer respectively to the ‘leftness’ of A and its separation from B. Applying ‘inversion’ to the monotransitive object (o) yields the examples in (12):

- (12) a. ? *Novyj prezident knigu čitaet interesnuju.*
 b. ? *Knigu novyj prezident interesnuju čitaet.*
 c. ?? *Čitaet knigu novyj prezident interesnuju.*
 d. ?? *Knigu čitaet interesnuju novyj prezident.*

As shown by the pattern of judgement marks, speakers were more or less as happy with inverted as with simple/pull orders; there is again a decline in acceptability from the S- (12a)(12b) to V-initial (12c)(12d) templates. This suggests that if there is a definite difference in derivation—as posited by Fanselow & Ćavar (2002), and by

Pereltsvaig (2008) specifically with respect to Russian—it does not have a significant effect on the acceptability of the results.

Staying with monotransitives but concentrating on ‘height’/‘distance’ yields (13)-(16):

(13) a. ?*Novyj prezident interesnuju čitaet knigu.*

b. ?*Interesnuju novyj prezident čitaet knigu.*

(14) ?*Interesnuju novyj prezident knigu čitaet.*

(15) a. ??*Čitaet interesnuju novyj prezident knigu.*

b. ??*Interesnuju čitaet novyj prezident knigu.*

(16) ??*Interesnuju čitaet knigu novyj prezident.*

The same pattern obtains in (13)-(16) as in (12), with V-initial-based orders (15)(16) slightly degraded as per (3)(5)(10) before them. Although here is the main place in this section where they are exemplified, inversion and distance/height should be seen as background options everywhere else. I return in 3.6, and again in 4.4, to the motivation behind the variability they introduce—which from a purely formal perspective seems arbitrary in degree and puzzling in the detail of what positions are targeted. The ‘position problem’ has had low uptake in the SS literature—Pereltsvaig (2008) gives a few examples to contradict Sekerina (1997)’s Periphery Constraint—but will be a theme of efi/m/l_s and section 5. It is compounded by the ability of traditionally diagnostic elements such as adverbs (cf. 3.4.2) to scramble. For the moment, the monotransitive SS possibilities considered so far may be captured in a descriptive linear template like (17):

(17) (Adj/N) S/V (Adj/N) V/S (Adj/N)

3.3.2 Multiplicity

The topic of ‘multiplicity’ is picked up by at least Sekerina (1997) and Pereltsvaig (2008) in the Russian context alone. There are two aspects to the issue. The first—whether a single constituent can be targeted more than once by SS—will come into view in the context of prepositions (20d) and phrase-level adverbs (21d). An opportunity I do not exploit here is the logical step from (21d) to at least predicate-if not sentence-level adverbs. I do however explore the avenue opened up by the recursivity of the *adjectival* node in NP. Pereltsvaig claims that multiple SS saves an otherwise ungrammatical construction—at least from the very brink; as shown in (18), my informants implicitly agree with Sekerina’s ‘Multiple Modifier Constraint’.

(18) a. **Novyj prezident tolstie čitaet interesnie knigi.*
 New-NOM.SG.M president-NOM.SG thick-ACC.PL read-3SG.IPFV
 interesting-ACC.PL books-ACC.PL

“The new president reads thick interesting books.”

b. **Novyj prezident interesnie čitaet tolstie knigi.*

c. **Novyj prezident tolstie interesnie čitaet knigi.*

- d. **tolstie novyj prezident interesnie čitaet knigi.*
 e. **Interesnie novyj prezident tolstie knigi čitaet.*

The other logical possibility with multiplicity is the application of SS to multiple constituents in the same sentence. Here Pereltsvaig upholds Sekerina's 'One-Split-per-Clause Constraint', and my data (19) supports it too:

- (19) a. **Interesnuju Ivan knigu novomu*
 Interesting-ACC.SG.F Ivan-NOM.SG book-ACC.SG new-DAT.SG.M
dal drugu.
 give-M.PFV.PST friend-DAT.SG
 "Ivan gave an interesting book to his new friend."
 b. **Ivan novomu dal interesnuju drugu knigu.*

It is unclear whether the constraint is a matter of competence or performance; there may be a confound in such examples with the immunity of IO to SS as diagnosed in 3.2.3.

3.4 Categorisation

3.4.1 Prepositions

3.2 made manifest contrasts in susceptibility to SS on a 'relational' basis, i.e. between S and O as well as between their subsets. In 'categorial' terms, however, the targets of (5)-(19) have been uniformly composed of Adj and N. Such structures are the mainstay of the analyses in section 5, but by no means exhaust the SS data space. I therefore turn to other elements that may enter into the composition of O, starting with prepositions:

- (20) a. ??*Novyj prezident čerez bol'suju proxodit*
 New-NOM.SG.M president-NOM.SG in big-PREP.SG.M pass-3SG.IPFV
komnatu.
 hall-PREP.SG
 "The new president passes through the large room."
 b. ??*Novyj prezident bol'suju proxodit čerez komnatu.*
 c. ???*Novyj prezident čerez bol'suju proxodit čerez komnatu.*
 d. **Novyj prezident čerez proxodit bol'suju komnatu.*

The results of baseline permutations from the SVO of (20) and of inversion from its 'simple' SS are both less acceptable overall and more changeable in acceptability than those of (5). The case of prepositions is further complicated by the fact that P is another variable in the SS equation: it can appear in A (20a) or B (20b) more or less interchangeably; in both (20c), though this option is stylistically marked to varying degrees (cf. 3.6.2); but not in isolation (20d). These data are in line with [Goncharov \(2012\)](#), whose interest is in P-doubling cases like (20c) and whose theoretical agenda I will not expound upon; and, in (20d), with [Sekerina's \(1997\)](#) Preposition-First Constraint.

3.4.2 Adverbs

A similar story may be told for phrase-level adverbs in NP, which [Pereltsvaig \(2008\)](#) exemplifies only in passing. On further inspection as in (21) Adv appears preferentially with Adj rather than N (21a)(21b), doubling is excluded (21c) outright, and isolation is a marginal possibility (21d):

- (21) a. ??*Novyj prezident očen' interesnuju čitaet knigu.*
 New-NOM.SG.M president-NOM.SG very interesting-ACC.SG.F
 read-3SG.IPFV book-ACC.SG
 “The new president reads a very interesting book.”
 b. ???*Novyj prezident interesnuju čitaet očen' knigu.*
 c. **Novyj prezident očen' interesnuju čitaet očen' knigu.*
 d. ???*Novyj prezident očen' čitaet interesnuju knigu.*

Predicate-level adverbials (22) do not appear in [Pereltsvaig \(2008\)](#), and seem somewhat freer:

- (22) ?*Očen' Ivan medlenno čitaet knigu.*
 Very Ivan-NOM.SG slowly read-3SG.IPFV book-ACC.SG
 “Ivan reads the book very slowly.”

Most sentence-level adverbs are single, unmodifiable words in Russian—meaning that it is difficult to complete the ascent up the adverbial ladder. A general conclusion from 3.4.1 and 3.4.2 might be that prepositions and adverbs are less susceptible to SS than adjectives. Possible explanations offer themselves from morphology (both sets being invariable, unlike agreeing adjectives) and phonology (with some members also being prosodically dependent on their ‘modifiee’); a different cause, and one which makes these cases key to the argumentation of the paper, is identified in 4.2.2 and treated further in section 5.

3.4.3 Adjective-like categories

In Russian a number of categories which traditionally belong to different ‘parts of speech’ behave in various ways and degrees like the prototypical adjectives used in 3.2-3.3. There are ‘determiners’ and ‘quantifiers’ like *vsjakij*, *ljuboj*, and *každyj*, which between them cover the territory of ‘any’ and ‘every’; *nekotoryj*, *kakoj-to*, and *kakoj-nibud'*, which range across the shades of ‘some’; *nikakoj* ‘none’; *ves'* ‘whole’ and *vse* ‘all’; *kakoj* ‘which, what sort of’ and *takoj* ‘such a’; *tot* ‘that, the’ and *ëtot* ‘this, the’. These are uniformly adjectival in agreeing with their nouns for number, case, and gender (three ways in the singular and one in the plural). Of the ‘possessive pronouns’, the first- and second-person forms of which—sing. *moj* and *tvoj*, pl. *naš* and *vaš*—are in the same morphological boat, whereas in the third person—m./n. *ego*, f. *eë*, pl. *ix*—there is invariability. The ‘numerals’ are a different mixed bag: *odin* ‘1’ is morphologically adjectival; the rest are nominal with varying degrees of paradigmatic eccentricity, governing various cases and numbers in the nominative and accusative but agreeing in case in oblique contexts.

The point of this enumeration was in the parallel and divergent behaviour of its members in SS constructions vis-à-vis adjectives. Given that they cannot be modified—apart from by each other in some combinations—direct interchangeability with the adjectives in (21) is impossible. As for nominal (5) and prepositional (20), the exchange is agreed to be by and large permissible, though there are concomitant changes in meaning which go beyond the aforementioned ‘nuances’ of 3.6.

3.4.4 More categories

Another path to pursue in the search for ever more data space is population of the clause with material besides the nucleus of predicate, arguments, and adjuncts built up over the course of this section. Russian is rich in interjections and particles, but with them—including the negative *ne* (and by extension Negative Polarity Item constructions, which incorporate it) and the conditional *by*—the confound of morphological and phonological ‘parasitism’ from 3.4.2 is critical. SS is therefore either impossible (in most cases) or undetectable, if the parasitism is also semantic. One last case is the future tense of the imperfective, which is composed from forms of the defective copula *byt’* plus the infinitive:

- (23) %??*/* *Novyj prezident budet interesnuju čitat’ knigu.*
 New-NOM.SG.M president-NOM.SG be-3SG.PFV interesting-ACC.SG.F
 read-3SG.IPFV book-ACC.SG
 “The new president will read an interesting book.”

(23) shows that, for some speakers at least the imperfective future displays syntactic as well as morphological analyticity.

3.5 Complexification

3.5.1 Nominal

Most of the constructions in this section have involved a ‘nominal juncture’ of some kind—with verbs in 3.2, prepositions in 3.4.1, adverbs in 3.4.2, and with adjectives more or less everywhere. Two types which have not appeared so far are commonly set aside from the rest as ‘complex’. One configuration the epithet is applied to is the embedding of a nominal as the complement of an adjective, as in (24):

- (24) a. ??? *Ivan novomu veren prezidentu.*
 Ivan-NOM.SG new-DAT.SG.M loyal-M president-DAT.SG
 “Ivan is loyal to the new president.”
- b. * *Novyj prezident v sobstvennoj uveren politike.*
 New-NOM.SG.M president-NOM.SG in own-PREP.SG.F sure-M
 policy-PREP.SG
 “The new president is sure of his own policy.”

SS from a complex adjectival in which the nominal complement is ‘bare’ (24a) appears marginally permissible, whereas if it is encased by a preposition SS is blocked. The second type is doubly nominal, and illustrated in (25):

- (25) a. **Ivan čitaet novogo knigu prezidenta.*
 Ivan-NOM.SG read-3SG.IPFV new-GEN.SG.M book-ACC.SG president-GEN.SG
 “Ivan sees the the new president’s book.”
- b. **Novyj prezident projavljaet v sobstvennoj uverenost’ politike.*
 New-NOM.SG.M president-NOM.SG display-3SG.IPFV in own-PREP.SG.F
 certainty-ACC.SG policy-PREP.SG
 “The new president displays certainty in his own policy.”

In complex nominals, it seems, SS from both NP and PP complements is bad. These contrasts in SS susceptibility between NP complements to verbs and to other categories anticipate the topic of ‘locality’ 4.3.2.

3.5.2 Clausal

Adjectival and nominal projections may also be ‘complexified’ by the embedding of a clause:

- (26) a. **Novyj prezident uveren, interesnuju čto čitaet knigu.*
 New-NOM.SG.M president-NOM.SG sure-M interesting-ACC.SG.F that
 read-3SG.IPFV book-ACC.SG
 “The new president is sure that he is reading an interesting book.”
- b. **Menja ne interesuet fakt, interesnuju čto novyj prezident čitaet knigu.*
 I-ACC.SG not interest-3SG.IPFV fact-NOM.SG interesting-ACC.SG.F that
 new-NOM.SG.M president-NOM.SG read-3SG.IPFV book-ACC.SG
 “I am not interested in the fact that the new president is reading an interesting book.”

In (26) the contrast between adjectivals and nominals disappears: SS from the clausal complement of either is excluded.

Sentences can in turn be ‘complex’ by virtue of clauses being embedded within each other. Unlike their adjectival and nominal counterparts, complex sentences have received some dedicated attention: [Pereltsvaig \(2008\)](#) contradicts [Sekerina’s \(1997\)](#) ‘Short-Distance Constraint’ about the clause-boundedness of SS in Russian; the next couple of example sets collate my own results from a range of ‘experimental conditions’, starting with the indicative-subjunctive juncture in its ‘simple’, distance/height permutations with an SVO base in (27):

- (27) a. %???/**Ivan* *xočet,* *interesnuju* *čtoby*
 Ivan-NOM.SG think-3SG.IPFV interesting-ACC.SG.F that
Marija *čitala* *knigu.*
 Marija-NOM.SG read-F.PFV.PST book-ACC.SG
 “Ivan wants Marija to read an interesting book.”
- b. %???/**Ivan interesnuju xočet, čtoby Marija čitala knigu.*
- c. %???/**Interesnuju Ivan xočet, čtoby Marija čitala knigu.*

Such sentences were in fact rejected by the majority of my speakers. Those few who accepted them marginally associated them with a certain—‘uneducated’, ‘illiterate’—sort of speaker and/or speech situation—‘casual’, ‘careless’ (cf. 3.6.3).

There are as many clausal junctures as there are types of clauses and ways of combining them. Core cases include indicative-indicative

- (28) **Ivan* *dumaet,* *interesnuju* *čto* *Marija*
 Ivan-NOM.SG think-3SG.IPFV interesting-ACC.SG.F that Marija-NOM.SG
čitaet *knigu.*
 read-3SG.IPFV book-ACC.SG
 “Ivan thinks that Marija is reading an interesting book.”

as well as indicative-relative

- (29) **Ivan* *vidit* *devušku,* *interesnuju* *kotoraja*
 Ivan-NOM.SG see-3SG.IPFV girl-ACC.SG interesting-ACC.SG.F which-NOM.SG.F
čitaet *knigu.*
 read-3SG.IPFV book-ACC.SG
 “Ivan sees a girl who is reading an interesting book.”

and indicative-interrogative

- (30) **Ivan* *dumaet,* *interesnuju* *kto* *čitaet*
 Ivan-NOM.SG think-3SG.IPFV interesting-ACC.SG.F who read-3SG.IPFV
knigu?
 book-ACC.SG
 “Ivan thinks who is reading an interesting book?”

all of which were unanimously rejected. ‘Long-distance’ (i.e. extra-clausal) SS thus seems to be a very marginal phenomenon, though Pereltsvaig still has a point. The ‘discontinuity’ in acceptability when the clause boundary is crossed, like the splitting of complex adjectivals and nominals with nominal complements in 3.5.1, hints at some constraint beyond the descriptive scope of this section; again, cf. 4.3.2.

3.6 Interpretation

3.6.1 Neutrality

The most basic division that emerges from my data separates orders which speakers feel to be ‘neutral’ for context of utterance and speech register from those that are somehow and somewhat ‘marked’. An overwhelming consensus isolates the unsplit SVO configuration (3a) as the most neutral of all S,V,O permutations—in accord

with almost all the literature from [Isačenko \(1966\)](#) to [Bailyn \(2011\)](#) apart from [King \(1993\)](#), who favours VSO.

3.6.2 *Style*

In this and the next subsection I note some features that were prominent in speakers' commentary on my example sentences, but which do not play any further role in the exposition and have been little remarked upon apart from in 'traditional' Russian texts (e.g. [Sirotinina 1965, 1970](#), [Zemskaja 1973](#), [Krylova & Xavronina 1984](#), [Švedova 1980](#))—because they fall outside the pale of pragmatics as defined in the Anglo-American (as opposed to the Continental) tradition. The first is 'style', for which two broad categories were invoked—'folksy' and 'poetic'. Into 'folksy' falls above all the preposition-doubling pattern of (20c); into the 'poetic', primarily the V-initial models of (5c) and (5d). In such cases the stylistic effect often seemed so strong that unacceptability would result from removing them from the relevant context.

3.6.3 *Sociolinguistics*

A thin line separates 'style' from 'sociolinguistics', but in the second class I place examples which were attributed to 'uneducated', 'illiterate' speakers and 'careless' registers—i.e. accorded negative prestige. Here the relevant examples are those which some speakers—who one might label, in a reverse sociolinguistic judgement, 'conservative'—simply rejected: above all, the scrambling of an adverb without the adjective it modifies (21d); and indicative-subjunctive clausal junctures (27). In light of the remaining subsections in 3.6 and of 4.3.2 as well as 4.4.1, one might hypothesise that the hypothetical speakers who allow such examples have a grammar which (systematically, not 'carelessly') gives higher priority to discourse factors than the 'norm'.

3.6.4 *Saliency*

By 'saliency' I mean the prominence of a sentence constituent in virtue of how new or old it is in the context of utterance. All other things being equal (which they are often not—see 3.6.5) moving from left to right in the range of configurational positions seems to correlate negatively with this property. To return to the canonical order of (3a), the subject is supposed to be the an entity more familiar in the circumstances than the contents of the predicate. The default nature of that state of affairs explains SVO's 'neutral' status, which is confirmed by seeing what order of words comes out when the equivalent of the 'zero-theme' question "What happened" is posed.

3.6.5 *Significance*

There was a caveat in 3.6.4 about all other things being equal in the context of utterance. The countervailing force hinted at there I refer to here as 'significance', a prominence which derives not from *familiarity* but from *importance* in the context of utterance. It is "countervailing" because it appears to take priority over

‘salience’ when the two compete for ‘pole position’ at the beginning of the sentence; in sentence-final position, significance seems to have an undisputed monopoly—although there may be a distinction between the two structural extremes, with a shade of contrastivity—‘x as opposed to other possible members of a set’—adhering to the left whereas the right is more ‘simply’ significant. Both ‘significance’ and ‘salience’ have been widely covered in the literature under a variety of names; cf. 4.4.1.

3.6.6 *Scalarity*

Implicit in 3.6.4 and 3.6.5 is the notion of ‘scalarity’, although it is more prevalent for ‘significance’ than for ‘salience’—an appearance possibly not unconnected with its double positional association. In the monotransitive object-splitting sentences of (13)-(16), for example, speakers say that there is more ‘emphasis’ on ‘A’ the further it is removed from ‘B’. On the basis of such a general formulation it is difficult to say whether this effect of scalar significance is produced by the ‘height’ of A in the sentence or rather its ‘distance’ from B (to employ the key terms of 3.3.1). However, when pressed speakers seem to lean on the side of ‘distance’ rather than ‘height’ alone. To conclude this section, I present a rough meaning model along the lines of (17):

(31) (Focus_{contrastive}) ... (Topic) (... (Focus_{medium})^{*} ... (Focus_{simple}))_{non-Topic}

3.7 *Conclusion*

In this section I have exercised my knowledge of Russian syntax and my combinatoric capabilities in order to present a picture of Russian Split Scrambling which can both be independent of but also translate into the remainder of the paper. In the process some ends have been left untied, particularly with respect to the *restrictions* on what otherwise seems like an operation which ‘gets its way’ most of the time, thus generating a bewilderingly large data space to explore and map. Particular points on which discussion has been postponed until Chapter 4 include subjecthood (3.2.1, 3.2.2) and the variegated embedding phenomena of 3.5, which together form a background to the treatment of locality in 4.3.2; the range of ‘landing sites’ which seem to be targeted by SS, an issue prominent in all the sections of the section to varying degrees but maybe particularly so in 3.3.1, and which relates to the A/A’ distinction of 4.3.3; the set of ‘extraction targets’ for SS, which speaks to the concept of constituency as developed over the course of 4.2; and, last but not least, the nuances of meaning that accompany the permutations of form and were sketched in improvised terms close to those used by my speakers in 3.6.4-3.6.6 before being fitted into a ready-made framework in 4.4. The references given to passages in the section 4 lead in turn to more distant ones in section 5, where empirical observation and rationalist imagination will meet in theoretical analysis and the data mentioned here will be of most moment. One link in ‘the chain’ at a time, however.

4 ISSUES

4.1 Introduction

This section is central to the paper—ideologically as well as structurally, because it is the forum for the issues which emerge from the data of 3 to confront the analyses of 5 and vice versa; the home of many of the intermediary links in the chain image of 3.7. It is highly referential, in that it both points in many places to either the previous or the next section; and presents issues in the context of them rather than in the abstract. Section 4.2 is where topics in the segmentation of strings are addressed: Subsection 4.2.1 starts off with constituency as a concept; 4.2.2 turns to its instantiations in the Phrase Structure framework(s); and 4.2.3 pauses over the DP Hypothesis, a milestone in the history of PS. The emergent picture is of a particular relation becoming ontologised into the basis for syntactic derivations. Section 4.3 is typological in spirit and movement-orientated in subject: Subsection 4.3.1 is concerned with the notion of movement itself and the broad classification of movement operations by ‘what’, ‘whence’, ‘whither’; 4.3.2 is expands the ‘whence’ question into locality, with Russian SS showing a Strong-Weak Island split; and 4.3.3 is devoted to diagnostics of the A/A’ variety, on which SS again gives mixed results. The emergent picture is of quite specific symptoms being ontologised into diagnostics to provide labels for unknown and uncomprehended phenomena. Section 4.4 is somewhat more miscellaneous, though it gravitates around meaning: in Subsection 4.4.1 the formalisation of the observations in 3.6 is undertaken; in 4.4.2, the term ‘interface’ itself comes under scrutiny; and in 4.4.3 the related properties of ‘optionality’ and ‘gradience’ are fleshed out. The emergent picture is of discourse being ontologised into a modular, post-syntactic interface which makes discourse-configurationality ‘a thing’. Section 4.5 is the conclusion.

4.2 Constituency

4.2.1 Foundations

At various points in section 3 the question was what can be targeted by Split Scrambling in Russian—which, allowing for confounding semantic, syntactic, morphological, and phonological factors, seemed to be quite a range of words in the sentence. The point of the exercise was to see what perspective SS can give us on the enterprise of segmenting strings in linguistic representations, a task which has in Chomskyan Generativism been carried out with the help of the concept of constituency. Constituency itself has been put on various footings, the general gist of which might be summarised as a drive towards ‘uniformity’, to emphasise Culicover & Jackendoff’s (2005) term—from the operationalisation of ‘parts of speech’ into Phrase Structure centred on major ‘lexical’ categories, to the increased incorporation of ‘functional’ classes into the universal schema of X-Bar Theory, to the ‘simple’ merging and labelling of Bare Phrase Structure.

More particulars will be given in 4.2.2, and a special section of the story will be reopened in 4.2.3. For the moment, however, a pause to revisit the foundations of

the framework. Unlike Dependency Grammar (Tesnière 1959), which was contemporary to *Syntactic Structures* (Chomsky 2002/1957) and intuits various connections between elements, the only principle recognised by the Immediate Constituent approach and its Chomskyan offspring is distributionalism. That monogamous commitment entails that if two elements pattern the same then they belong to the same category, and from there constituents of varying sizes are built up hierarchically on the same principle. The devices deployed in deciding whether substrings ‘pattern the same’ are in operational terms substitutional, whether via like for like/pro-forms or via co-ordination and displacement. A question then is: does SS respect constituency as defined by the original ‘tests’? The (negative) answer is in the ‘Split’ of SS as well as in the data of section 3; and the (involved) reason in 4.2.2.

4.2.2 *Phrase Structure*

An important part of filling out the constituency narrative begun in 4.2.1 is recognising that it is in an important sense not the whole story. From the beginnings of Chomskyanism, constituency as represented by Phrase Structure was not intended to account for all the distributional facts pertaining to the composition of strings. It was designed to encompass a subset of sentences called the Kernel; the rest were the responsibility of a roster of operations known as Transformations. Although the distribution of labour has shifted over the course of time and Transformations have morphed into Move-alpha then Internal Merge, the basic division has remained in force and been extended to Universal Grammar. Constituency congealed into ‘configurationality’.

It is in that climate that ‘Scrambling’ and a fortiori its ‘Split’ subtype have made names for themselves as problematic phenomena. In the present context the salient point is that SS does not conform to the X-Bar template on the one hand—so how can it be assimilated into the system?—and on the other is derived with relative ease in BPS—so why is it limited to languages like Russian? Indeed, Russian is maybe a particularly difficult case because it seems to present conflicting evidence on constituency. Bailyn (2011) is at pains to prove that the major projections of the Chomskyan inventory—DP, VP, CP—display signs of configurationality according to traditional diagnostics. Yet the data in section 3 demonstrate not only that those configurations can be disturbed by SS, but—e.g. in the prepositional (20) and adverbial (21)(22) cases, as well as simple/inverted variants (12)—that it can work in multiple as well as mysterious ways. The pattern just observed—of SS decidedly not fitting the mould but also not being entirely beyond the pale—will become something of a motif in this section, and be of great importance in 5.

4.2.3 *DP Hypothesis*

Speaking of sections: the ‘special’ one promised in 4.2.1 specifically relates more forward to the analysis of 5.3 than backward to the data in 3.2-3.5, though it is ultimately relevant to almost all the examples of Russian Split Scrambling. The section concerns what has come to be known as the DP Hypothesis, which was not only proposed but also justified and extended by Abney in his dissertation (Abney

1987). What he did to give the work its title was to draw parallels between the sentence and the nominal projection, especially along the lines of *Caesar destroyed Rome—Caesar destroying Rome—Caesar’s destroying of Rome—Caesar’s destruction of Rome*. Such signs that NP has the ‘sentential aspect’ of embedding a proposition went some way to explaining its GB barrierhood and justifying its MP phasehood.

More of Barriers, Phases, and difficulties from SS in 4.4. For present purposes, however, a more important consequence of the sentential-nominal parallelism is the extension of the X-Bar template it was used to license. It is the *DP* Hypothesis because Abney proposed that ‘determiners’—articles, quantifiers, possessive pronouns, etc.—should be taken to head their own projection with an NP as its complement rather than as specifiers of the nominal projection as thitherto. A special categorial status was thus formalised for the ‘functional class’ of determiners, thus predicting that there should be asymmetries between them and adjectives—which had occupied the same slot in the NP system, albeit in phrases of their own rather than in isolation. A further step that Abney took in that regard was to make AdjPs complements in Quantificational Phrases. The first move runs up against the absence of articles in Russian as well as the existence of adjective-like categories (3.4.3); and the second against the vagaries of 3.2.

4.3 Movement

4.3.1 Foundations

The *raison d’être* of ‘movement’ or ‘displacement’ in Chomskyan theory is the derivationalist desire to discern structure preservation between ‘similar’ strings, as already hinted at in 4.2.2. Whether to see the forest among the trees when the derivational component of the syntax consisted of many different Transformations, or the trees among the forest after the conversion into a ‘Move-alpha machine’, there has long been a demand in Chomskyan theory for some means of classifying the operations that are meant to take place in the grammar. One approach has been to concentrate on *what* is ‘moved’, thence the terminology of XP- (‘phrasal’) and X- (‘head’); another on *whence* the extraction occurs, thence ‘locality’; and a third angle on *whither* the movement takes place, thence ‘A-’ (argument) vs. ‘A’-’ (non-argument) movement.

The first point was covered in 4.2; the second, which follows on from it in that impermeable domains must be wholes of some kind in their own right, is the topic of 4.3.2; and the third, in turn connected with 4.4 by their common interest in the positions targeted by movement, is the focus of 4.3.3. The point in both cases is that over their lifetimes the distinctions have become conventionalised as diagnostics; and in turn that the notion of ‘movement’ has been ontologised in much the same way as the concept of ‘constituency’, so that constructs of quite particular provenance are enshrined in Universal Grammar. It remains to be seen on the strength of 4.3.2-4.3.3 and section 5 whether the Russian Split Scrambling data seem to break the mould as in 4.2.

4.3.2 *Locality Typology*

Locality is, similarly to if more so than the A/A' Binding-based diagnostics of 4.3.3, a multifaceted phenomenon. It finds its place at this point of the present paper because in the sense that it prescribes domains which bound movement operations to a greater or lesser extent it is about 'where things can move', in line with the rest of 4.3. But equally it could have been at home in 4.2, because insofar as it describes the relevant domains in their own right it is still about 'what can move'. This dual status is perhaps partly responsible for locality's enduring popularity as a topic of linguistic inquiry; below I copy some details from Szabolcsi (2006)'s picture of the scene.

The defining divide in her presentation separates 'strong' (32) from 'weak' (33) islands; their internal variety are illustrated here by adaptations of Szabolcsi's leading examples for each member of the locality menagerie (a few of which were examined for Russian SS by Pereltsvaig (2008), and some more for Russian XP Scrambling by Bailyn (2011). In the spirit of Nossalik in 4.3.3, I hasten to qualify the validity of some of the data given: sentential subjects (32b) as well as clausal complements to nouns and adjectives (33d) are more marginal in Russian than English, and scopal ambiguities less available (336); without articles (cf. 4.2.3) definiteness (32c) is not so easily manipulated in an unmarked fashion.

- (32) a. **Kakogo sleduet vyzvat' učitelja,*
 Which-GEN.SG.M follow-3SG.IPFV call-PFV teacher-GEN.SG
kotoryj nakazal reběnka?
 which-NOM.SG.M punish-M.PFV.PST child-GEN.S
 "Which kid must you call the teacher who punished?"
 (Complex Noun Phrase Constraint)
- b. **Kakoj čtenie knigi zainteresovalo*
 Which-GEN.SG.F reading book-GEN.SG interest-N.PFV.PST
novogo prezidenta?
 new-GEN.SG.M president.GEN.SG
 "Which book did reading interest the new president?"
 (Sentential Subject Condition)
- c. **O kakom ty našel stixi*
 About which-PREP.SG.M you find-M.PFV.PST verse-ACC.PL
Marii čeloveke?
 Marija-GEN.SG person-PREP.SG
 "Which man did you discover Marija's poem about?" (Definites)
- d. **O kakoj ty ušël potomu, čto*
 About which-PREP.SG.F you leave-M.PFV.PST because Marija
Marija pogovorila teme?
 talk-F.PFV.PST topic-PREP.SG
 "Which topic did you leave because Marija talked about?" (Adjuncts)

The geometrical structure of linguistic theory

- e. **Kakogo ty priglasil Mariju i*
 Which-GEN.SG.M you invite-M.PFV.PST Marija-ACC.SG and
človeka?
 person-GEN.SG
 “Which man did you invite Marija and?”
 (Coordinate Structure Constraint)
- f. **Kakogo sprosila Ivan, kto priglasil*
 Which-GEN.SG.M ask-M.PFV.PST Ivan-NOM.S who invite-M.PFV.PST
človeka?
 person-GEN.SG
 “Which man did Ivan ask who invited?”
 (Tensed Constituent Wh-Complements)
- g. **Kakogo ty uvidel kartinu človeka?*
 Which-GEN.SG.M you see-M.PFV.PST picture-ACC.SG person-GEN.SG
 “Which man’s picture did you see?” (Left Branch Condition)
- (33) a. %?/??? *Kakuju ty somnevaešsja, čitat’ li*
 Which-ACC.SG.F you-NOM doubt-2SG.IPFV read-IPFV whether
knigu?
 book-ACC.SG
 “Which book are you wondering whether to read?” (Wh-)
- b. %?/??? *Ja sprosila, kakuju Ivan ne čitaet*
 I ask-M.PFV.PST which-ACC.SG.F Ivan-NOM not read-3SG.IPFV
knigu.
 book-ACC.SG
 “I asked which book John wasn’t reading.” (Negatives)
- c. %?/??? *Kakuju ty rad, čto on čitaet*
 Which-ACC.SG.F you-NOM happy-M that he-NOM read-3SG.IPFV
knigu?
 book-ACC.SG
 “Which book are you happy that he is reading?” (Factives)
- d. **Kakuju skandal’no to, čto on čital*
 Which-ACC.SG.F scandalous-N that-N that he-NOM read-M.IPFV.PST
knigu?
 book-ACC.SG
 “Which book is it a scandal that he was reading?” (Extraposition)
- e. %?/??? *Kakuju on často čitaet knigu?*
 Which-ACC.SG.F he-NOM often read-3SG.IPFV book-ACC.SG
 “Which book does he read a lot?” (VP Adverbs)

- f. **Kakuju* *čital* *každyj* *rebėnok*
 Which-ACC.SG.F read-M.IPFV.PST every-NOM.SG.M child-NOM.SG
knigu?
 book-ACC.SG
 “Which book did every boy read?” (Scope)

The overall picture from (32) and (33) is that Russian SS is affected by SIs but not by WIs. Interestingly, I detected an additional disparity in grammaticality between split and non-split versions of the WI sentences, though the directionality of preference varied between speakers; semantic factors seemingly responsible for the spectrum effect. The data here are not so much a problem for the Chomskyan movement typology as an indication as to where SS fits into it. In theories, that ‘reality’ is reflected in the character of the devices proposed to account for locality effects. The term ‘island’ itself was invented by Ross (1967) for the strong variety, and has comparably colourful successors in ‘barriers’ (Chomsky 1986b) and ‘phases’ (Chomsky 2000); ‘intervention’ and ‘relativisation’ (Rizzi 1990) are the watchwords on the weak side of the equation.

4.3.3 Landing Site Typology

Another step along the way is to fulfill 4.3.1’s promise of taking a closer look at the constraints invoked as movement diagnostics. The origin of the A/A’ movement typology is in two observations: firstly, that some operations—such as passivisation—seem to target positions usually filled by arguments and therefore associated with certain grammatical features, whereas others—e.g. wh-question formation—do not; and secondly, that the two groups of operations seem to be affected more or less coherently by certain constraints, which were the subject of 4.3.2. Over time the connection between the two groups of facts has become fossilised, so that operations are routinely placed in one class or the other solely on the basis of which set of constraints they seem more bound by.

In the context of Russian (non-Split) Scrambling, Nossalik (2005) gives the standard A/A’ set established on the basis of Binding, a syntactic relation that sometimes holds between coindexed elements. It is her examples that I use in (34). First there is the creation of new binding relations:

- (34) a. A: Yes
 Peter seems to himself_i t_i to be smart.
 b. A’: No
 **Whom_i does it seem to himself_i that Mary likes t_i?*

Secondly and conversely there is Reconstruction of old binding relations:

- (35) a. A: No
 Pictures of John_{ij} seem to him_i t_j to be on sale.
 b. A’: Yes
 **It’s pictures of John_{ij} that he_i sells t_j.*

Thirdly there is Weak Crossover effect with an operator and a non-c-commanding coindexed variable:

The geometrical structure of linguistic theory

- (36) a. A: No
Who_i seems to his_i mother t_i to be happy?
 b. A': Yes
 * *Whom_i does his_i mother love t_i?*

Fourthly and finally there is the licensing of Parasitic Gaps, i.e. traces which are bound by another trace and only via it by an overt antecedent:

- (37) a. A: No
 * *John_i was killed t_i by a tree falling on pg.*
 b. A': Yes
Which articles_i did John file t_i without reading pg?

While recognising these tests as more or less robust for English, Nossalik takes issue with their applicability to Russian. Specifically, its anaphoric and pronominal inventories are not isomorphic to their English counterparts: the reflexives having larger domains and a subject orientation; the anaphors being on the contrary subject-avoiding; and the language as a whole having an aversion to backward anaphora. Together with the dubiousness of the existence of Parasitic Gaps in Russian, this reduces the list in (34)-(37) to Reconstruction and Weak Crossover, which can both work with Binding Condition C (thus avoiding the confounding properties of reflexives and anaphors in A and B respectively). Nossalik still makes a case for the A/A' hybridity of Scrambling, remade in (38) for SS.

- (38) a. * *Knigi Ivan_i prodaët on_i interesnie.*
 Books-ACC.PL Ivan-GEN.SG sell-3SG.IPFV he-NOM interesting-ACC.PL
 "He_i is selling Ivan_i's interesting books."
 b. *Každogo_i ljubit ego_i mat'.*
 Every-GEN.SG.M love-3SG.IPFV own-NOM.SG.F mother-NOM.SG
človeka_i.
 person-GEN.SG
 "His_i mother loves every man_i."

The case is that the Reconstruction example in (38a) is excluded, indicating SS is A'-movement; whereas WCO as in (38b) is allowed, suggesting SS is A-movement: paradox! The dilemma will resurface in *passim* in section 5, particularly in 5.5.3.

4.4 Discourse

4.4.1 Foundations

In this subsection I turn to the formalisation of native speaker intuitions about the nuances of meaning associated with the Russian SS constructions of 3.2-3.5 and reported on in 3.6. Over the years the 'structure' associated with such properties and the 'theory' that explicates it have been labelled as 'Discourse' or 'Information'. It will also be recalled that three approaches—those of Yokoyama (1986), Lambrecht (1994), and Erteschik-Shir (2007) have been particularly relevant to recent treatments of Russian such as King (1993), Bailyn (1995, 2011), Sekerina (1997), Slioussar (2007) and Dyakonova (2009). As he occupies something of an axiomatic middle ground

between the other two and for purposes of illustration I go with Lambrecht in the following paragraphs.

The centrepiece of Lambrecht's IS framework is the overarching concept of 'Information' itself. 'Presupposition'—of shared prior knowledge, that is—and 'assertion'—of a proposition about entities either presupposed or not—are the 'forces' in his universe of discourse. Other important occupants include the properties of 'identifiability' and 'activation', both predicated of co-resident entities. From the forces and properties combined are derived the 'pragmatic relations' of 'topic'—a presupposed, identifiable entity about which an assertion is made—and 'focus'—an entity which crucially has high activation—which have already featured as Praguian heirlooms and which hold between 'discourse referents' in 'mental representations'.

What if we try to fix the 'facts' of 3.6 into Lambrecht's 'constructs'? What I there called 'saliency' tallies tolerably well with topicality, and 'significance' fits the bill of focality. The two relations seem to be independent in Russian discourse-configurationality, as sentence-initial position can be associated with either the topic or the focus—though sentence-finality is monopolised by focality. Intermediary positions are also marked by the discourse duality—with topicality decreasing from left to right, and proximity to the edges correlating positively with focality. Although the nature of SS yields a quantitatively significant number of configurations on the basis of which to make such observations, it is in terms of relative rather than absolute position that it has a *qualitatively* distinctive contribution to make to the description of Russian IS—in that speakers seem to have an intuition that a larger 'distance' between A and B in an SS construction corresponds to greater 'emphasis' (i.e. focality).

4.4.2 Interfaces

In 4.4.1 the skeletal properties that emerge from the Russian SS data were given some theoretical flesh. In this subsection an issue is addressed which, in turn, comes half-formed out of the discussion of Discourse/Information Structure/Theory: interfaces, or the interaction between DS/IS and e.g. syntax or phonology, taken as separate 'components' of the grammar—an ontological difference of the next order up from that between the base and derived structures connected by movement in 4.3.1.

One aspect of the 'interface relationship' is 'proximity'. Lambrecht (1994) frames the question in terms of 'autonomy' and 'motivation' and, as will be recalled from 4.4.1, is agnostic about the answer. For Erteschik-Shir (2007), the debate is about 'discourse-configurationality' and thus syntax-specific. The raw and processed data patterns from 3.6 and 4.4.1, respectively, suggest that Russian may be numbered among discourse-configurational languages, though not to the degree of identification required by a simplistic interpretation of Miyagawa (2010)'s typology opposing them to agreement-based ones.

In addition to the *proximity* of the interaction, another necessary detail in the interfaces picture is its *directionality*. To deconstruct that—is there an ordering between the components of the grammar or not, i.e. is the derivation serial or

parallel? If there is an ordering, what are the relations of priority between the parts of the system? And is the derivational progression from first to last in the list singular or iterative, cyclical or otherwise shaped? Chomskyan Generativism—and therefore all the analyses in section 5 apart from RT—assumes serialism, a post-syntactic interpretive component, and a single derivational cycle.

The preceding paragraphs find their place in the exposition not because their contents is in itself decisive in the analysis of SS as that it forms the bedrock on which the other assumptions arising out of this section, which *will* be of some moment in section 5, are founded.

4.4.3 *Optionality & Gradience*

The twin themes of ‘optionality’ and ‘gradience’ may be drawn out of both 4.4.2 and 3.6. On the one hand, the characteristics of indeterminacy and continuousness could have been included as part of the interfaces parcel, a product of the proximity and directionality of the interaction between components of the grammar. On the other hand, one might perceive a certain amount of ‘free choice’ or ‘margin for error’ in the range of Russian surface word orders, the discourse and/or informational properties of which were detailed in 4.4.1. As concepts optionality and gradience also have their own contexts in linguistic theory and significance for Russian SS, which are the subjects of the following paragraphs.

One of them is the ‘computational’ nature of Chomskyan theory, an implicit if not an express design feature of the Generativist drive towards explicitness and formalism—with the notions of a ‘grammar’—a machine that produces a ‘language’ (a set of strings with certain properties)—inscribed in the foundations of both Generative linguistics and computer science. Classical computation has no place for ‘maybe’ and ‘sort of’, which from the intuitions of native speakers reported in 3.6 seem to be at work in Russian SS—in that a change in meaning is not necessarily marked, at least in the syntax, and a change in the syntax does not necessarily have a straightforward correspondence in interpretation.

The other, not altogether unrelated context I wanted to touch on here is the Chomskyan separation of meaning from form—that is, of semantics and pragmatics from the distributionally defined level of constituency of 4.2.1. Although many devices from formal semantics—i.e. elements of propositional and predicate logics—have been co-opted more or less covertly by Chomskyan syntax, pragmatics—into which DS/IS is usually absorbed, though it possesses certain truth-conditional powers—has remained largely unassimilated. The phenomenon of discourse-configurationality, which only just made an appearance in 4.4.2, might be used to bridge the form-meaning divide—a feat which will be counted among the attractions of Representation Theory in 5.5.4. As in 4.2 and 4.3, the main message from the discussion of this section is about the Chomskyan ontologisation of theoretical concepts.

4.5 *Conclusion*

The presentation of ‘issues’—of Phrase Structure (4.2.2) and the DP Hypothesis (4.2.3) in the context of constituency (4.2); of strong and weak islandhood aspects

of locality (4.3.2) along with landing sites and diagnostics for their involvement (4.3.3) as part of the typology of movement (4.3); and of interfaces (4.4.2) as well as optionality and gradience (4.4.3) in connection with discourse/information structure (4.4)—in this section has been more or less disparate and descriptive in nature. It has nevertheless imposed some linear and conceptual order on the data of section 3; and also provided pointers on possible inroads into the territory of architectural features. Both processes will be continued in section 5. A side-effect of the transition from general issues to specific analyses should be to make clear connections which might not have emerged fully here, particularly between constituency and movement—with a shift of ground underneath one, the other, or both being the way out of the ‘SS problem’ in one way or another for all of the accounts considered.

5 ANALYSIS

5.1 Introduction

This section takes the final step through the substantive core of the paper, thereby reaching the application of concrete theoretical analysis to Russian Split Scrambling. Following the lead of the papers discussed in the following sections, particularly Fanselow & Ćavar (2002) and Pereltsvaig (2008), I review representatives of the ‘mainstream’ Chomskyan alternatives, in each case progressing across three subsections from constituency to movement to discourse (in parallel with section 4). Section 5.2 starts the ball rolling with Fanselow’s Base Generation (Fanselow 1988), which is found simultaneously too permissive and too restrictive with regard particularly to constituency and discourse respectively. 5.3 continues with Bošković’s (2005) Left Branch/Direct Extraction version of Simple Movement, which offers a neat way of parameterising constituency but does not extend beyond the simple NP cases and has no discourse account at all. And 5.4 brings us to Fanselow & Ćavar’s (2002) Distributed Deletion, which has been taken on in Russian contexts by e.g. Pereltsvaig (2008) and Goncharov (2012); it purports to solve the problems of both ‘movement’ and ‘non-movement’ analyses, which it can if given the power it demands but then is difficult to stop. I claim that there are still both empirical and conceptual reasons to seek a different solution; and in Section 5.5 I sketch the outlines of such an analysis in the Representation Theory framework of Williams (2003), addressing constituency in 5.5.1, locality in 5.5.2, landing sites in 5.5.3, and discourse in 5.5.4—a slight modification of the previous template. In the light of RT SS the ‘exceptionality’ implied by both the ‘Split’ and the ‘Scrambling’ in SS disappears. The point of the exercise is not to have the last word on SS in Russian, let alone cross-linguistically and all the less so on the potential structural problems in the architecture of linguistic theory the phenomenon brings to light. Concluding the section in 5.6, however, I do summarise why I believe RT has the right *sort* of features to accommodate Russian SS.

5.2 Base Generation

5.2.1 Constituency

In an immediate sense BG offers an easy solution to the conundrum of constituency, as it ‘escapes’ the X-Bar template in favour of its own ‘twin’ NPs. It is thus in principle possible to obtain any word order one wants on the basis of a discontinuous nominal projection. In practice, it is difficult to see how one could keep the system from undergenerating without letting it overgenerate, or vice versa if extended beyond the simple nominal case. A noteworthy feature of Fanselow’s account is the virtue he makes—in the German context—of it being able to capture ‘imperfect splits’ (where A and B would not make a viable constituent if put together, either because of the constituents themselves—as in (1), in fact—or relations between them), which is at best redundant and at worse a liability in light of the Russian SS data, where the *presence* of agreement between A and B seems not only obligatory but also a licensing factor; and prepositional doubling (20c) is marginal.

5.2.2 Movement

The case seems to be open and closed in the opposite direction for movement: surely BG is ‘BG’ because it does not resort to movement? And sits on one side of the fence set up by Pereltsvaig (2008), with symptoms of ‘non-movement’—such as mixed islandhood behaviour (see (32), (33))—in its favour? Alternatively, one might see the binding relation which Fanselow *does* posit between the two positions as the potential beginning of a new outlook on ‘movement’. If one were to pursue that path, however, the question would arise as to beyond what point BG would to all and intents and purposes stop being BG. As it is, there is already a problem in the shape of landing sites: the A/A’ diagnostics of 4.3.3 do not straightforwardly support Fanselow A=A’ and B=A identification. And—something also true of 5.3 and 5.4—there is little in BG to tell us *what* the multiple positions occupied in (3.3.1) are and *why*.

5.2.3 Discourse

Strikingly given its ‘profligacy’ with *positions*, if maybe unsurprisingly given its vintage, the ‘discourse vocabulary’ of BG is very small—limited, in fact, to a [focus] feature. In weak or non-discourse-configurational languages that might be enough, and coming from a particular point of view on the nature and desirability of theoretical parsimony, the operational minimalism it betokens might be seen as a forte. Given the range of pragmatic effects observed, described, and explained to various degrees in 3.6 and 4.4.1, however, it seems unlikely that BG can match up to the task of capturing the nuances of meaning involved in Russian Split Scrambling (cf. 3.6, 4.4.1). BG is thus a curious mixture of restrictions and freedom, which does not seem to be as effective as could be desired for our purposes.

5.3 Simple Movement

5.3.1 Constituency

This is why there was the interest in Russian adjective-like words in 3.4.3 and the DP Hypothesis in 4.2.3. The fact that the Russian equivalents of what would be the D in the DP have a similar distribution as adjectives by the ‘test’ of SS lends some credibility to the LBE story—but note ‘similar’ instead of ‘same’. Although it distinguishes itself from at least BG by its approach cross-linguistic variation, however, Bošković’s version of events is far from uncontroversial—being explicitly and energetically rejected by Bailyn (2011) for Russian. Like Fanselow (1988), moreover, Bošković (2005) is only really designed to deal with the Adj-N splitting of a nominal—which, as was mentioned in 5.2.1, is far from exhausting Russian SS. Neither of the templates in Figure 3 (see Figures 4 and 6 for follow-ups), for instance, would extend to having P and Adj/N as a constituent (20), and Bošković’s NP would actually make the Adv examples (21) more difficult to accommodate than an Abney’s DegP (in which Adv and Adj would be somewhat closer). There is thus a paradox as to whether Abney’s or a more standard X’ constituency model is better for Russian SS.

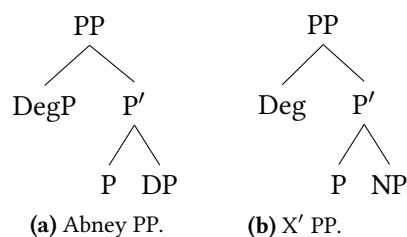


Figure 3 PP templates.

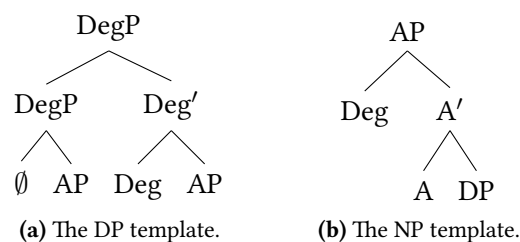


Figure 4 AdvP templates.

5.3.2 Movement

SM is targeted by Pereltsvaig’s (2008) arguments equally and oppositely to BG 5.2.3. On the one hand, consequently, the SM approach has on its side the morphological agreement between A and B in SS constructions—suggesting, at least in the conventional localistic conception of agreement, according to which the two elements were adjacent at some point in the derivation. Furthermore, SS displays

some sensitivity to islandhood (cf. 32), which is another diagnostic of movement in the Chomskyan sense. By the same token, however, SM is then flummoxed by SS's apparent hybridity between SIs (32) and WIs (33), as per 5.2.3, as well as A and A' status, cf. (38). It also leaves some loose ends in the sphere of 'complex constituents' (3.5) and their inconsistent susceptibility to SS.

5.3.3 *Discourse*

Finally, there is at first glance even less to be said about the treatment of discourse in SM than in BG as Bošković (2005) saves no space whatsoever for the consideration of such matters; one might wonder whether his purely formal focus is entirely in keeping with the spirit of the Minimalist Program. For justice's sake it should therefore be noted that Bošković (2004) has detailed discussion of 'topicalisation' and 'focalisation', albeit in the context of polemic with Bailyn and on the understanding that 'Scrambling' proper is something different. Both in terms of the distinctions made and the place assigned to them in the system, therefore, SM does not seem suited to Russian SS.

5.4 *Distributed Deletion*

5.4.1 *Constituency*

In the case of constituency DD bears certain similarities to BG, as two NPs are involved—though DD's twins are identical in the eyes of the syntax, with the dissimilarity appearing when they reach the interfaces. Although the difference of detail will be important in 5.4.3, the comparison may be continued by raising the counterpart of the problem which faced BG, and which delves into a dilemma running through all the analyses in this section (Representation Theory not excluded) between mechanical permissivity and restrictivity. Without further work DD like BG errs too far on the side of laissez-faire, whereas SM is too strict. All three, however, suffer through not accounting—whether liberally or conservatively—for discontinuity outside simple NPs; although DD does at least have an account for the 'inversion' of 3.3.1 (Fanselow & Ćavar 2002: 88ff).

5.4.2 *Movement*

For its adherents DD, nevertheless, is also the way out of a paradox they perceive proceeding from the 'movement' vs. 'non-movement' opposition between previous treatments of SS such as SM and BG, respectively. Where the SS evidence—which my Russian data corroborated in (32)-(33), (38)—is split between the two poles, the DD claim is that the 'copy and delete' mechanism can have it both ways. On the one hand, that is, DD does involve movement—therefore SS should be subject to 'distance-related' locality effects; and so it is, as demonstrated by the marginality of LD constructions (30)-(27). On the other hand, the copying mechanism should obviate constraints associated with the locus of extraction; here the S-O and IO-DO asymmetries of 3.2, the other complex constituents of 3.5, and strong islandhood (32) seem problematic.

5.4.3 Discourse

Fanselow & Ćavar (2002) are more generous than either Fanselow (1988) or Bošković (2005) in the ‘discourse compartment’ of their feature inventory, including both [link-topic] and [focus]. For them the right-hand part of the the discontinuous constituent (‘B’ in the terms of 3.1) must be focalised, whereas the left-hand one (‘A’) may be either a topic or a ‘second focus’. Such ‘duality of focality’ is in line with the expectations of 3.6.5 and 4.4.1, but its content is not elucidated in the paper. For that we will have to wait until Representation Theory’s turn, in 5.5.4. The other main feature of the DD discourse package is the assertion that splitness only surfaces if there is a demand (which overcomes the ‘inertia of adjacency’) for one constituent to fulfil two functions, labelled ^p and ^q in (43)—where $XP=[a^p [b c]^q]$. This idea, which has parametric ramifications, features in a different form in RT and is integral to rather than stipulated in its architecture. RT also has things to say about the competition between discourse factors and their gradience (cf. 3.6.4-3.6.6, 4.4), on which DD (plus BG and SM a fortiori) is silent. To it I now turn, therefore.

5.5 Representation Theory

5.5.1 Constituency

The relevance in the constituency context of RT’s overarching architectural features as per 5.5.1 is that each Structure has its own ‘constituency’ and ‘headedness’, the properties of which are determined by the nature of the relation(s) responsible for it. The closest match to conventional Chomskyan constituency is therefore to be found at the earliest, smallest levels—CS-TS and PS-?S—where relations such as argumentation, agreement, and binding are encoded; at the other end of the scale, in SS-QS and AS-FS the discourse relations of 4.4.1 come into their own (cf. 5.5.4). The highly local ‘priority’ of Phrase Structure is thus not ontologised in the same way by RT as by GB or even MP. And all the less so because Williams explicitly derives the X’ schema by axiomatisation:

(39) *Axiom 1 (The juncture types of X-bar theory)*

There are just three juncture types:

- a. mother node = X > daughter node (embedding)
- b. daughter node = mother node > X (satisfaction)
- c. mother node = daughter node (adjunction)

(40) *Axiom 2 (Pollock-Cinque Functional Hierarchy)*

There is a universal set of elements (T, Agr_O, Agr_S, ..., V) that are in a fixed chain of complement-of relations:

(T > Agr_S > ... > Agr_O > V)

Labels must be a subsequence of this hierarchy.

where previously, at least until the promulgation of Bare Phrase Structure (more or less contemporary with Williams 2003) the template had been taken as fundamental. His model of the simple clause following from (39) and (40) is in Figure 5 (Williams 2003: 180), with Figure 6 providing a Structure-based equivalent (21).

The geometrical structure of linguistic theory

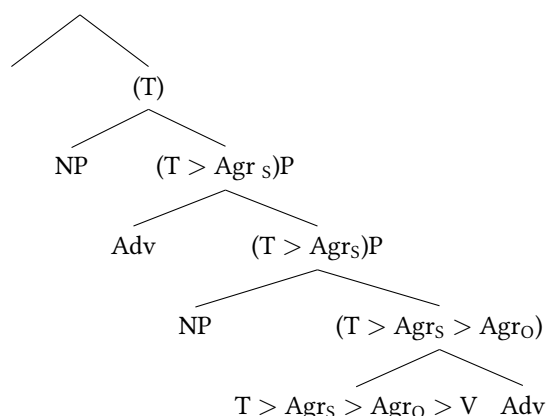


Figure 5 The RT model of the simple clause with a single main verb (PCFH).

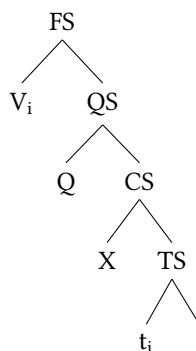


Figure 6 The RT model of the simple clause with a single main verb (Structures).

In light of this SS is ‘split’ only in the relative sense that there are mismatches between constituencies and mismappings between their respective Structures. Elements that live together at one level of representation—such as Adj and N (5), P and NP (20), Adv and NP (21), Adv and Adv (22)—need not cohabit at all the others if some feature will be satisfied by scrambling away, to a place where non-constituents at lower levels may be fully entitled to constituent status. This is particularly pertinent in the adverbial and prepositional cases, which will not fit any of the constituency templates of 2.2-2.4—nor state of the art BPS—but have in virtue of the ‘parasitism’ of the invariable functional word on its inflected lexical partner (3.4.2) the sort of ‘wholeness’ that Williams (2003: 29–30) regards as beyond MP Checking Theory. At the same time, there are some clear cases—e.g. the majority of agentive subjects (4) and complex constituents (3.5), and strong islands (32)—where by contrast earlier, more local bonds trump later, more global ones. These will be taken up in their own right in 5.5.2—and it will be seen there that RT does have its blind spots. For now, however, the model seems equal to the core constituency facts as well as to those which pose problems for the previous analyses.

5.5.2 Locality

Although I separate them for the sake of exposition (in parallel with the segregation of 4.3.2 and 4.3.3) the substance of this subsection and 5.5.3 is to a large extent overlapping in the RT worldview. That is to say, the same architectural principle—the ‘Level Embedding Conjecture’, according to which “an item can be embedded exactly at the level at which it is defined, and no other” (Williams 2003: 64)—is invoked to account for locality as for the A/A’ facts which will be the focus of 5.5.3. Insofar as they are incorporated into the framework, both cases crucially involve the relative ‘timing’ of the establishment of relations, with operations unable to “extract from structures that are not embedded until after the level at which [they] apply” (p. 165).

Williams’ own exposition of the LEC’s consequences speaks to a number of the apparent constraints on Russian SS observed across section 3 and in 4.3.2. The intersection of domain size and relation type in level embedding makes Structures analogous to but more flexible than ‘propositional’ Phases c2001; the ‘junctures’ of 3.5, particularly the clausal cases of (26)-(30), and many of the islands in (32) and (33) might be seen as resisting the advances of SS because of their relational completeness and integrity, with incomprehensible fragmentation or even ambiguity the result of extraction from them. The S-O asymmetry emerging from (4), (5), and 3.2.1 might in turn be associated with the multiplication of subjecthood resultant from the LEC, with the ‘earlier’ thematic subjects less susceptible to SS than the ‘later’ surface ones. Williams’ partial subject taxonomy is shown in Figure 7 (Williams 2003: 85); given that Williams himself does not consider ditransitives and their DO-IO asymmetry (10)(11), one might assume that similar reasoning applies to them.

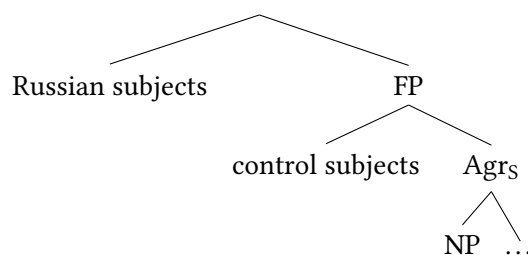


Figure 7 Williams’s (2003) partial subject taxonomy

There are, however, lacunae in the RT locality landscape—which are openly declared in Section 6.4 of the monograph as “restrictions [...] that fall under the traditional rubrics of Subjacency and the ECP” (Williams 2003: 165), in the concrete context of the *Wh* Island Constraint. After mooted a solution on the same page Williams decides to leave the issue open, positing the existence of some unknown locality constraints on embedding “beyond those predicted by RT” (p. 166). Given that the *Wh* Island Constraint falls on the ‘weak’ side of the islandhood strength scale from 4.3.2, one might salvage something from this admission of defeat—in that it indicates some basis for the SI-WI contrast with Russian SS in (32) and (33). As its

author is also keen to admit elsewhere in the monograph, RT is very much a work in progress.

5.5.3 Landing Sites

At the beginning of this subsection as in 5.5.2, I make a few remarks relevant to the topics of both. RT is in a sense all about ‘scrambling’: it will be recalled from 5.5.1 that its creator uses *that* word, not ‘movement’, to refer to the force behind the mismappings between RT’s multiple Structures, and thus behind the theory’s driving dynamics. This terminological departure from the Chomskyan norm is symptomatic of a conceptual fresh start of the sort envisaged in 5.5.4, in some sense approximating a state of equilibrium between the derivational and representational extremes in the debate on the nature of the grammar, between which linguistic thought has oscillated over time as per Lasnik (2001).

Turning now specifically to landing sites: of all the issues which serve as theoretical testing grounds in this and the other sections of the section, it is here that RT probably presents the most impressive front—as may be deduced, indeed, from the structural summary of 5.5.1. All three of what were there labelled the “core” sections of Williams (2003) are relevant to the A/A’ distinction of 4.3.3, with all four of Nossalik’s (2005) diagnostic phenomena covered in some detail apart from parasitic gaps. As also mentioned in 5.5.1, Williams’ strategy in all cases is to generalise the relation or opposition in question across Structures. The A/A’ discrepancy displayed by Russian SS in (38) is thereby defused, because binding and quantification relations are defined on different levels (PS and QS, respectively) and thus in distinct A/A’ terms. Table 1 charts the reconstruction possibilities into which they thus fit (Williams 2003: 136).

This → reconstructs for this ↓	Focus	Wh movement	Long scrambling	Short scrambling	NP Movement	Movement for Case
Wh movement	✓	✗	?	✗		
Long scrambling	?	?	?	✗		
Weak quantifiers	✓	✓	✓	✗	?	?
Anaphor binding	✓	✓	✓	✗	✗	
Short scrambling	?	?	✗	✗	✗	
NP movement	✓?	✓	✓	?	✗	
Q-float	✓?	✓	?	✗	✗	✗
Theta relations	✓	✓	✓	✓	✓	✓

Table 1 What reconstructs for what

Moving beyond the level of debate in the last paragraph—that is, beyond inferences made about SS’s *type* of landing site on the basis of A/A’ tests—BG, SM, and DD are for the most part quite coy about the *exact* identity of the positions in question.

This agnosticism was both reflected in the data presentation of section 3 and partly justified by pointing out (in 3.3.1) that a classic device such as the insertion of adverbs is undermined by their own susceptibility to Scrambling. Given that the representation of a sentence in RT is a compound constructed from level embedded ‘subtrees’ (cf. Figures 5, 6), the categorical identification of landing sites is made both more manageable and less urgent as a problem. The bewildering range of configurations (see 3.3.1) can be restrictively accommodated by tying them in ‘step by step’, from L to H, with successively larger Structures, as per the Phase comparison in 5.5.2 and the detail of 5.5.4.

5.5.4 Discourse

Although they have been held back until last in this section (to parallel their relative positioning in sections 3 and 4), Williams almost *starts* and *ends* with Topic and Focus—in a way reminiscent of the double configurationality of Russian focus from 3.6.5 and 4.4.1. The first point of substance in this subsection comes back to Figure 2 in 5.5.1, where QS and AS—the formal strongholds of discourse meaning—appeared casually alongside the other RT Structures. That is really the point in itself: in RT discourse meaning is recognised as a force both in its own right and on a par with the rest, not as subordinate features like in BG or DD, and not as a watertight module like Bailyn (2011)’s FF.

With regard to its typology of discourse relations, RT is ‘quantitatively’ rather than ‘qualitatively’ ahead of the established analyses to which I juxtapose it. For Williams topics are primitive in that they come in one shape and size; foci, on the other hand, can be either ‘logical’ (elicited by the question-answer method of 3.6.5, which generates a propositional presupposition) or ‘informational’ (linked with the main sentence accent, Information Structure, and sometimes with contrastivity or metalinguistic correction). The intuitions from 3.6.5, formalised somewhat in 4.4.1, about a three-way split in the discourse meanings associated with Russian SS, are thus naturally accommodated. In (41a) Williams (2003: 31) and (41b) Williams (2003: 253) linear models are given after the style of (31); Figure 8 Williams (2003: 249) zooms in on the relevant part of Figure 2:

- (41) a. [Topic XP*][non-Topic ...X...XP*]
 b. [saw]_{IFocus} [[John]_{LFocus} [Bill xed someone]_{LPresup}]_{IPresup}

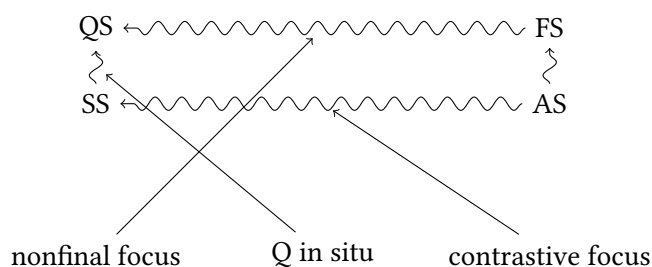


Figure 8 RT Architecture: Right-hand side.

RT has more yet to offer, however: according to Williams, the assignation of the three discourse functions to different RT Structures combines with the already familiar architecture of the framework to produce effects that it would be difficult—still *pace* Williams—to capture in Chomskyan Checking Theory (Williams 2003: 251ff). He has in mind the subordination of LFocus to IFocus and the ‘projection’ of IFocus across the sentence, but one can imagine how the same or similar mechanics might be used to derive the scalarity of 3.6.6 and gradience of 4.4.3. And it does not end there: Williams also ties in topic and focus with the empirical fact of cross-linguistic variation (2.4-8, 9.3.3-4) and the theoretical concept of compositionality (2.9, 9.1), as well as semantic phenomena as seemingly diverse as Heavy Noun Phrase Shift (2.3), scope (2.6) subjecthood (2.8), blocking (9.2), and ellipsis (9.4). Buying into the RT discourse story thus brings other things ‘for free’.

5.6 Conclusion

In the eponymous RT monograph of 5.5 Williams anticipates his critics by being the first to admit there are many loose ends in the tapestry of his framework, and that its presentation therein is most definitely programmatic rather than comprehensive. One's immediate thought might then be to dismiss it and move on, as indeed I advocated with Base Generation (5.2), Simple Movement (5.3), and Distributed Deletion (5.4) in succession. I plead that an exception be made for Williams (2003), however—because its underdevelopment is that of a wide-ranging framework in its nascent stages rather than of a more or less closed and phenomenon-specific analysis; and because, as I averred in 5.1, it has the right *sort* of features to bring it success in the SS arena—with respect to constituency (5.5.1), locality (5.5.2), landing sites (5.5.3), and discourse (5.5.4): the key in my mind is its ecumenical recognition of multiple structures on the ontological level monopolised by ‘constituency’ in the Chomskyan sense within the Chomskyan worldview; from there one moves with relative ease onto the relationship between them and between form and meaning. In short: I neither deny the importance of Chomskyan mental categories nor propose a fully-fledged alternative to the system which has grown up around them; I instead try to identify some structural flaws in the paradigmatic edifice and suggest some ways one might think of alleviating them. That is all.

6 CONCLUSION

6.1 Retrospective

I structure this section and 6.2 with respect to the three aims stated back at the end of 1.2. The first was to offer an overview of the ‘facts’ regarding Russian Split Scrambling. This was the task of section 3 (which organised most of the data I collected along the lines of predication, permutation, categorisation, complexification, and interpretation), with some additional input from 4.3.2 and 4.3.3. The overall picture was one of great power but also of great responsibility: the space of possible constructions is indeed large, but seems to be constrained by countervailing principles.

The next aim was to review the background issues associated with, and existing analyses formulated of, SS. This was inevitably implicit in the pursuit of the first, given that the right sort of data needed to be collected for this next step to be made. The focus of the review was, however, on sections 4 and 5, which match the sub-aims more or less one to one: the questions considered cluster around the issues of constituency, movement, and discourse structure/interfaces—on all of which SS sheds an interesting light; the accounts chosen from the list in ?? were Base Generation, Simple Movement, and Distributed Deletion—which move to accommodate the revelations of SS by modifying some architectural characteristic(s) of the Chomskyan framework in which they all operate.

The third and final aim was to provide a preview of an analysis to accommodate the facts of Russian SS and of a framework for thinking about the design features of linguistic theory. This was chiefly the concern of 5.3: the analysis was formulated in Representation Theory, dealing with the constituency, movement, and discourse in turn; and the framework was built around the aspects of RT, as well as other non-Chomskyan paradigms. With regard to the all three aims but to this last more than the others, there is much still to be done. I explore some possible avenues of further inquiry in 6.2 by way of valediction.

6.2 *Prospective*

Returning to the first aim, it is immediately clear that the empirical scope of this paper is somewhat parochial. Even within the sphere of Split Scrambling in Russian, there are doubtless many nooks and crannies of possible data space that space, time, and imagination prevented me from inspecting—in particular where complicated clauses, complex sentences, and the complexities introduced by confounding semantic, morphological, and phonological factors are concerned. Beyond *Split Scrambling* in Russian, there are the larger syntactic spheres of word order and Scrambling, as well as the alternative means to linguistic ends identified in morphology and phonology; beyond Split Scrambling in *Russian*, there are many other languages which might well offer useful points of comparison; and ‘beyond’ more generally, the vistas of Universal Grammar itself beckon.

The range of issues drawn from the data and approached by the analyses, in turn, was of course influenced by my own interests as well as those of those whose work I cited. So pervasive is that bias, in fact, that it is difficult to escape it now. An easier way can be found out of the problem of how to expand on the analysis front, as there is a large if incomplete list of options in from which one selection was made here but many others are possible and probably would be informative, whether because of new individual perspectives or of novel perspectival combinations—always bearing in mind the necessity of comparability, which is no small ask given scholars’ highly diverse theoretical assumptions and empirical targets.

If there are many ‘old’ analyses that have not found their way into this paper, there are maybe even more ‘new’ ones waiting to happen—not only because the ratio of coverage old:new was 3:1 but also because one might always modify an existing analysis to the point where an outside observer could take it to be a different

one; or develop different accounts within the same framework; or compare and contrast treatments derived in different ones. Starting from RT, an obvious way to go would be via the acknowledged similarities from 5.5.1. But on this, and on anything else, I now must now be silent tractatus.

REFERENCES

- Abney, Stephen Paul. 1987. *The English noun phrase in its sentential aspect*: MIT dissertation.
- Bailyn, John Frederick. 1995. *A configurational approach to Russian free word order*: Cornell dissertation.
- Bailyn, John Frederick. 2011. *Russian Syntax*. Cambridge: Cambridge University Press.
- Bošković, Željko. 2004. Topicalization, Focalization, Lexical Insertion, and Scrambling. *Linguistic Inquiry* 35(4). 613–638.
- Bošković, Željko. 2005. On the Locality of Left Branch Extraction and the Structure of NP. *Studia Linguistica* 59(1). 1–45.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. London: Blackwell.
- Chomsky, Noam. 1975/1955. *Logical Structure of Linguistic Theory*. London: Croon Helm.
- Chomsky, Noam. 1986b. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. *Ken Hale: A Life in Language* chap. Derivation by Phase, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2002/1957. *Syntactic Structures*. Berlin: Mouton de Gruyter.
- Collins, Chris. 1997. *Local Economy*. Cambridge, MA: MIT Press.
- Csirmaz, Anikó. 2004. Book review [of Williams (2003)]. *Acta Linguistica Hungarica* 51(3–4). 419–435.
- Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler Syntax*. Oxford: Oxford University Press.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.
- Dyakonova, Marina. 2009. *A phase-based approach to Russian free word order*: University of Amsterdam dissertation.
- Emonds, Joseph & Rosemarie Whitney. 2006. *Blackwell Companion to Syntax* chap. Double Object Constructions, 73–144. Blackwell.
- Erteschik-Shir, Nomi. 2007. *Information Structure: The Syntax-Discourse Interface*. Oxford: Oxford University Press.
- Fanselow, G. & C. Féry. 2006. Prosodic and Morphosyntactic Aspects of Discontinuous Noun Phrases. *Linguistics in Potsdam* 26.
- Fanselow, Gisbert. 1988. Aufspaltung von NPn und das Problem der “freien” Wortstellung. *Linguistische Berichte* 118. 91–113.
- Fanselow, Gisbert & Damir Ćavar. 2002. *Theoretical Approaches to Universals* chap. Distributed Deletion, 65–107. Amsterdam: John Benjamins.

- Fillmore, Charles. 1965. *Indirect Object Constructions in English and the Ordering of Transformations*. The Hague: Mouton.
- Goncharov, Julie. 2012. P-Doubling in split scrambling: A renaissance analysis. In *Actes du congrès annuel de l'Association canadienne de linguistique*, .
- Hornstein, Norbert & Andrew Nevins. 2005. Book reviews [Williams (2003)]. *Linguistic Society of America* 81(3). 757–761.
- Huang, James. 1982. Move wh in a language without wh-movement. *The Linguistic Review* 1. 369–416.
- Isačenko, Aleksandr. 1966. O grammatičeskom porjadke slov. *Voprosy jazykoznanija* 6. 27–34.
- King, Tracy Holloway. 1993. *Configuring topic and focus in Russian*: Stanford dissertation.
- Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, MA: MIT Press.
- Krylova, Ol'ga Alekseevna & Serafima Alekseevna Xavronina. 1984. *Porjadok Slov v Russkom Jazyke*. Moscow: Russkii Jazyk.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lasnik, Howard. 2001. *The Handbook of Contemporary Syntactic Theory* chap. Derivation and Representation in Modern Transformational Syntax. Oxford: Blackwell.
- Lebeaux, David. 2011. *Where Does Binding Apply?* Cambridge, MA: MIT Press.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.
- Miyagawa, Shigeru. 2010. *Why Agree? Why Move? Unifying Agreement-Based and Discourse-Configurational Languages*. Cambridge, MA: MIT Press.
- Nossalik, L. 2005. The A/A'-status of Russian scrambling. *McGill Working Papers in Linguistics* 19(2). 65–102.
- Nunes, Jairo Morais. 1995. *The copy theory of movement and linearization of chains in the Minimalist Program*: University of Maryland dissertation.
- Pereltsvaig, Asya. 2008. Split phrases in colloquial Russian. *Studia Linguistica* 62(1). 5–38.
- Prince, Alan & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Maiden, MA: Blackwell.
- Reinhart, Tanya. 1996. *Interface Strategies*. Cambridge, MA: MIT Press.
- Richards, Norvin. 2010. *Uttering Trees*. Cambridge, MA: MIT Press.
- van Riemsdijk, Henk & Edwin S. Williams. 1986. *Introduction to the Theory of Grammar*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, MA: MIT Press.
- Ross, John R. 1967. *Constraints on variables in syntax*: MIT dissertation.
- Sekerina, Irina A. 1997. *The syntax and processing of scrambling constructions in Russian*: City University of New York dissertation.
- Sirotinina, Ol'ga Borisovna. 1965. *Porjadok Slov v Russkom Jazyke*. Saratov: Izdatel'stvo Saratovskogo Universiteta.

- Sirotnina, Ol'ga Borisovna (ed.). 1970. *Russkaja Razgovornaja Reč'*. Saratov: Izdatel'stvo Saratovskogo Universiteta.
- Slioussar, Natalia. 2007. *Grammar and information structure: a study with reference to Russian*: University of Amsterdam dissertation.
- Smolensky, Paul & Géraldine Legendre (eds.). 2006. *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.
- Steedman, Mark. 1996. *Surface Form and Interpretation*. Cambridge, MA: MIT Press.
- Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10(1). 80–126.
- Stroik, Thomas S. 2009. *Locality in Minimalist Syntax*. Cambridge, MA: MIT Press.
- Švedova, N. Iu. (ed.). 1980. *Russkaia Grammatika. Tom 2: Sintaksis*. Moscow: Nauka.
- Szabolcsi, Anna. 2006. *Blackwell Companion to Syntax*, vol. I chap. Strong vs. Weak Islands, 479–531. Oxford: Blackwell.
- Tesnière, Lucien. 1959. *Éléments de Syntaxe Structurale*. Paris: C. Klincksieck.
- Uriagereka, Juan. 1999. *Working Minimalism* chap. Multiple Spell-Out, 251–282. MIT Press.
- Williams, Edwin S. 1974. *Rule ordering in syntax*: MIT dissertation.
- Williams, Edwin S. 1976. Underlying tone in Margi and Igbo. *Linguistic Inquiry* 7(3). 463–484.
- Williams, Edwin S. 1994. *Thematic Structure in Syntax*. Cambridge, MA: MIT Press.
- Williams, Edwin S. 2003. *Representation Theory*. Cambridge, MA: MIT Press.
- Yokoyama, Olga Tsuneko. 1986. *Discourse and Word Order*. Amsterdam: John Benjamins.
- Zemskaja, Elena Andreevna (ed.). 1973. *Russkaja Razgovornaja Reč'*. Moscow: Nauka.
- Zubizarreta, Maria Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

Thomas Jolyon Wood
University of Edinburgh
tom.wood@ed.ac.uk