Journal of Universal Language 7
March 2006, 49-89

The Word Order of Estonian: Implications to Universal Language

Martin Ehala
Tallinn University

Abstract

Estonian is most often considered a SVO language. However, in main clauses, SVX and XVS are equally frequent, which indicates that Estonian has the V2 phenomenon. In subordinate clauses, Estonian shows a quite high proportion of verb final order. The picture becomes more complex when the position of the auxiliary verb (I) in verbal complexes is taken into account: Estonian seems to have a SIOV order. Historically this order has been stable for several centuries. The results of the study cast a doubt on the generality of word order universals proposed in the literature. The article argues that instead of grammatical principles of word order, the rules of ordering given and new information are universal. It is proposed that V2 is the preferred constituent order for artificial languages and that the thematic roles rather than grammatical categories should be linguistically coded.

Key words: word order, Estonian, syntactic universals, universal language

1. Introduction

Estonian is known as a free word order language (Vilkuna 1998, ja Remmel 1968).[1] In the literary language, main clause SVX and XVS are equally frequent (about 25% and 24% respectively, see Tael 1988). This indicates that the word order in Estonian is determined by the needs of organizing known and new information rather than by the purely syntactic criteria. Although XVS seems to be mainly a feature of the standard language (in the oral discourse and dialects it occurs only about 12-19%, see Lindström 2000), most of the researchers would agree that the word order permutations in Estonian main clauses are due to V2 phenomenon (Vilkuna 1998, Huumo 1994).

This opens an interesting field of discussion over the base (or basic or unmarked) word order of Estonian. A default view is that the base word order of Estonian is SVO. As SVO is the most frequent surface word order in Estonian main clauses, and it is also the order of neutral declarative sentence, the position seems well motivated. It is also consonant with the view that the original OV order of the Finno-Ugric languages was replaced by VO order in Finnic languages, including Estonian (Kiparsky 1996). However, in subordinate clauses, Estonian shows a quite high proportion of verb final order (about 20%, see Tael 1988) and SOV order is also common in main clauses with negations (see Vilkuna 1998, Valgma & Remmel 1968). The picture becomes more complex when the position of the auxiliary verb (I) in verbal complexes will be taken into account: in sentences with verbal complexes, the auxiliary usually takes the second position and the main verb the final one. Even in subordinate clauses, the auxiliary tends to be in the second

[1] This study was supported by an Estonian Science Foundation grant No 5066. I would like to thank the anonymous reviewers of this paper whose comments and criticism enabled me to improve it considerably.

position, not at the absolute end of the sentence (see Sahkai 1999). This indicates that Estonian seems to have a typologically rare SIOV order rather than the common SOVI characteristic, for example to German that has influenced Estonian over centuries. This raises the question what actually is the base word order of Estonian and what implications it has for language typology and the theory of universal language. In order to be able to answer this question, the notion of base word order must be explicated first.

2. The Notion of Base Word Order

It seems that there are at least three possibilities of defining the notion of base word order. First, the base word order might be defined as statistically the most common word order. I call it WO_1. Second, it could be understood as the word order of a neutral declarative main clause. I call it WO_2. Third is the "underlying" word order from which the different surface word orders are derived —WO_3.

The first notion is empirical—one can discover WO_1 from a sufficiently large corpus by statistical means. In this sense the Estonian base word order is SVO as it is the most common word order. The problem with this is that XVS is almost as common in Estonian as SVO. One might wonder what validity does the claim that the base word order for Estonian is SVO have if there is another word order that is only marginally less "base" than SVO.

Or, let us look further: WO_1 for Finnish is also SVO, but the XVS pattern is much less frequent than in Estonian (14.3% against Estonian 24.6%, see Huumo (1994: 23)). Does it mean that the Estonian and Finnish base word order is the same, and the differences are just of a quantitative nature? A simple illustration can help to answer this question. Huumo (1994) provides a number

of examples of translation equivalents drawn from a corpus of Finnish-Estonian translations. The example (1) below, quoted from Huumo (1994: 26) is a good case for illustrating the difference in Estonian and Finnish word order:

(1) Fin: No jos ei ole halua kauppakäyntiin, Pietari sanoi.
Est: No kui te just ei taha kaubelda, ütles Pietari.
'But if there is no urge for trading, Pietari said/said Pietari.'

The Finnish example has a SV order in the underlined main sentence, Estonian equivalent has VS. If the difference were purely quantitative, we could easily adjust the translation by changing the word order:

(2) Est: ??No kui te just ei taha kaubelda, Pietari ütles.

As the resulting sentence (2) is not a normal Estonian sentence, the difference between Estonian and Finnish could not be purely quantitative one, but has to have a categorical foundation. If this is the case, WO_1 is not a very useful notion, except for very basic descriptive purposes. Similarly as at some level of description it makes sense to say that "the Earth is flat" as it certainly appears so for our perception, for a more adequate understanding, one should concentrate on small discrepancies that seem not to fit to the broad picture. It is the same with word order: in order to achieve an explanation for the word order patterns described above, a more generalised notion of WO is needed.

Perhaps WO_2 could be used, as it is based on native speaker intuitions and should therefore capture the essence of this language. In WO_2 sense the base word order is to be understood as the order of a neutral declarative main clause. As the example in (3a) proves, Estonian has SVO:

(3)	a.	Mart	leidis	noa.
		Mart.NOM	found	knife.GEN[2]
	b.	Metsa-s	leidu-b	hunte.
		Forest-INESS	found-3SG	wolf.PL PART
	c.	Mardi-l	on	nuga.
		Mart-ABL	is	knife.NOM

Yet in which sense a transitive sentence is more basic than for example an existential sentence (3b) or a possessor sentence (3c)? All of the examples in (3) are neutral declarative sentences and in some other languages like in English, the different thematic roles do not affect the word order in a way as in Estonian, for example *Someone found a knife* and *Someone has a knife* has the same order. To what extent we are justified to say that Estonian WO_2 is SVO, not for example XVS? It seems that in the case of Estonian, the WO_2 notion of base word order excludes so many contexts from consideration that it could not be used to characterise the Estonian word order as a whole. Yet the idea behind the notion "base word order" rests on the assumption that for each language there is a word order that is in some sense essential to this language, the order which we can use if we want to characterise the way this language linearises the constituents.

Provided the wealth of word order configurations (at least in the so called free word order languages), there is no way one can do it without a significant abstraction. WO_3 captures this idea in full, defining the base word order as the "underlying" order of constituents from which the surface variants are derived. Most explicitly, this concept is elaborated in the Chomskyan type

[2] The following abreviations of grammatical categories are used: singular–SG; plural–PL; infinitive-INF; imperfect–IMP; impersonal–IMPERS; past participle–PP; persons are indicated by numbers 1, 2, and 3; genitive–GEN; partitive–PART; illative–ILL; inessive–INESS; ablative–ABL.

generative transformational grammar (for example Chomsky 1981, 1982), but the distinction of deep and surface structures is used much more broadly here for a good empirical reason. I try to explicate this.

It is not hard to see that there are two types of principles or rules that influence the way languages linearise linguistic material. The one is grammatical and it can be characterised in terms of the ordering of heads in respect of complements, the other is discourse-configurational and it can be characterised in terms of topic, focus, verb, rest and so on. The grammatical linearization of constituents is easily detected in languages with a rigid word order. Languages with free word order have several possibilities, some of which are equally widespread. This has leaded some researchers to abandon the task of characterising the word order of such languages in grammatical terms. Instead, these languages are called "discourse-configurational" and the word order patterns are characterised in discourse-informational terms (see Kiss 1987, Vilkuna 1989). Although implicitly, this has led to a belief that there are two types of languages: 1) the languages where syntax is run by grammatical means and 2) the languages where syntax is run by discourse-informational means.

This obviously is an oversimplification. Let us take English as an example. Although English belongs to the first group of languages, it is clear that discourse-configurational rules operate in English, too—it has passivisation, dummy subjects, topicalisation and other means of configuring the given and new information. And if English has both the grammatical as well as discourse-configurational means for ordering constituents, why could we assume that there are languages (the second type) which have only one means (the discourse-configurational one). It is reasonable to assume that all languages possess both types of rules, but the one or other rule type may just be more prominent in a particular language.

If this is so, there should be a base word order for each language,

and it is the order provided by the grammar, the order that gets distorted by discourse-configurational rules in some languages more than in others. This is the base word order in WO_3 sense. To be able to find out this base word order, we need to know the purely grammatical means for linear organisation of constituents. In order to analyse this I will use the notion of X'-structure.

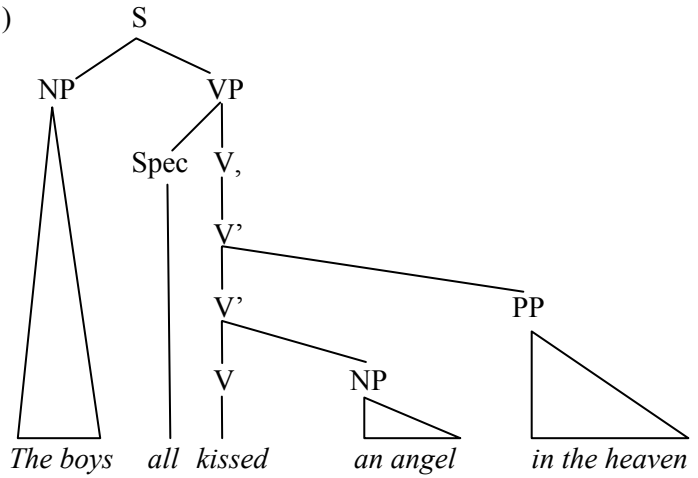
3. X'-theory

The main hypothesis of X'-theory (called x-bar theory, first introduced by Chomsky 1970, see for a survey in Stuurman 1985, and introduction in Haegeman 1991) is that all phrasal categories show in principle a single uniform structure which can be expressed by the following set of rules:

(4) XP → Specifier; X'
X' → X'; YP
X' → X; YP

The theory brings out the common properties of phrases (X stands for any lexical category, capable of forming a phrase, for example N or V). According to this theory, every phrase is headed by a lexical head that gives the name for the phrase. The head may be complemented by an YP where Y stands for any possible lexical category that can complement the given head X. There are two levels of projection, which means that the complements of X are ordered hierarchically. This can be exemplified by (5):

(5)

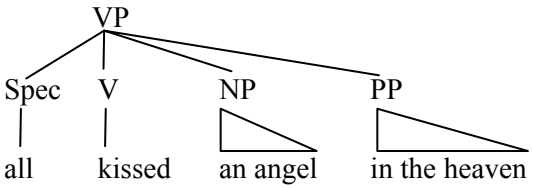


Here the prepositional phrase *in the heaven* complements the V' *kissed an angel* and the NP *an angel* complements the head of the phrase *kissed*. What this notation allows is to specify more precisely the notion of constituent. This can be shown by a simple interrogation test in (6):

(6) a. What did the boys do?
b. What did the boys all do?
c. What did the boys all do in the heaven?
d. *What did the boys all do an angel in the heaven?

Here the verb *do* can substitute the whole VP and two different V'-s but not the V. This means that XP and X' are constituents, but not X. By being able to specify different constituency levels below the phrase level, the X'-notation is superior to a flat notation like (7) which can offer no structural explanation to the phenomena described in (6).

(6)



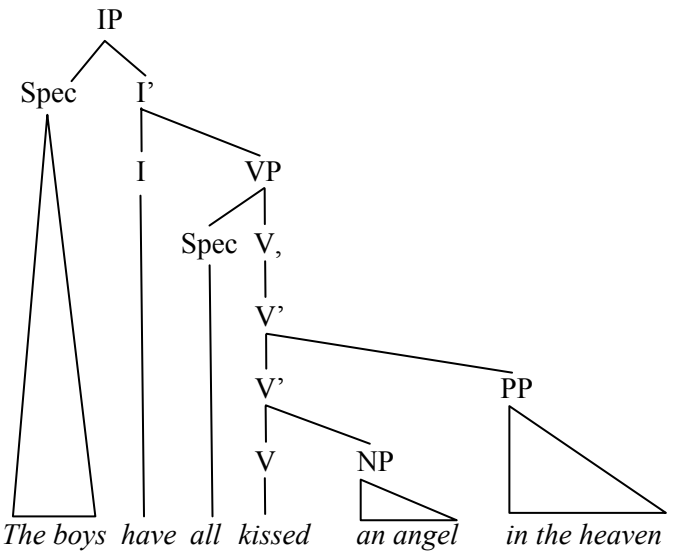
Furthermore (5), but not (7) can explain why VP **all kissed in the heaven an angel* is ungrammatical: only the direct object can be the sister of V, all other verbal complements can complement only V'. As no complement can stand in between of V and its direct object, **all kissed in the heaven an angel* is ungrammatical. The flat notation in (7) cannot differentiate structurally between different types of complements as all of them are sisters of V.

The X'-notation in (4) does not specify the order of categories. The semicolon between them (for example X'; YP) means that either order X'—YP or YP—X' is possible in principle, yet any single language chooses one or the other for each particular phrase type. As the branching is strictly binary, the number of possible order permutations is theoretically quite limited.

What makes the X'-theory particularly interesting for our purposes is that it is extended to cover the structure of clauses, too (see Chomsky 1981, 1982). This extension creates an interesting symbiosis of phrase structure grammar and dependency grammar (of a type of Tesnière 1959). In the latter, the head of a sentence is the main verb. In X'-theory the head of the sentence is not the verb but a part of it—its inflectional specification for tense and aspect. The tense and aspect are the features of a non-lexical category I. This category is often expressed by auxiliary verbs, if not, it is expressed by V that carries inflectional features. Provided that I is the head of sentence, the sentence itself is an IP and it concords to the X'-rules as does any other phrase. To illustrate the point, let us see the structure of (5) as (8) in these terms (I have changed the tense to

make it more illuminative):

(8)



In (8), the subject NP of the sentence has become a specifier of IP phrase and the VP a complement of I. Whatever one thinks of generative grammar in general, the X'-notation is an adequate generalisation of dependency structures in natural languages' syntax and has successfully stood against empirical scrutiny over several decades. For this reason it is taken in this article as a basis for specifying the WO_3 for Estonian.

4. The Position of V in Estonian Clause

Most of researchers would agree that Estonian has the V2 phenomenon, see (9):

(9) a.	Lapse-d	**söö-vad**	täna	suppi.
	children-PL	eat-3PL	today	soup.PART
b.	Täna	**söö-vad**	lapse-d	suppi.
	today	eat-3PL	child-PL	soup.PART
c.	Suppi	**söö-vad**	lapse-d	täna.
	soup	eat-3PL	child-PL	today

'Today, the children eat soup.'

The main function of V2 is to allow smooth organisation of given and new information in the sentence. Thus, it is a discourse-configurational rule not a rule of grammatical dependency as the X'-rules are. V2 can be found both in the languages that order heads before the complements (SVO-languages) as well as languages that order heads after complements (SOV-languages). In both cases the syntactic word order of the given language is blurred by the discourse-configurational rules. For example, in German and Dutch that are widely considered SOV languages in WO_3 sense, the most common word order in main clauses is SVO, in fact around 60% (Gerritsen 1984). Similarly, the Old English had also a SOV in WO_3 sense, but in the main clauses of the Anglo-Saxon Chronicle the SOV amounted from 18 to 44% in different sections of the chronicle (Bean 1983). The apparent dominance of SVO in SOV languages with V2 is due to the fact that the subject is the most likely topic. Thus, in languages that have V2, WO_3 could not be determined by the order of S, V and X in main clauses only. Instead, more subtle indicators of the base order of constituents should be taken into account.

Lightfoot (1991) argues that there are many indirect clues in main clauses indicating the base word order that has been disrupted by V2. First, according to X'-theory, heads are adjacent to their complements before they are moved to satisfy V2. Thus, in V2 language the position of objects can indicate the place where the verb was initially. Similarly, the non-finite parts of verbal complexes

and verbal particles usually do not move with the finite verbs, but remain in their initial position in dependency tree. According to Lightfoot (1991), these indicators can be sufficient for establishing SOV as the base word order in V2 languages such as German or Dutch.

According to X'-theory, the object stands next to the verb, and then the other verb complements come, thus the order is either VOX or XOV where X stands for a free modifier or a complement other than the direct object. In declarative sentences, the discourse-functional rules may easily reorder the constituents. Thus the word order is more easily detectable in infinite clauses. The order of constituents within such a phrase is not subject to discourse-configurational rules as the whole phrase is operated as a single unit at the sentence level. In (10) the infinite VP *täna suppi süüa*, is a nominal complement in a object NP of an impersonal sentence:

(10)	X	O	V
[Otsus	täna	suppi	süü-a]
decision-NOM	today	soup.PART	eat-INF
[teh-ti]	[üksmeelselt].		
made-IMPERS	unanimously		

'The decision to eat soup was made unanimously.'

That the neutral order of constituents in this deeply embedded VP is XOV, gives support to the position that the word order in finite VPs at the sentence level is also XOV. Further evidence is provided by the sentences with verbal complexes: the auxiliary verb is in the second position and the non-finite part at the end of the sentence as in (11):

(11) X O V

a. Lapse-d on täna suppi söönud.
child-PL have today soup.PART eaten

b. ???Lapsed on söönud suppi täna.
children have eaten soup today

The alternative sentence ???*Lapsed on söönud suppi täna* is very unusual and it is hard to imagine the context where this sentence might sound acceptable.

The same regularities hold for verbal particles. In (12) the verbal particle *üle* is at the final position in the sentence while the finite verb is in the second position:

(12) X O V

Lapse-d värvi-vad$_i$ täna maja üle t_i.
child-PL paint-3PL today house.GEN over
'The children will overpaint the house today.'

It is reasonable to assume that *üle värvima* 'to overpaint' is a single lexical unit which is located at the V node. Thus, when the verb is moved to the second position, the particle is left behind. That the verb is moved, not the particle is confirmed by changing the tense from present to perfect as in (13). The alternative ***Lapsed on üle värvinud maja täna*** is ungrammatical, not just unusual.

(13) Lapse-d on täna maja üle värvinud.
child-PL have today house.GEN over painted
'The children have overpainted the house today.'

The examples above indicate that V2 in Estonian main clauses is derived from a base SOV rather than SVO. However, things are not that straightforward. The XOV order in VP is prevalent only in cases where X is a free modifier. Verbs that require another

obligatory complement besides the direct object have different order in VP. Although the verb is at the final position, it is not preceded by the direct object, as would be predicted by X'-theory, but by its other complement as in (14):

(14)

		O	X	V
Ema	on	lapse-d	kooli	viinud.
mother	has	child-PL	school.ILL	taken

'The mother has taken the children to the school.'

In fact, as much as 42,5% of the bound complements follow the direct object in Estonian (Sahkai 1999). In sentences that have simple verbs and bound complements, the V2 phenomenon creates an appeareance of SVO, as in (15):

(15)

S	V	O	X
Ema	viis	lapse-d	kooli.
mother	took	child-PL	school.ILL

'The mother took the children to the school.'

The question arises, how one can be sure that the pattern in (15) is actually not an instance of genuine SVO in Estonian. Here a parallel with German would be helpful. In her work on word order in Estonian VP, Sahkai (1999) notes that the linear order of verbal complements in Estonian sentences closely resembles that of German. In this language the constituents in the sentence follow the pattern: subject—finite verb—free adjuncts—direct object—bound adverbial (directionality)—verbal particle—infinitive or participle (Helbig & Buscha 1994: 583). The order of Estonian verbal complements is finite verb—lative possessor adverbial—free adverbial—object or predicative—bound adverbial—verbal particle —infinite part of a complex predicate (Sahkai 1999: 29). Both languages question the validity of the X'-principle that objects stand

next to the verb in VP.

In the case of German, two solutions are proposed to eliminate the discrepancy with X'-theory. Müller (1999) suggests that the surface OXV order is derived by scrambling the direct object to the left of the adverbial (XOV → OXV). Fanselow (2003) proposes that the constituent order should be relaxed, so that different word order configurations in free word order languages need not be derived by movements, but can be base generated. According to this proposal, free order of constituents is universal; languages with a rigid word order just need a special device to eliminate the possibility of multiple configurations. Be as it may, neither of the authors has questioned that German is characteristically a SOV language, despite the fact that some bound adverbials can separate the direct object from the verb. As the word order in Estonian VPs is very similar to that of German, we should not rule this possibility out for Estonian either.

But let us look for additional evidence for WO_3 in embedded clauses. According to X'-theory, the base word order is the same in main clauses as well as in subordinate clauses. In subordinate clauses V2 is impossible, as the structural position where the verb has to move for V2 to be created is often filled with a complementiser. In these cases the verb will remain in its underlying position, providing thus another clue for specifying the WO_3 of the language (see Lightfoot & Hornstein 1994). This method works well for German or Dutch, but not equally well for Estonian. The problem is that in Estonian like in Icelandic and Yidish (see Vikner 1994), V2 is possible also in embedded clauses with overt complementisers, although not in all clause types. For example in *if*/*when*-clauses, the verb is in the final position (C stands for the position of a complementiser):

(16) C S X O V
… kui lapse-d lõpuks supi ära söö-vad.
… if child-PL finally soup.GEN away eat-3PL
'… if the children will finally eat the soup.'

In *that*-clauses V2 is permitted, as show (17a) and (17b). (17c) indicates that the base position of the verb is at the end of the sentence:

(17)
a. … et lapse-d söö-vad täna suppi.
… that child-PL eat-3PL today soup.PART
'… that children will eat soup today.'
b. … et täna söö-vad lapse-d suppi.
that today eat-3PL child-PL soup.PART
c. … et lapse-d on täna suppi söönud.
that child-PL have today soup.PART eaten
d. … et ema on lapse-d kooli viinud.
… that mother has child-PL school.PART taken
'… that mother has taken the children to the school.'

Thus, it seems that V2 is quite normal in subordinate clauses with the complementiser *that*. And as (17c) and (17d) indicate, the word order in VP is XOV if X is a free modifier, and OXV when it is a bound complement. From the examples above it can be concluded, contrary to the common belief, that Estonian is a V2 language that has verb final VPs, and consequently the SOV base word order. Therefore it seems that the Estonian word order is much closer to the German word order than to the Finnish one. Certainly there are also differences between the Estonian and German word orders, one of the most interesting of them being the position of auxiliary verbs, or to use X'-terminology, the position of the functional head I of IP.

5. The Position of I in Estonian

As about the position of the auxiliary or I in respect of S, V, and O, Greenberg's (1966) formulates its 16th universal, repeated here as (18):

(18) In languages with dominant order VSO an inflected auxiliary precedes the main verb. In languages with dominant order SOV, an inflected auxiliary always follows the main verb.

Taking into account only the OV order in respect of auxiliary, four logical possibilities arise. All of the four types have been also attested, but the distribution of languages between types is significantly uneven. The table 1 illustrates this distribution in Dryer's (1992) genera:

Table 1. Auxiliary Orders in Dryer's (1992) Genera

Type	Order	Number	Percentage
A	[Aux [V NP]]	28	39.4%
B	[[NP V] Aux]	36	50.7%
C	[Aux [NP V]]	3	4.2%
D	[[V NP] Aux]	4	5.6%
Total		71	

According to Table 1, there are two typologically common types of auxiliary order (A and B) that account about 90% of languages and two rare types (B and C) that account for the rest of 10%. I assume that types A and B represent unmarked word order possibilities and C and D the marked possibilities. Good examples of type A and B languages are English and German. English is a

well known SVO language that has auxiliary preceding the main verb (it has SIVO order in X'-terms, see (19a)); German is a SOV language that has auxiliary following the main verb (SOVI order, see (19b)):

(19) S I V O

a. Father has said to me that mother has taken the children to the school.

S O

b. Der Vater hat mir gesagt daß die Mutter die Kinder in
V I
die Schule gebracht hat.

Provided that Estonian has SOV, as argued in this paper, two possibilities are open: either Estonian is like German SOVI or it belongs to the rare type C and has SIOV. It is not easy to determine the right position of I in Estonian, as Estonian has V2 in both main and subordinate clauses. In sentences with V2, the auxiliary is in the second position of the clause, but this position is not its original position. For example in (19b), the auxiliary *hat* is at the second position of the main clause; giving an impression that German belongs to the marked SIOV type. As I argued above, the V2 rule is a discourse-configurational rule that distorts the structural order of constituents in order to satisfy the needs of organising the given and new information. Thus one cannot count for sentences with V2 to specify the position of auxiliary in the WO_3 sense.

Fortunately not all Estonian subordinate clauses have V2. Remmel (1963) argues that in Estonian, there are two types of subordinate sentences which differ from each other by communicative weight. The one type he calls "ordinary subordinate clauses", the other "subordinate clauses with the function of a main clause" (Remmel 1963: 244-245). Ordinary subordinate clauses are communicatively less prominent than main clauses, they do not

provide new information, but just give some necessary details, needed to specify the content of the main clause. In ordinary subordinate clauses V2 is not possible and the main verb is at the final position in the sentence. The subordinate clauses that have the weight of a main clause provide the main bulk of the new information while their main clause only specifies the conditions. For example, *that*-clauses often belong to this class, particularly if the main clause gives the frame for an indirect quotation.

The example in (20a) provides an ordinary subordinate sentence with a neutral order. (20b) is an ordinary subordinate clause that has an unacceptable V2 with inverted VS order. Examples in (20c and d) show that V2 is possible in subordinate clauses that have the function of the main clause:

(20)

a. See juhtu-s pärast seda, kui valitsus
this happen-IMP after that, when government
Eesti-s palju töö-d oli teinud.
Estonia-INESS much work-PART has done

b. *See juhtu-s pärast seda, kui Eesti-s
this happen-IMP after that, when Estonia-INESS
oli valitsus palju töö-d teinud.
has government much work-PART done
'This happened after the government has done much work in Estonia.'

c. Ta ütle-s, et valitsus on Eesti-s
he say-IMP that goverment has Estonia-INESS
palju töö-d teinud.
much work-PART done

d. Ta ütle-s, et Eesti-s on valitsus
he say-IMP that Estonia-INESS has goverment
palju töö-d teinud.
much work-PART done

‘He says that the governent has got much work done in Estonia.’

Thus, the ordinary subordinate clauses that do not allow V2 can be used to determine the position of I in respect of S, O and V. Let us compare the following sentences:

(21) a. … kui lapse-d on lõpuks supi ära söönud.
… if child-PL have finally soup.GEN away eaten
‘… if the children have finally eaten the soup.’
b. … kui lapse-d lõpuks supi ära on söönud.
… if child-PL finally soup.GEN away have eaten
c. ??… kui lapse-d lõpuks supi ära söönud on.
… if child-PL finally soup.GEN away eaten have

(21a and b) provide two possible sites for an auxiliary. In (21c) the auxiliary is at the final position in the sentence like in German, but the whole sentence sounds awkward because of this. In this respect Estonian is clearly different from German, and so the SOVI order is a dubious option for Estonian. As (21a) is the most neutral, it is possible that Estonian might have a marked SIOV word order. The situation is even more complicated, as in (21b) the auxiliary intervenes between V and its complements. This appears to be challenging to the word order typology since the SOIV type is not even mentioned as a possibility.

Actually, Estonian word order is not that exceptional as it might first appear, as both these rare orders occur also in Dutch, West Flamish and in some German dialects. Furthermore, both SIOV and SOIV can be handled in generative grammar using different types of movement rules. This would suggest that if Estonian is similar to the languages just mentioned, it might use the same movement rules and consequently have the same WO_3, too, namely SOVI. Let us look at this possibility in a more detail.

For example in German there is a phenomenon known as *oberfeldbesetzung* (Bech 1955) where 3-verb clusters have the auxiliary in the left of its verbal complements (22a) rather than in the right (22b):

(22) a. dass er das Buch wird lessen müssen
b. ??dass er das Buch lesen müssen wird

A similar phenomenon is also found in Dutch (23a) and West Flemish (23b, all Dutch and West Flemish examples from Haegeman 1998):

(23) a. dat Jan een huis wil kopen
that Jan a house wants buy
b. da Valère em een brief ee zien schrijven
that Valère him a letter has see write
'that Valère saw him write a letter'

In all of these examples auxiliary (or finite verb) intervenes between main verb and its object, resulting in a SOIV order. Traditionally this contradiction to the X'-theory is solved by assuming a SOVI base order and the so called Verb Raising—a rightward movement of the main verb over the I node. There is a vast literature on Verb Raising (see for example Besten & Rutten 1989, Haegeman & Van Riemdijk 1986, Haegeman 1994) and I will not go into its details here, I note only that a similar solution might be a possibility for Estonian too, if Estonian had a SOVI order. This depends on whether the surface SIOV order is derived or not.

For Dutch and West Flemish (WF) the surface SIOV is assumed to be derived from an underlying SOVI (see Haegeman 1998). For example in WF *Infinitivus Pro Participio* (IPP) constructions the finite auxiliary may be at the end of the clause or after the Subject:

(24) a. da Valère [$_{XP}$ willen Marie dienen boek geven] eet
that Valère want Marie that book give has
b. da Valère ee [$_{XP}$ willen Marie dienen boek geven]
c. [$_{XP}$ INF1 – (XP) – INF2] Aux$_{perf}$
d. Aux$_{perf}$ [$_{XP}$ INF1 – (XP) – INF2]

The order in (24a, c) is assumed to be the base SOVI order. The order in (24b, d) is achieved by an operation called extraposition which is a rightward movement of the whole complement of the auxiliary. In WF, extraposition is triggered by certain morphological features of the finite verb. (cf., Haegeman 1998)

Extraposition is a sensible solution to SIOV surface order in case this order is in some sense a deviation of the dominant SOVI order, and the conditions of its appearance can be defined. Thus, in order do decide whether Estonian uses SOVI and extraposition or SIOV, we should analyse the distribution of SOVI and SIOV in Estonian in a more detail and compare it with the similar phenomenon in WF. Let's consider the following examples:

(25) a. kui Jaan on midagi söö-nud
if Jaan has something.PART eat-PP
'if Jaan has eaten something'
b. kui Jaan midagi on söö-nud
if Jaan something.PART has eat-PP
c. kui Jaan midagi söö-nud on
if Jaan something.PART eat-PP has

(26) a. kui Jaan on midagi süü-a taht-nud
if Jaan has something.PART eat-INF want-PP
'if Jaan has wanted something to eat'
b. kui Jaan midagi süüa on taht-nud
if Jaan something.PART eat has want-PP
c. ??kui Jaan midagi süüa taht-nud on
if Jaan something.PART eat want-PP has

(27) a. et Jaan peab seda tööd teha taht-ma
that Jaan must this job do.INF want-INF
'that Jaan must want to do this job'
b. et Jaan seda töö-d pea-b teha taht-ma
that Jaan this job-PART must-3SG do.INF want-INF
c. ?et Jaan seda töö-d teha pea-b tahtma
that Jaan this job-PART do.INF must-3SG want-INF
d. *et Jaan seda töö-d teha taht-ma pea-b
that Jaan this job-PART do.INF want-INF must-3SG

(28) a. et Jaan on pida-nud seda
that Jaan has must-PP this
töö-d teha taht-ma
job-PART do.INF want-INF
'that Jaan must have wanted to do this job'
b. ?et Jaan seda töö-d teha taht-ma
that Jaan this job-PART do.INF want-INF
on pida-nud
has must-PP
c. *et Jaan pida-nud seda töö-d
that Jaan must-PP this job-PART
teha taht-ma on
do.INF want-INF has

The examples in (25)-(28) indicate that the auxiliary can only be at the sentence final position if the sentence has one verb. If a clause has a verbal cluster of two or more verbs, auxiliary cannot occur at the final position, it preferably occurs immediately after S or immediately before the main verb.

As SOVI is the least likely surface order for sentences having auxiliaries in Estonian, it would be hard to argue for SOVI base word order on either empirical or acquisitional grounds. First, unlike in WF, in Estonian, there are no morphological conditions that

trigger Verb Raising and extraposition, as both of these operations are almost obligatory in Estonian. Rather, the SOVI order seems to be somehow marked, as it sounds natural only if the auxiliary is stressed, i.e., the main verb is focused. Second, as SIOV and SOIV are also statistically more prominent, it would be hard to imagine how the SOVI base order could be acquired. Considering the data and theoretical arguments it is much more likely that Estonian has SIOV and the SOIV is achieved by scrambling in Middle Field, i.e., moving O to the left of I.

The other possibility for explaining the SOIV order is that the auxiliary is cliticized to the verb, indicating that it has become or is becoming a verbal prefix. Some support to this hypothesis comes from the fact that the cliticized auxiliary can occur between the verbal particle and the main verb, see for example (21b).

If Estonian has SIOV as the data suggest, it is different of both Finnish that could be considered a SIVO language (Holmberg et al. 1993) and German, a SOVI language. Estonian is like a kind of mixture of Finnish and German type. In this case it would be interesting to know how this word order may have emerged.

6. Changes in Word Order of 20th Century Estonian

A common assumption is that Uralic languages have historically had SOV word order, which at present has been lost in several of the languages, including the Finnic. As for Finnish the assumption seems to be tenable, but the situation is more complex for Estonian. One possibility is that Finnic changed from Proto-Uralic SOV to SVO already before splitting to sister-languages. Thus, Estonian was initially a SIVO language just like Finnish, but the German influence that lasted around 700 years turned Estonian half-way back again resulting in the current SIOV.

The second possibility is that the German influence caused Estonian to change its word order completely to SOVI until Johannes Aavik who led an extensive language renewal campaign got it turned half-ways back. (One of the goals of Aavik was to replace the German-like word order that was extremely common in the beginning of the 20th century by moving the verb forward. For Aavik's reforms see Tauli (1983)).

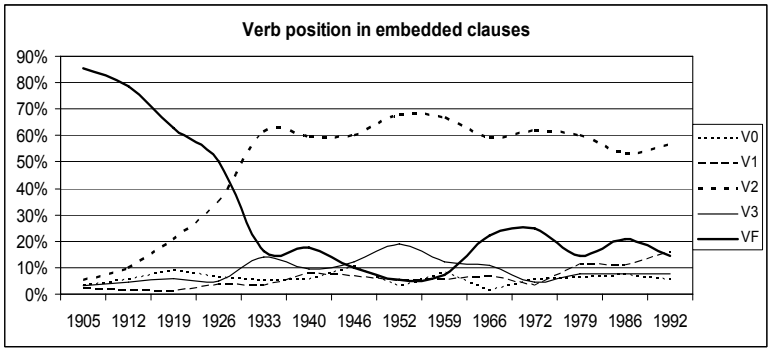
The third possibility is that Estonian has never been a SVO language and that it has basically retained its original base SIOV order for all times, except that the German influence might have caused some increase of SOVI order in the 19th century.

I do not have a definite answer to this question, but let us see what the data suggest. In 1996 until 2000 I and my students created a database of Estonian newspaper language of 20th century. The database consists of 14 small corpora, each of 600 clauses which were analysed for surface word order and verb position.

First I'll tackle the base word order problem. If we look at the dynamics of verb position in the Estonian embedded clauses during the last century we see that at the beginning of the century the final position for the main verb was absolutely dominant in the embedded clauses. Then the situation changed quite rapidly in the first half of the century which quite certainly was due to Johannes Aavik's language renewal campaign. Thus, by the end of thirties the VF embedded clauses have dropped to 20% and so they have remained (see Figure 1).

It can be argued (Ehala 1998) that this change constituted the change from the German influenced SOV base order to the present SVO. Although the change Aavik initiated was quite rapid and extensive in written language, it cannot be taken as evidence for a change of base word order for two reasons. First, the base word order change would cause changes in the constituency order in main clause VPs, too (see Battye & Roberts 1993), but this has not happened, as show examples in (10)-(13). Second, the change has

Figure 1. Word Order Change in the 20th Century Estonian



not happened in oral language, where V2 and VF are of approximately equal frequency in the embedded clauses (around 35%), although educated speakers tend to use more V2 and less VF than average (Käi 1997).

This suggests that no base word order change has taken place in modern Estonian, and what actually happened is a rise in V2-movement in subordinate clauses due to the fact that the verb final order in subordinate clauses got stigmatised in literary language and well educated users deliberately employ movement operations to hide the property, mainly in writing, but to a lesser extent also in speech.

Let us see what can be said about the position of auxiliary verb: has Estonian always had a SIOV order or was it SOVI at some times.

To look for the position of I in respect of V and O, the Tartu University untagged corpus of Estonian, available at http://www.cl.ut.ee/ee/corpusb/ was used. From the corpus of Estonian of 1900, 100 random embedded clauses were collected that contained the auxiliary verb form *oli* 'was' in a verbal complex. The most frequent (79% of the total) was the SOIV order. The second was the SOVI order with the frequency of 16%. There were also a few instances of SIOV and SIVO order.

Table 2. The Position of I in 1900

order	number	percentage
SOIV	79	79%
SIOV	3	3%
SIVO	2	2%
SOVI	16	16%
total	100	100%

Bearing in mind that Estonian was heavily influenced by German at the beginning of 20th century, it is significant that the SOVI order is represented only as a minor order in subordinate clauses. This suggests that SOVI is not characteristic to Estonian and the base order is either SIOV or SOIV. What the table suggests is that there was a kind of conflict between the Estonian original base word order and the German SOVI, so that the most dominant, but typologically unattested order SOIV is a compromise between the original SIOV of Estonian and German SOVI.

Another possibility is that Baltic German spoken in Estonia, did not concur with the word order of Standard German, but carried some dialectal features that are now reflected in the Estonian word order. As most of the Estonian colonisers from 13th-15th century came from the area of Low-German, mainly from Westfalen, Holstein, Mecklenburg and Pommer that are related to Dutch and West Flamish, it is likely that the variant of German spoken in Estonia might have SIOV and SOIV word order patterns. However, after reformation the Low German variety spoken by Baltic Germans was replaced by Standard German to the extent that in early 20th century SOVI was the only word order in written newspaper language. From a hundred subordinate clauses having either an auxiliary or a modal, collected from a random issues of the *Revalsche Zeitung*'s 1900 volume, the 100% had a SOVI order. This is in a sharp contrast with Estonian newspapers that had SOVI only 16%. Thus it seems that Estonian word order is not a reflection of

Standard German, but might on its early stages have got a Low German influence.

7. Word Order in Early Estonian

As about the early Estonian, there is not much evidence available. The earliest documented samples of Estonian go back to 13th century, but these are just a few phrases. The number of written Estonian starts to grow from 16th century onwards. But this is the period when the Baltic German dialect shifted from Low German to Standard German.

To look for the evidence of Estonian word order, I checked the Tartu University corpus of old literary Estonia of 16th century, available at http://ee.www.ee/filosoft/wakk/. The corpus is small because of lack of texts, and even these texts have mostly been written by non-Estonians. Thus, these data may not be very reliable.

I collected 76 clauses from this corpus and analysed them as for the surface word order and verb position. The results are presented in Table 3.

Table 3. The Position of V in 16th Century

Verb Position	Number	Percentage
V1	2	4.5%
V2	13	29.5%
V3	6	13.6%
VF	23	52.4%
TOTAL	44	100%

The word order is very close to the German SOVI, 6 clauses which have VF order have also a verbal complex and in all of them the auxiliary is at the absolute end of the clause, for example:

(29) k[uth] J[u]mal sen essymesen ynymysen l[onu]th oly
when God this first man created had

There is also a clause from an informal private letter which has V2:

(30) et synnu-la pehe-p keick assia Jella same
that you-ADESS must-3SG all things again become
'that you must get all the things again'

To summarise, the 16th century written Estonian has clear indications of SOVI, but the word order generally is very variable which makes it hard to draw any firm conclusions. The fact that the early written Estonian has more indicators of SOVI than the language of early 20th century seems to suggest that this feature was brought to Estonian texts by non-native speakers. When Estonians themselves started to produce written texts, this SOVI decreased, despite the fact that these native Estonians had obtained their education in German and knew German very well.

Besides written texts, it is likely that the ancient word order has survived also in parables, idioms and folk verse. For this, the word order in Estonian folklore was studied. First, the word order in runo-songs was targeted. It is obvious that one should be careful in drawing conclusions from such material because of *licentia poetica*. As the language of runo-songs does not contain many complex sentences, and surface word order is quite often ambiguous as about the base word order, it is no surprise that good evidence is hard to find. In the main clauses there are a number of clear cases of V2:

(31) a. Siis tuli pukki Pohjamaa-lta.
then came ox North-ABL
b. Sinna kasvi-s suuri saari.
There.ILL grow-IMP 3SG large ash tree

c. Meil aga kasva-s suuri saari.
we.ADESS but grow-IMP 3SG large ash tree
'A large ash tree grew at our place.'

In subordinate clauses, the SOV order is quite common:

(32) a. …kus mina laulu-d lahuta-si-n
where I song-PL separate-IMP-1SG
b. …kuni ma kodu-je jõua-n
until I home-ILL arrive-1SG
c. …*kus meie viimaks või-da sõ-i-me*
where we finally butter-PART at-IMP-1PL
d. …*kuhu mina noa unust-i-n*
where I knife.GEN forget-IMP-1SG

It is harder to say anything about the position of I, since the verbal complexes are not common at all. I found no cases of SOVI and only one case of SIVO, however in this sentence the metrum requires the particular order (there is no other possible position for the monosyllabic *põld*):

(33) …kes põld põlves põldu kün-nud
who hadn't ever field-PART plough-PP

Remmel (1963) presents the results of his quantitative study of subordinate clause word order in Estonian folk tales and parables. He counted subordinate clauses from 28 random pages of "Eesti rahvanaljandid" (Põldmäe 1941), a collection of folk tales from different counties in Estonia. From among 171 subordinate sentences in 152, the verb is at the end of the sentence or precedes the infinitive. "Valimik eesti vanasõnu" (Normann 1955) is a collection of Estonian parables. From among 381 subordinate sentences containing a simplex verb, 322 had a verb final order. In

59 cases the verb is not in a final position. From among these 59, 19 cases contain juxtaposition which distorts the neutral order and in 9 cases, the rhyme requirements are obviously influenced the order and in some cases the rhythm (Remmel 1963: 248).

Considering comparative evidence from other Uralic languages, it must be admitted that the family is notoriously divergent as about word order. There are languages with consistent SOV order like Mari and Udmurt as well those that have a flexible SVO order, mainly Finnic (Vilkuna 1998). As about the position of auxiliaries, Mari and Udmurt have SOVI which is the typologically unmarked option. Interestingly, in Southern Saami, auxiliaries take a medial position in sentence yielding SIOV (Trosterud 1996, Vilkuna 1998). Trosterud (1996) argues that there is a word order continuum from Southern Sami to Northen Sami, the later having a consistent SIVO order. Pitesaami and Lulesaami show variation between SIOV and SIVO. This continuum is a reflection of the diachronic process of the change "von der ururalischen Komplement-Kopf-Struktur zu einer westeurasiatischen Kopf-Komplement-Struktur" (Trosterud 1996: 109). Undoubtedly Estonian also takes a position in this continuum, and if my argumentation is correct, it takes a position at the rather conservative end of the scale.

8. Implications to Word Order Typology and to the Universal Language

First of all, the controversy over the Estonian word order is a vivid example of the unreliability of word order typology. In the literature, Estonian is widely believed to be a SVO language. Probably it is also indexed in some typological databases as such. Yet SVO in Estonian is just a consequence of discourse-configurational principles, namely of V2. The actual word order in

Estonian shows a large amount of variability, the deeper analysis of which reveals that the grammatical order of constituents in Estonian is SIOV.

Typologically, SIOV is considered a rare and marked order. However, it is not clear to what extent the markedness of SIOV is real. Odden (2003) argues thoroughly that the current crosslinguistic databases are not statistically reliable for making claims about linguistic universals, as they cover only a relatively small amount of nonrandomly chosen languages whereas most of linguistic diversity remains undescribed or underdescribed. For example, it is well known that German is SOVI, but it is not so known that several German dialects have SIOV, as does West-Flemish, and of course, Estonian. It is not known how many languages are inadequately described in databases that are made to draw conclusions on linguistic universals. Thus, the markedness of SIOV may not be as evident as it has been claimed.

Also a common claim is that the marked word order types are diachronically unstable, being just transitionary stages from one unmarked and stable type to the other (Vennemann 1974, 1981). The diachronic analysis of Estonian word order shows that SIOV has been stable in Estonian over several centuries, having shown no tendency to change even under persisting influence of German or in the course of radical language planning campaign.

If we look at word order studies theoretically there have been proposals from two absolute opposites. In the tradition of generative grammar, Kayne (1994) has proposed that underlyingly all languages have the same order of head and complements, namely that complements follow the heads. This proposal is known as Universal Base Approach and it is currently widely accepted amongst generative linguists. For the word order typology, the Kayne's proposal would mean a universal SVO order for all languages.

On the other hand, Dryer (1997) has argued that there are no

word order universals, grammatical relations are language particular. This means that cross-linguistic notions are just handy conventions that enable to create order in chaos. Although this view is not very widely accepted amongst typologists, there is evidence that languages do not use word order universally for expressing grammatical relations. For example LaPolla (2002) argues that even if there are apparent similarities, languages actually use word order in very different ways and to different extent to express grammatical relations. He shows that for example Chinese could not described properly using the notions of subject and object, thus it does not make much sense to describe its word order regularities in this manner. Instead, Chinese word order could better be characterized as "topical and non-focal NPs occur preverbally and focal or non-topical NPs occur post-verbally" (LaPolla 1995: 10).

Actually the same regularity occurs in Estonian and in many other languages being known as V2. I argued in the first section of the paper that V2 is a discourse-configurational rule that organizes the known and new information in the sentence. It is a fact that all languages organize given and new information in the discourse. This is absolutely necessary as text is linear and the constituent parts of an utterance can only be presented piece by piece. Thus, it seems that discourse-configurational principles are far more universal than grammatical word order regularities. Whereas grammatical relations can be expressed either morphologically or syntactically, discourse-configurational principles pertain necessarily to linear alignment of constituents, i.e., to word order.

Let us see what this generalization means to the theory of universal language. If we assume that linear ordering is the main and unmarked means for organizing given and new information, the ordering of the constituents in the clause level should follow this principle. Undeniably the most flexible principle for configuring information in the sentence is V2. However, if V2 is used, the order of constituents is free which raises the question of how to signify

grammatical relations. This cannot be done by syntactic means as the variable position of constituents does not allow determining their grammatical roles.

A large number of languages using V2 have also rich morphology. Thus, a morphological ending can reliably signify the thematic role of the constituent whatever position in the sentence it appears. This could be a solution to universal language, too, but it is also well known that languages with rich morphology are hard to learn; also the core morphological cases (nominative, accusative, genitive, ablative) signify grammatical relations not semantic roles. Thus, a morphological system would be a real disadvantage for a universal language.

Furthermore, there is no need for a language user to know which constituent is a subject which is an object. The reason is that both subjects and objects can fulfill very different semantic roles. For example in *I sing*, the subject is an actor, in the sentence *I have a knife*, it is a possessor, in a sentence *I was seen*, it is a patient. As we have seen, in the case of Estonian, the same semantic roles can be expressed by different syntactic categories. What is interesting, is that languages with rich morphology have two types of morphological cases—the structural cases (Nominative, genitive, accusative, or ergative, absolutive) and semantic cases (like inessive, ablative, and so on). The latter cases are structurally very much like adpositions (pre- or postpositions) in a sense that they add a simple well-definable meaning to the word they attach to whereas the structural cases fulfill multiple functions—they signify syntactic relations and to some extent also semantic roles (like genitive). The intriguing question is—to what extent it is necessary at all to know the syntactic relations between constituents. For understanding, it is necessary to know thematic roles, if they were signified, grammatical relations at the sentence level become in fact redundant.

What I am suggesting is that a universal language would do without syntactic marking of subjects and objects if there is some

sort of marking for thematic roles. For example, assuming that the sentence is organized by V2 and constituents are marked for semantic roles, the hypothetical sentences in (34) would allow a free configuration of topical and given information:

(34) a. Actor-John kiss Patient-Mary In-morning
'John kisses Mary in the morning.'
b. Actor-John kiss In-morning Patient-Mary
'It is Mary who is kissed by John in the morning.'
c. Patient-Mary kiss Actor-John In-morning
'It is in the morning when Mary is kissed by John.'
d. Patient-Mary kiss In-morning Actor-John
'It is John who kisses Mary in the morning.'
e. In-morning kiss Actor-John Patient-Mary.
'In the morning, John kisses Mary.'
f. In-morning kiss Patient-Mary Actor-John
'In the morning, Mary is kissed by John.'

The question is in what form the thematic roles would manifest in the universal language. There are two possibilities, either they are expressed by morphological markers or alternatively by prepositions or postpositions. As the number of semantic roles is likely to be larger than the number of cases in most common morphological systems in world's languages, the morphological system of the universal language would be rather marked. Thus, thus expressing the semantic roles by prepositions is much more preferable.

Although the sentence level constituent order can be modeled by discourse-configurational rules and it is possible to free the structure of universal language from the notions of subject and object, the problem of constituent order is not disposed yet: the constituents are ordered also in the phrasal level where the discourse-configurational rules do not apply. Basically, this is the question of the order of heads in respect to their complements.

Already in sixties Yngve (1960) has proposed on the basis of English data that there is a universal bias against the depth of left-branching constructions. His hypothesis has found a partial confirmation on a corpus based study (Sampson 1997). Recent psycholinguistic studies (Kempler et al. 2004) have shown that left-branching constructions impose a greater burden on working memory during production. From this it follows that the preferred structures on the phrasal level should be right-branching, i.e the head-complements order.

Thus, the results of this study would suggest that the preferred word order structure for a universal language would be V2, the thematic roles usually signaled by word order or by inflectional morphology would be signified by prepositions. The word order in subsentential level would be head-complement.

9. Conclusion

Estonian word order is anything but straightforward, if not the notion of free word order is employed to get rid of the need to search for regularities. In this paper I have been relying on the assumption that there is a distinction between the rules of grammar and the rules of pragmatics. By eliminating the latter we could get to the core of grammar in order to see what the essential principle is how a language organizes its constituents. As about Estonian, its WO_3 is to a substantial degree obscured by V2, so that main clauses can provide only indirect clues. The subordinate clauses where discourse-configurational rules operate less freely provide a better grasp to the regularities of base word order, and it appears that Estonian has SIOV base word order. Historically this order has been stable in Estonian, despite lasting German influence and deliberate attempts in the beginning of 20^{th} century to change the word order.

Further theoretical analysis revealed that at the sentential level, the discourse configurational rules seem to be more universal than the grammatical principles of ordering constituents. This has direct implications to the theory of universal language—the preferred word order at the sentence level is V2 which is the optimal linearization strategy for ordering given and new information. At the phrasal level the preferred order is head-complements.

If word order is used for discourse configurational purposes, the relations between constituents need to be signaled by other means than word order. The present article proposes that the thematic relations between constituents be signaled by prepositions. This proposition also raises an interesting question for the future research: if thematic roles and information structure are linguistically expressed, is there any need for the notions of subject and object in an artificial language at all?

References

Battye, A. & I. Roberts. 1993. Introduction. In A. Battye & I. Roberts (eds.), *Clause Structure and Language Change* 1-30. New York & Oxford: OUP.

Bech, G. 1955. *Studien über das Deutsche Verbum Infinitum*. Niemeyer: Tübingen.

Besten, H. den & J. Rutten. 1989. On Verb Raising, Extraposition, and Free Word Order in Dutch. In D. Jaspers, W. Klooster, Y. Putseys, & P. Seuren (eds.), *Sentential Complementation and Lexicon. Studies in Honour of Wim de Geest* 41-56. Dordrecht: Foris.

Bean, M. 1983. *The Development of Word Order Patterns in Old English*. London & Canberra: Croom Helm.

Chomsky, N. 1970. Remarks on Nominalization. In R. Jacobs & P. Rosenbaum (eds.), *Readings in English Transformational Grammar* 184-221. Waltham, MA: Ginn.

_____. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

_____. 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.

Dryer, M. 1992. The Greenbergian Word Order Correlations. *Language* 68, 81-138.

_____. 1997. Are Grammatical Relations Universal? In J. Bybee, J. Haiman, & S. Thompson (eds.), *Essays on Language Function and Language Type* 115-143. Amsterdam & Philadelphia: John Benjamins.

Ehala, M. 1998. How a Man Changed a Parameter Value: The Loss of SOV in Estonian Subclauses. In R. Hogg & L. van Bergen (eds.), *Historical Linguistics II: Germanic Lnguistics* 73-88. Amsterdam: John Benjamins.

Erelt, M., R. Kasik, H. Metslang, H. Rajandi, K. Ross, H. Saari, K. Tael, & ja S. Vare. 1993. *Eesti keele grammatika II*. Tallinn: TA Keele ja Kirjanduse Instituut.

Fanselow, G. 2003. Free Constituent Order: A Minimalist Interface Account. *Folia Linguistica* 37, 191–231.

Gerritsen, M. 1984. Divergent Word Order Developments in Germanic Languages: A Description and a Tentative Explanation. In J. Fisiak (ed.), *Historical Syntax*. Mouton de Guyter: Amsterdam.

Greenberg, H. 1966. *Universals of Language*. Cambridge: The MIT Press.

Haegeman, L. & H. van Riemsdijk. 1986. Verb Projection Raising, Scope, and the Typology of Rules Affecting Verbs. *Linguistic Inquiry* 17, 417-66.

_____. 1991. *Introduction to Government and Binding Theory*. Oxford: Basil Blackwell.

_____. 1994. Verb Raising as Verb Projection Raising: Some Empirical Problems. *Linguistic Inquiry* 25, 509-522.

_____. 1998. V-positions and the Middle Field in West Flemish. *Syntax* 1.3, 259-299.

Helbig, G. & J. Buscha. 1994. *Deutsche Grammatik*. Leibzig: Langenscheidt.

Holmberg, A., U. Nikanne, I. Oraviita, H. Reime, & T. Trosterud. 1993. The Structure of INFL and the Finite Clause in Finnish. In A. Holmberg & U. Nikanne (eds.), *Case and Other Functional Categories in Finnish Syntax* 176-206. Berlin: Mouton de Gruyter.

Huumo, T. 1995. Näkökulma Suomen ja Viron sanajärjestyseroihin. In H. Sulkala & H. Laanekask (eds.), *Lähivertailua 8*, 21-39. Oulu: Oulu University Press.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge: MIT Press.

Kemper, S., R. Herman, & C. Liu. 2004. Sentence Production by Young and Older Adults in Controlled Contexts. *Journals of Gerontology Series B: Psychological Sciences and Social Science* 59, 220-224.

Kiparsky, P. 1996. The Shift to Head Initial VP in Germanic. In H. Thrainsson (ed.), *Studies in Comparative Germanic Syntax* 2, 140-179. Dordrecht: Kluwer Academic Publishers.

Kiss, K. 1987. *Configurationality in Hungarian*. Budapest: Reidel & Akademiai Kiado.

Kroch, A. 1989. Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change* 1, 199-244.

Käi, P. 1997. *Eesti keele kõrvallause sõnajärg eri eagruppide kõnes*. An Unpublished MA Dissertation. Tallinn: Tallinn Pedagogical University.

LaPolla, R. 1995. Pragmatic Relations and Word Order in Chinese. In P. Downing & M. Noonan (eds.), *Word Order in Discourse* 299-331. Amsterdam & Philadelphia: John Benjamins Pub. Co.

_____. 2002. A Problem with Word Order Typology: Differing Uses of Word Order among Languages. *Research Center for Linguistic Typology* 1-14. Melbourne: Institute for Advanced Study, La Trobe University.

Lightfoot, D. 1991. *How to Set Parameters: Arguments from Language Change*. Cambridge, MA: MIT Press.

Lightfoot, D. & N. Hornstein. 1994. (ed.) *Verb Movement*. Cambridge: CUP.

Lindström, L. 2000. Narratiiv ja selle sõnajärg. *Keel ja Kirjandus* 3, 190-200.

Müller, G. 1999. Zur Ableitung der NP Adv V Stellung im Deutschen. In J. Bayer & C. Römer (eds.), *Von der Philologie zur Grammatiktheorie* 117-138. Tübingen: Niemeyer.

Normann, E. (ed.) 1955. *Valimik eesti vanasõnu*. Tallinn: Eesti Riiklik Kirjastus.

Odden, D. 2003. Languages and Universals. *Journal of Universal Language* 33-74.

Põldmäe, R. (ed.) 1941. *Eesti rahvanaljandid, I. Inimese eluperioodid*. Tartu: Teaduslik Kirjandus.

Roberts, I. 1997. *Comparative Syntax*. London: Arnold.

Remmel, N. 1963. Sõnajärjestus eesti Lauses. Deskriptiivne käsitlus. *Eesti keele süntaksi küsimusi* 216-271. Tallinn: Eesti Riiklik Kirjastus.

Sahkai, H. 1999. Eesti verbifraasi sõnajärg. *Keel ja kirjandus* 1, 24-32.

Sampson, G. 1997. Depth in English Grammar. *Journal of Linguistics* 33, 131-151.

Stuurman, F. 1985. *X-Bar and X-Plain. A Study of X-Bar Theories of the Phrase Structure Component.* Dordrecht: Foris.

Tael, K. 1988. *Sõnajärjemallid eesti keeles (võrrelduna soome keelega).* Tallinn: TA Keele ja Kirjanduse Instituut.

Tauli, V. 1983. The Estonian Language Reform. In I. Fodor & C. Hagége (eds.), *Language Reform. History and Future* 2, 309-330 Hamburg: Buske Verlag.

Tesnière, L. 1959. *Elements de Syntaxe Structurale.* Paris: Klincksieck.

Trosterud, T. 1996. Die Südsamische Wortfolge al Seine Kombination der Deutschen und der Marischen Wortfolge Analysiert. In *Lapponica et Uralica. 100 Jahre Finnisch-Ugrischer Unterricht an der Universität Uppsala. Studia Uralica Upsaliensia* 26, 103-112. Uppsala: Almqvist & Wiksell International.

Valgma, J & N. Remmel. 1968. *Eesti keele grammatika*. Tallinn: Valgus.

Vennemann, T. 1974. Theoretical Word Order Studies: Results and Problems. *Papiere zur Linguistik* 7, 5-25.

_____. 1981. Typology, Universals and Change of Language. Paper Presented at the International Conference on Historical Syntax,

Poznan, March 1981.

Vikner, S. 1994. Finite Verb Movement in Scandinavian Embedded Clauses. In D. Lightfoot & N. Hornstein (eds.), *Verb Movement* 117-147. Cambridge: CUP.

Vilkuna M. 1989. *Free Word Order in Finnish: Its Syntax and Discourse Functions*. Helsinki: Suomalaisen Kirjallisuuden Seura.

_____. 1998. Word Order in European Uralic. In A. Siewierska (ed.), *Constituent Order in the Languages of Europe* 173-233. Berlin & New York: Mouton de Gruyter.

Yngve, V. 1960. A Model and a Hypothesis for Language Structure. *Proceedings of the American Philosophical Society* 104, 444-466.