

Head directionality – in syntax and morphology¹

Hubert Haider, Univ. Salzburg

Dept. of Linguistics & Centre for Cognitive Neuroscience

Abstract

The particular position of the head in a phrase correlates with a wide range of syntactic properties. Mutually exclusive subsets of these properties characterize head-final ('OV'), head-initial ('VO') or 'freely positioning' ('third type') languages. The linearization of the head is a function of a category-specific directionality value. The interaction of the parametric directionality property of lexical heads with the cross-linguistically invariant directionality property of phrases (right-branching) accounts for major syntactic differences between OV-, VO-, and T3 languages. Analogous effects are located in properties of morphological structuring.

1. Introduction

This chapter is organized as follows. Section 2 briefly introduces two aspects of directionality, as a property of *heads* and as a property of *structures*. Section 3 presents a set of syntactic contrasts that distinguish head-initial and head-final phrases. The explanation for these contrasts is derived and formulated in terms of a directionality theory in section 4. Section 5 deals with syntactic properties of Slavic languages as examples of *unspecified* directionality. Section 6 briefly compares the directionality properties of phrasal structure with word structure. The appendix summarizes the major differences between two competing accounts.

Reference to the position of the head is a convenient criterion for grouping languages. There are strictly head-final languages like Japanese. This means that for any phrase, the head *follows* its dependents. On the other hand, in strictly head-initial languages (e.g. Austronesian languages), any head *precedes* its dependents. SVO is the most frequent subtype of head-initial languages. A language of this type is head initial apart from the fact that the position of the subject as an argument of the verb is apparently exceptional. It precedes the verb while the other arguments follow.

Cross-categorical uniformity is not obligatory, however. There are categorially differentiating languages. The head position may co-vary with the category of the head. In German and Dutch, verbal and adjectival heads follow; nouns and prepositions precede their complements. Finally, there are languages with seemingly variable head positions. In Slavic languages, for instance, a verb may not only precede or follow its arguments; it may also be 'sandwiched' by its objects. Here, scrambling is not the cause but rather the effect in a syntactic setting with characteristics of its own. Arguably, in these languages the directionality of the head relative to its complements is un(der)specified.

At least since Greenberg (1963), the search for cross-linguistic correlations has been guided by particular attention to the order of heads relative to other grammatical properties. Typically, inductive generalizations have been formulated by correlating cross-category patterns. Take for instance Greenberg's universal #5: "*If a language has dominant SOV order and the*

¹ This paper has benefitted greatly from advice and criticism by anonymous reviewers and the editors. My gratitude does not imply any responsibility on their side for whatever shortcomings may have remained.

genitive follows the governing noun, then the adjective follows the noun.” Had German and Dutch (Greenberg 1963:67, appendix II:10) been correctly identified as OV plus V2, ‘universal’ #5 could not have been formulated this way. In German, attributive adjectives obligatorily precede; the genitive follows the noun and the nominal arguments of the verb must precede the verb (1). Greenberg was misled by the V-2 property that typologists had not recognized before Mallinson & Blake (1981).²

- (1) dass man [den *ehemaligen* Präsidenten *des Nachbarlandes*] nicht *einreisen* ließ
that one [the *former* president (of) *the neighboring country*_{Gen.}] not *enter* let

A correlation like universal 5 is an inductive generalization. It describes a clustering of properties of a sample of languages. If correct, inductive generalizations call for an explanation. This means that they have to be paired with general principles of a grammar theory that cover the inductive findings.

In grammar theory, linearization properties have been captured in various ways. The position of an item in linear order may be characterized as a reflex of its original phrase-structural position, or as the result of a syntactic process that places an item in a derived position, or even as the outcome of autonomous (post-syntactic) rules of linearization.

What would this mean for OV versus VO, for concreteness? In the Principles & Parameter paradigm, based on Chomsky (1981), the position of the verb in the VP has been characterized by a parametric directionality requirement.³ The head determines whether a complement has to precede or follow it. This property is seen as a parametric property of the classes of lexical heads. For VPs with two objects, a parametric choice was assumed between the structural option (1a) and the option (1b). The latter turned out to be empirically inadequate.⁴

- (1) a. [XP [YP V°]_V]_{VP} ‘OV’-type
b. [V° XP]_V [YP]_{VP} ‘VO’-type (1st guess) (see Chomsky 1981:171)

In (1a,b), the linear position of the verb relative to its objects is determined by its directionality requirement. However, most of the syntactic predictions entailed by the left-branching structure (1b) turned out to be empirically inadequate, as Barss & Lasnik (1986) pointed out. Consequently, Larson (1988) suggested replacing (1b) by a derivation ending up at a right-branching shell-structure as an empirically more adequate structure; see sect. 4.

The current model of Generative Grammar evades directionality parameters especially in those variants that pursue Kayne’s (1984) views. The relation between OV and VO is captured derivationally instead. The base position of heads is assumed to be identical in head-initial and in head-final phrases. As for the VP, the allegedly universal structure is that of the

² “The order used for a stylistically unmarked version of ‘John saw Mary’ in German would be SVO, too, but to simply call German an SVO language would disguise the verb-second nature of its word order.” Mallinson & Blake (1981:129).

³ Parametric directionality has been assumed for case or theta-role assignment (see for instance Koopman (1984:120)).

⁴ Not only in typological schools is head-initial serialization confounded with the *structural* notion ‘left/right-branching’, as in Dryer (2009:185): “I argued that the word order correlations reflect a tendency for languages to be consistently left-branching or consistently right-branching.” He obviously means that in these languages phrases tend to be head-initial (‘left’-headed) or head-final (‘right’-headed). This is a statement on word order, not on structure. Positive evidence for the alleged left-branching structures is not delivered, of course.

English VP. What appears to vary across languages is believed to be merely the position of the dependents at “spell-out”. In OV, as for instance in Dutch (Zwart 1994), the originally post-verbal dependents are assumed to have moved in order to end up in higher, pre-verbal positions.

Once this conjecture is put to test seriously, with sufficient evidence from OV languages,⁵ straightforward central predictions turn out to be in conflict with empirical evidence; see for instance Haider (2010, ch.1) and Haider (2013, ch.9). For a summary of the major empirical shortcomings please consult the appendix of this chapter.

The head-final and the head-initial architecture of phrases produce distinct sets of syntactic properties. Representative examples can be found easily in languages that contain both types of phrases. In German, the VP is head-final while the NP is head-initial. In addition, German word formation provides a fully productive category conversion paradigm. The infinitive form of a verb may be used as a noun. Category conversion therefore yields the closest possible minimal pair context for V° and N° as phrase-final or phrase-initial heads, respectively:

- (2) a. [ein Verb_{Acc} in ein Nomen *konvertieren*_{V°}]_{VP}
a verb into a noun convert_{Inf.}
b. das [*Konvertieren*_{N°} eines Verbs_{Gen.} in ein Nomen]_{NP}
the [convert_{Inf.} a verb into a noun (,the converting of a verb into a noun‘)

In the NP, the *preverbal* accusative object of the verb corresponds to the *post-nominal* genitive object of the noun. These are the regular forms of the structural case of the object in a VP and the structural case of an object in an NP. In the VP, the object precedes but in the NP it follows the head. These are the structural aspects of the conversion. And here come some of the concomitant syntactic differences. The German NP as a head-initial phrase behaves syntactically just like an English head-initial phrase, that is, an NP or a VP with respect to *scrambling* (3), *compactness* (4) and *adjunctions* (5), to name just a few syntactic properties that sharply differentiate between head-initial and head-final phrases.

- (3) a. [ratlose Syntaktiker an ihren wilden Hypothesen erkennen]_{VP}
perplexed syntacticians at their wild hypotheses recognize
‘recognize perplexed syntacticians by their wild hypotheses’
b. [an ihren wilden Hypothesen ratlose Syntaktiker erkennen]_{VP}
c. das [Erkennen der/von ratlosen Syntaktiker/n an ihren wilden Hypothesen]_{NP}
the [recognize_{Inf.} the_{Gen.}/of perplexed syntacticians by their wild hypotheses
d.*das [Erkennen an ihren wilden Hypothesen der/von ratlosen Syntaktiker/n]_{NP}
e.*recognize by their wild hypotheses perplexed syntacticians

Note that the contrasts in (3) clearly show that scrambling is *phrase specific* rather than language specific. It is not a holistic property of a language. The constraint is nothing but a corollary of the compactness of head-initial phrases (4). For the verb and its object, a scrambled phrase is an intervener and therefore on a par with an intervening adverbial like in (4).

⁵ Note that the vast majority of syntacticians working in the Generative Grammar paradigm are native speakers of English or some other Indo-European SVO language and that “*every p-theory* [present day grammar theory_{H.H.}] *was developed by a native speaker of English.*” (Dixon 2011:182).

- (4) a. [ein Abstract für eine Tagung nach der Deadline einreichen]_{VP}
an abstract to a conference after the deadline submit⁶
b. das [Einreichen *eines/von Abstracts* für eine Tagung nach der Deadline]_{NP}
the submit_{Inf.} a_{Gen.}/of abstract(s) to a conference after the deadline
‘the submitting of an abstract / of abstracts to a conference after the deadline’
c. *das [Einreichen für eine Tagung *eines/von Abstracts* nach der Deadline]_{NP}
d. *das [Einreichen nach der Deadline *eines/von Abstracts* für eine Tagung]_{NP}
e. *submit for a conference *an abstract* after the deadline

The following set of data illustrates a restriction on adjuncts that occurs only with head-initial phrases as adjunction sites. The head of the adjoined phrase must be adjacent to the adjunction target. This excludes the adjoining of phrases with post-head material to head-initial phrases. In Haider (2000, 2004, 2013 ch.6) this property has been named ‘edge effect’.

- (5) a. that he had [much more carefully (*than others (thought)) [analyzed the problems]]_{VP}
b. dass er viel sorgfältiger (als andere (dachten)) die Probleme analysiert hatte
that he much more carefully (than others (thought)) the problems analysed had
c. eine [viel sorgfältigere (*als andere (dachten)) [Analyse der Probleme]]_{NP}
a much more careful (*than others (thought)) analysis (of) the problems

English (5a) contrasts with German (5b). As in German the VP is head-final, the edge effect is absent in (5b), but it shows up with the head-initial NP (5c). In English, both the VP and the NP are head-initial and therefore the edge effect applies to preverbal adverbial phrases as well as to prenominal phrasal attributes. As a consequence, complex attributes with post-head material are invariably post-nominal⁷ if a language allows attributes on either side of the NP, like in Romance languages. What these facts suggest is a robust syntactic impact of the position of the head within the phrase on the particular syntactic properties of the phrase.

Directionality is at issue whenever a structure becomes complex in phrase structure as well as in word structure. In morphology, “*perhaps the largest theoretical question is why suffixes are more frequent than prefixes.*” (Dryer 2013). What this question addresses is a directional asymmetry in word structure. Haspelmath (1992) among others suggests regarding suffixes as heads of complex words, on a par with the right-hand constituents of compounds. The head of a phrase determines the category of a phrase, and the suffix of a complex word determines the grammatical properties of the word, as for instance the category and grammatical gender.

Nominal compounding provides a good illustration of the point. In Romance, nominal compounds are head-initial⁸ but this word formation pattern is strongly constrained (6b). It cannot be applied recursively (Haider 2001). This contrasts with the productive recursive compounding in Germanic languages (6a).

⁶ All six permutation variants of the three phrases are grammatical.

⁷ In Romance, APs are head-initial, hence APs consisting of head plus complement are invariably postnominal. In English, the prenominal position (i) is ungrammatical for a complex AP, just like in Romance languages. So an appositive post-nominal variant is the last resort option (ii). In German (iii.), the complement precedes A°.

i. * a [familiar with Kayardil]_{AP} linguist
ii. a linguist, [familiar with Kayardil]_{AP}
iii. ein [mit Kayardil vertrauter]_{AP} Linguist

⁸ For instance Fr. *café filtre* (filter coffee); It. *conferenza stampa* (press conference); Sp. *hombre rana* (frog man)

- (6) a. [N N_h], [N [N N_h]], [N [N [N N_h]]], ... (head-final recursive compounding)
 b. [N_h N], *[N_h N] N], *[[N_h N] N] N], ... (head-initial recursive compounding)

Recursive head-initial compounding would produce left-branching word structures (6b). However, unlike the productive head-final type (6a), the head-initial type (6b) is not used recursively.⁹ Stacking to the left is productive¹⁰; stacking to the right is not. Stacking to the right would result in left-branching structures. If – as will be discussed in detail below – the structure of complex linguistic structures is right branching, recursive head-initial compounds are restricted to the trivial format of [X_N+Y]_N.

2. Directionality of lexical heads and directionality of structures

For lexical heads, directionality information is part of the lexically stored information. The directionality of structures, on the other hand, is a property regulated by the system that determines the properties of phrase structure. Directionality as a lexical property is easy to demonstrate. In German, for instance, relational particles are prepositions plus a handful of postpositions. Some of them alternate between preposition and postposition (7a), but some are grammatical only as postpositions (7b).

- (7) a. [der Vorschrift nach_{P°}]_{PP} – [nach_{P°} der Vorschrift]_{PP}
 the regulation_{Dat.} according-to according-to the regulation
 b. [Gerüchten_{Dat} zufolge_{P°}] – *[zufolge_{P°} Gerüchten]
 rumors according-to

It is clearly not a property of the noun ‚Gerüchte‘ that it has to precede the postposition ‚zufolge‘. It is a property of ‚zufolge‘ as a lexical head that it is a postposition. Phrased in directional terms, ‚zufolge‘ requires a complement that precedes rather than one that follows.¹¹

The correlation between lexical (classes of) items¹² and their directionality property is not limited to relational particles (pre- or post-positions), of course. It is a property of all major lexical categories as well as of functional lexical heads. For instance, there are languages with two sets of complementizers, namely head-initial and head-final ones. An example is Bengali (Singh 1980; Bayer 2001; 1998: 245). This language is representative of a group of Eastern and South-Eastern Aryan languages (Oriya, Assamese, Marathi, Dakkhini-Hindi). Bayer (1998) emphasizes that the final complementizers are base-generated in the clause-final position and argues in detail against hypotheses that surmise that these complementizers might happen to end up in the final position in the course of a derivation that moves the complement of a clause-initial C° to the left, thereby surfacing the complementizer in an apparently clause final position. Koopman (1984, sect. 3.4.2) discusses a similar property in the complementizer

⁹ Here is an example of recursive N+N compound formation: stone – [corner stone] – [granite [corner stone]] – [ersatz [granite [corner stone]]], ...

¹⁰ E.g. “Space-speak is an *engineering technology concept expression manuscript sentence grammar device*.” David McNeill. URL: <http://recremisi.blogspot.co.at/2010/08/linguistics-speaking-of-space.html> (18.3.2014).

¹¹ ‘Zufolge’, ‘halber’, ‘zuliebe’ are genuine postpositions. Whatever complement they select precedes. This must not be confused with the fact that the defective pronoun ‘es’ is avoided in the Germanic OV languages as a complement of prepositions. Instead, there is a pre-clitic: [mit *es*] → *damit* (Ge.); [met *het*] → *ermee* (Du.).

¹² Directionality may even be *item specific* as in the case of ‘enough’ and any of its cognates (Du. *genoeg*, Sw. *nog*, Da. *nok*) It is the ‘odd-ball’ item. Unlike all other measure modifiers, which precede, it obligatorily *follows* the phrase it modifies: “strange *enough*” – “sufficiently strange”, “seltsam *genug*” – “*genügend* seltsam”, etc.

system of Vata. Clauses with initial C follow the selecting verb; clauses with clause-final C precede the verb. Bayer (1998) reports the very same distribution for Bengali.

As for directionality in structures, it is a graph-theoretic truism that for each binary branching structure, there is a mirror image structure that differs only in directionality. It is no truism but empirically well-founded that the structures of linguistic expressions do not exploit the mirror image potential of phrase structuring. This is easy to see when we examine the periphery of a clause. There are well-known syntactic phenomena that target the clause initial position, that is, the left periphery. In numerous languages, question formation or relative clause formation places a specifically marked item into the leftmost position. On the other hand, there is no counterpart as a set of languages that target the right periphery for the same class of items, that is, the clause final position.

A particular illustrative case is the V2-property of the Germanic languages. A declarative clause starts with a single ‘arbitrary’ constituent followed by the finite verb. Not a single language is known with a mirror image configuration, that is, one that places the finite verb in the penultimate position, followed by a single, arbitrary constituent. There is no evidence that V-penultimate languages could exist. In terms of phrase structure properties, this invites the conclusion that (7b), the mirror image of (7a), is no option for natural language grammars.

- (7) a. $[_{FP} \dots [_{F^\circ} \dots [\dots]]]$
b. $*[[[\dots] \dots F^\circ] \dots]_{FP}$

In (7), F° is a variable for a functional head position. In general, such a position can be lexicalized either directly by a lexical functional item (e.g. a complementizer) or in a secondary manner, by displacing an independent lexical head with appropriate features into this position, as for instance the finite verb. In the latter case, directionality of the functional head cannot be related to a *lexical* property of a specific class of lexemes. In German, for instance, the verb as the head of a VP is head-final, but the head position in the V2-structure (7a) is head-initial, that is, it precedes its complement as functional head.

The example of V2 is representative of all instances of functional head positions for which there is no primary lexicalization. An analogous case in Romance languages is the functional head position between the subject and the VP that accommodates the finite verb. In English, this position is open only for finite auxiliaries or quasi-auxiliaries.

Since the directionality of F° in (7) is not correlated with any lexical property, it must follow from a general system-internal property of phrase structures of human languages. There is no parametric variation of the position of secondarily lexicalized functional heads relative to their complement position, as indicated in (8a,b), contrary to empirically insufficiently backed handbook wisdom. Available evidence strongly supports the conclusion that (8b) does not exist. There are no head-final functional phrases that are lexicalized secondarily.

- (8) a. $\dots [_{F'} V_i [_{VP} \dots e_i \dots]]$
b. $*\dots [[_{VP} \dots e_i \dots] V_i]_{F'}$

It is telling that (8b) has been tacitly assumed to exist in OV, in spite of the complete absence of any positive evidence of the kind that motivated (8a). The assumption of ‘V-to-I’ in English or French was corroborated by the evidence from intervening items, as indicated in (9a).

Corresponding patterns are predicted for OV but they are absent. In head-final clauses, the sequence of clause-final verbs is compact. The predicted order (9b) is ungrammatical, without exception.

- (9) a. ... [_{F'} V_i [(XP) [_{VP} e_i ...]]]
b. ... [[[_{VP} ... e_i] (*XP)] V_i]_{F'}

In (9a), a typical instance of XP is an adjunct to the VP as for example a manner adverbial. A case for the XP in (9b) is material that is extraposed to the right edge of a VP. Although extraposition targets the right edge of a VP (10a), this pattern is clearly ungrammatical in the context of (9b). The predicted order would be the unacceptable order (10b). The only acceptable order is (10c).

- (10)a. [Warten können *auf ein Wunder*]_{VP} wird er nicht
wait can_{Inf.} for a miracle shall he not
b. *dass er nicht [warten können e_i *auf ein Wunder*] wird_i
that he not wait for a miracle can_{Inf.} will
c. dass er nicht warten können wird *auf ein Wunder*

Extraposition targets the right edge of the containing VP. This is easy to check when a VP is topicalized (10a). If the finite verb were to move out of the VP, up to a clause-final functional head position, the resulting order would be (10b). But this is an unacceptable order. For additional independent evidence¹³ see Haider (2010:59-68).

The source of the ungrammaticality of (7b) and (8b) is identical. The functional head position that is in need of secondary lexicalization must not *follow* the phrase whose head is supposed to be displaced into the functional head position. It must *precede*. This asymmetry is covered by a universal restriction against left-branching (extended) projections (see section 4, (21i)).¹⁴

3. Some correlates of the linear position of the head in relation to its dependents

This section lists and describes a set of syntactic properties that correlate with the head-initial vs. head-final positioning in Germanic languages. By the same token, the list is a challenge for grammar theories and their success in accounting for these phenomena in an insightful way. Directionality-based explanations for the listed properties will be derived in section 4.

Compactness: As illustrated already in the introduction section with examples (3d,e), head-initial phrases are *compact*. Intervenors between the head and the nominal complements are ungrammatical. This restriction rules out intervening *adverbials* just as well as *scrambled* phrases since they would qualify as intervenors as well. Head-final phrases are not compact.

Edge effect: The head of a phrase-initial adjunct of a head-initial phrase must be adjacent to

¹³ For German and Dutch, a post-VP functional head as a target of V-movement has been assumed by mere analogy to the pre-VP functional heads in English. Clear positive evidence has been missing from the beginning.

¹⁴ Note that this restriction covers the otherwise unexplained so-called Final-over-Final (FoF) constraint (Haider 1992, sect.5.2; 2013: 132-35). FoF says "For all heads { α , β , ...} on a single projection line, if α is a head-initial phrase and β is a phrase immediately dominating α , then β must be head-initial. If α is a head-final phrase, and β is a phrase immediately dominating α then β can be head-initial or head-final." Biberauer & Holmberg & Roberts (2007) readdress an issue that has originally been opened by Kiparsky (1996; his ex. 20c): There are no VO languages with a post-VP 'I'-position (8b). The BBC eliminates the unattested case: There are no left-sisters of functional heads targeted by head-movement (i.e. functionally extended VPs), neither in OV nor in VO.

the host phrase. Any material that may follow the head is ungrammatical when the phrase is in an adjunction position. This property can be studied with adnominal attributes or with adverbial phrases, as in example (5) in section 1. There is no edge effect for adjunct adjoined to a head-final phrase,¹⁵ and there is not edge effect for phrases in spec-positions either.

Particles and result predicates: Without any exception, these elements *precede* the head in OV and *follow* the head in VO (11a,b). If the given language allows particle stranding, a particle of a particle verb may be stranded in between two nominal arguments in VO (11c). In OV, the only stranding position is the clause-final position, that is, the base position of the verb in the VP. The parallel behavior of particles and result predicates is expected, since V+particle combinations typically denote a result relation (11).

- (11)a. You must cut it *loose/off*
b. Du must es *los/ab*-schneiden
you must it loose/off-cut
c. Let's send her *up* a drink

Higher-type adverbials and negation particles: In VO-languages, sentence adverbials precede the VP (12a). VP-internal placement (12b) is deviant. In OV languages, sentence adverbials arguably are VP-internal. They may be placed in between VP-internal elements and the VP-final verbal head. (12c) merely illustrates the claim. Sufficiently detailed argumentation tends to be lengthy (see Haider 2010:170f.).¹⁶ These patterns will be shown to follow from the directionality system to be explicated in section 4.

- (12)a. He has *obviously/unfortunately* returned it to the wrong place
b.*He has returned it *obviously/unfortunately* to the wrong place
c. Manchmal können Leute damit ja *leider/offensichtlich* nicht umgehen
sometimes can people with-it PRT unfortunately/obviously not deal
'sometimes, people unfortunately/obviously cannot handle it / deal with it'

The placement of negation particles follows analogous patterns. In VO, the negation precedes the VP. In OV it is VP-internal. In each case, the negation particle closely *precedes* the *head* of the VP.¹⁷ In both cases, negation minimally c-commands its semantic type domain (Haider 2013:138). This, but not the VO-pattern, is the cross-linguistically valid constellation.

In-situ wh-items of a higher type: It is uncontroversial that in English *why* or *how* are deviant if they are not fronted (13). *Where* and *when*, however, are perfect. Szabolcsi and Zwarts (1993) realized that this reflects a semantic difference. *Why* and *how* do not range over individual-type variables, that is, type <e> elements, but over higher types, like properties or propositions. Just like argumental wh-items, *when* and *where* operate over individual type variables, namely points of time and place, respectively.

¹⁵ Note that the edge effect is unexplained if an adverbial phrase is assumed to be accommodated by the specification of a covert adverbial functional head. Phrases in Spec do not display an edge effect.

¹⁶ A check by means of a simple corpus search count is instructive. A web search for "*sent it unfortunately to*" produced 2 hits while "*unfortunately sent it to*" produced 4 600. There was no hit for "*returned it obviously to*" but 1560 for "*obviously returned it to*" (Google search, March 21st 2014).

¹⁷ More precisely, negation particles c-command the item that is their semantic target. For constituent negation it is the negated phrase; for sentence negation, it is the item that situates the situation variable, that is, the (base position of the finite) verb.

- (13)a. (Tell me) who called when/*why?
b. (Tell me) what happened where/in what manner/*how?

In OV languages, this contrast is absent. Any type of adverbial operator may be left in situ (cf. Haider 2010:118-22). The causal difference is the relative order of adverbs and the verb. In OV, the in-situ wh-elements precede and c-command the (base position of the) finite verb, in VO they do not c-command the (base position of) the finite verb when they follow but they would have to, in order to receive their type-adequate semantic interpretation.

Obligatory functional subject position: In Generative Grammar, this property is covered by the EPP. According to Lasnik (2001:356) “*The ‘Extended Projection Principle’ (EPP) has been [...] a pervasive mystery since it was first formulated by Chomsky (1981).*” It demands that every clause have a subject in a functional spec position above the VP. The empirical basis for Chomsky’s generalization was English, and confirming evidence came from European SVO languages. A substantive part of the mystery is that OV languages do not recognize the EPP. The clearest symptom is the necessary presence or absence of an expletive subject.

In non-pro-drop SVO languages, a missing subject argument is compensated by means of an expletive. In SOV, subject expletives are ungrammatical in truly subjectless clauses.¹⁸ The passive of an intransitive verb produces instances of subjectless clauses. In Scandinavian languages,¹⁹ an expletive ‘saves’ the EPP (14a). In German, the expletive in the allegedly universal Spec position for the subject is ungrammatical (14b), but German is not ‘semi-pro-drop’, as the obligatory expletive subject in the intransitive middle construction shows (14c). ‘Es’ would be the appropriate expletive since it is also the obligatory expletive for the clause-initial spec position in a V2-clause (14d).

- | | |
|--|-----------|
| (14)a. Ofte vart *(det) telefonert
often was EXPL telephoned | Norwegian |
| b. Oft wurde (*es) telephoniert
often was EXPL telephoned | German |
| c. Hier telephoniert *(es) sich leicht
here telephones EXPL itself easily | |
| d. Es wurde oft telephoniert
EXPL ar often telephoned | |

What has been overlooked in the various endeavors of shielding EPP against problems raised by German or Dutch is the fact that *any* SOV language is a problem for EPP and that the EPP itself is merely the name of a problem but not its solution. An EPP feature, strong or weak, does not explain; it merely replaces an empirical generalization by a technical notation in terms of an unexplained feature. Why should OV systematically differ from VO with respect to an EPP feature? There must be a more profound explanation than an ad-hoc feature, especially in the absence of a theory of syntactic features. The auxiliary assumption that every OV

¹⁸ Quasi-arguments, like weather-verb subjects, are not expletive. They are semantically empty arguments and this information is part of the lexical entry of the respective verbs.

¹⁹ In English, the obvious candidate for a subject expletive, namely *there*, became unavailable once it had been tied to a postverbal subject argument in the ‘there-construction’. Grammar change with respect to available expletives has turned English into the only Germanic language without a passive of intransitive verbs.

language is a semi-pro-drop language obviously misses a generalization. Being semi-pro-drop is a contingent property, but the absence of an expletive subject in OV is a universal property.

Additional, independent evidence for the structurally different status of subjects in SVO and SOV comes from *wh*-subjects in situ. In English they are deviant (Hankamer 1974) and in other SVO languages too; in SOV languages, they are straightforward. Chomsky (1981:236) acknowledges that clauses like (15a) are ungrammatical because they contain an in-situ *wh*-subject. The German version is fully acceptable.²⁰

- (15)a. *It is not important who thinks (that) *who* won
 b. Es ist nicht wichtig, wer glaubt, dass *wer* gewonnen hat²¹
 it is not important who thinks that *who* won has

The restriction that accounts for the ungrammaticality of (15a) will simultaneously rule out (16a). In each case, there is a *wh*-subject in situ, but for Chomsky, (16a) is a standard case of the Superiority constraint. Strangely enough, the pattern (15a) has never again been re-examined in the discussions of Superiority violations. As the contrast (16a,b) illustrates, Superiority is challenged by data from SOV.

- (16)a. *It is entirely unclear who(m) *what* has irritated
 b. Es ist völlig unklar, wen *was* irritiert hat German
 it is entirely unclear whom *what* irritated has

In sum, the presence or absence of subject expletives as well as the acceptability of in-situ *wh*-subjects in combination with so-called Superiority violations that involve an in-situ *wh*-subject are areas of syntactic contrasts between SVO and SOV that call for a more principled explanation than the presently favored ones. In the MP perspective, the structures would be principally alike, with a *wh*-subject sitting in a pre-verbal spec position.

Order variation among auxiliaries and quasi-auxiliaries in SOV: In VO languages, the auxiliaries precede the main verb in an invariant relative order; in OV they follow.²² It is remarkable that typologists would not immediately recognize the connection with the directionality of the main verb, as Dryer (2009:204) frankly admits. “*The primary conclusion is that there is no obvious explanation for why auxiliary verbs tend to precede the main verb in VO languages but follow in OV languages.*” Here is an explanation. Bech (1955) coined the term ‘status government’ and saw the parallel to case government. The status of the dependent verb (e.g. participial or infinitival or aspectual form) is determined by the governing (auxiliary) verb. Just like case government involves directionality, status government does so, too. Hence, O+V languages are V+Aux languages, while V+O languages are Aux+V(P) languages. The main verb is a case governor; Aux is a status governor for the main verb. The

²⁰ In fact, the counterparts of each of the ungrammatical English sentences of Chomsky (1981:236) are grammatical in German. Chomsky (1981:237) attributes the deviance to the ‘*that-t*’ effect even though “*the presence of the complementizer is immaterial*”. See example (15a) and remember that the *that-t* effect is inoperative when ‘*that*’ is absent: *What_i do you think (*that) t_i explains this contradiction?*

²¹ It is easy to locate this pattern in corpora, even if it likely to be infrequent:

i. *Wer sagt, dass wer unaufmerksam ist?* <http://www.scilogs.de/hochbegabung/hochbegabt-und-aufmerksamkeitsgest-rt/>
 ii. *Wer meint, dass wer meint, er sei Hellseher?* <http://de.board.bigpoint.com/seafight/archive/index.php/t-173206-p-20.html>

²² Greenberg restricts his statement (Universal 16) to VSO vs. SOV because of misclassifications in his sample. German and Dutch are misclassified as SVO and their auxiliary verbs pattern in accordance with OV, obviously.

directionality of ‘government’ (viz. for the dependency relation between a verbal head and a depending element) is identical.

This was the easy part of verb ordering, and here is the tough issue: In VO-languages the relative order of auxiliaries and the main verb is invariant. In VO languages, the relative order is variable if the given language allows ‘V-movement’. Germanic OV-languages meet the latter requirement. The order variation in Dutch and German is illustrated in (17) and (18), respectively.

- | | |
|--|------------------------------|
| <p>(17) a. *dat hij niets gezien hebben <i>kan</i>
 that he nothing seen have can
 b. dat hij niets gezien <i>kan</i> hebben
 c. dat hij niets <i>kan</i> gezien hebben
 d. dat hij niets <i>kan hebben</i> gezien</p> | <p>Dutch (ANS 1984:1069)</p> |
| <p>(18) a. dass er nichts gesehen haben <i>wird</i>
 that he nothing seen have would
 b. dass er nichts gesehen <i>wird</i> haben
 c. dass er nichts <i>wird</i> gesehen haben
 d. *dass er nichts <i>wird haben</i> gesehen</p> | <p>German</p> |

As SOV languages with ‘immobile’ verbs (e.g. Korean, Japanese) show, (17a) and (18a) are the base orders, with the dependent verb preceding. A precondition for the variations is the ‘mobility’ of the verbs. In SOV languages without V-movement (e.g. Asian OV languages like Korean or Japanese), there is a single order, namely the order corresponding to (17a) and (18a). In standard Dutch, (17a) is not available any more. In standard German, on the other hand, the fully inverted Dutch order (18d) is not admitted. There are regional German variants, however, that permit this order (see Schallert 2014, sect. 5.3.1).

That-t-effect: In English and other SVO languages, fronting the subject across a complementizer is ungrammatical (19a). Swedish (Engdahl 1986:166) and Danish (Vikner 1995:12) pattern like English. In Dutch (Den Dikken et als. 2007:346) or German (19b), the counterpart of English (19a), is grammatical and acceptable.

- (19) a. What_i do you think (*that) *t_i* has happened in that night?
 b. Was glaubst du denn dass *t_i* passiert ist in jener Nacht?²³
 what think you PRT that happened has in that night

The *that-t* effect singles out subjects in SVO languages. In these languages, the subject position is preverbal whereas the objects are postverbal. This is a difference in the directionality relation between the verb and its arguments. In SOV, subject and objects are preverbal. There is no difference in directionality, hence no functional subject position and therefore no *that-t* effect. This is an indication that the appropriate understanding of the *that-t* effect as an SVO-specific restriction involves directionality.

²³ Den Dikken et als. (2007:345) report that acceptance is highest for Dutch informants when there is material in between the trace and the verbs following it. For me and informants I consulted there is no difference between (19b) and the version with the extraposed adverbial PP preceding the verb ‘passiert’. Both orders seem to be equally acceptable for speakers of standard German in Southern Germany and Austria. Northerners are known to shy away from extraction of any kind of arguments across a complementizer.

4. Universal and parametric directionality²⁴

Structural directionality is universal; head-directionality is parameterized. Its values may vary across lexical (sub)classes. Structural directionality is the directional relation between a projecting node on the projection line and its sister node. In principle, phrase structures could be right-branching (20a) or left-branching (20b). Evidently, the structures (20a,b) are mirror-images. All available empirical evidence points to the conclusion that a left-branching projection structure (20b) is not employed by human language grammars. A structure like (20b) is excluded on principled grounds. This is the *universal* directionality property of structures (21i).²⁵

- (20) a. [_{HP} X [_{H'} Y [_{H'} Z h°]]] right-branching
 b. [[[h° Z]_{H'} Y]_{H'} X]_{HP} left-branching

The *parameterized* property is a property of lexical heads. Complements either precede or follow the head of the phrase. As this formulation tells, this kind of directionality is a property of a relation between the lexical head of a phrase and its complements. Let us call this relation the licensing relation. The directional licensing relation in combination with the universal directionality of structuring is the grammatical source of the syntactic differences between the head-initial and head-final organization presented in the preceding section. The theoretical core assumptions are the following (Haider 1992; 2010: 26; 2013: 3f.):

- (21) i. Projection lines are universally *right-branching*²⁶ and endocentric.
 ii. A dependent phrase is licensed in the *canonical direction*.
 iii. The position of a dependent phrase P is *licensed* =_{Def.} the (projection of the) phrase head h and P minimally and mutually c-command each other.

It is the *minimal & mutual* c-command condition (21 iii) that is directly causal for many of the OV/VO contrasts described in the preceding section, namely compactness, scrambling, the edge effect and the need of a functional subject position in VO, with the concomitant syntactic properties of the functional subject position. Let me emphasize that it is the *very same* principle (viz. merger is universally to the left) applied under parametric directionality that produces the different outcomes for OV and VO.

Let us begin with compactness. In keeping with (21i.), the structural grid of a phrase is that of (22a) but not (22b). (22a) offers two alternative foot positions for a head, namely x or y. The actual choice depends on the directionality value of the head.

- (22) a. ... [... [... [x y]]]
 b. *[[[x y] ...] ...] ...

The value for the canonical licensing direction is parametrical, that is, it is either progressive (“→”) or regressive (“←”). The two implementations in (23) illustrate the directionality difference in the sub-tree that contains the head.

- (23) a. ... [... [... [ZP ← V°]]] OV

²⁴ What follows is a condensed explication of the theory of parameterized directionality developed in Haider (1992; 2010 chapter 1; 2013, chapters 3 and 5).

²⁵ This claim has been published for the first time in Haider (1992): BBC = *Basic branching conjecture*.

²⁶ In other words, the direction of merger *within* a phrase is universally to the left.

b. ... [... [... [V^o → ZP]]]

VO

The crucial differences between OV (23a) and VO (23b) become ‘visible’ when the phrase becomes more complex. In OV, the canonical direction of licensing is *congruent* with the universal direction of merger (21i.). In VO, however, the canonical licensing direction is *opposite* to the universal direction of merger. This is the source of VP-shell formation which in turn is the source of compactness.

When a second object is merged in VO (24a), its position is not in the directionality domain of the head. Hence the head needs to be re-instantiated. This amounts to the formation of what is called a *VP shell* in Generative diction, namely (24b).

(24) a. ... [YP [V' V^o → ZP]]

b. ... [V_i → [YP [V' e_i → ZP]]]

c. ... [V_i → [(^{*}π) [YP [(^{*}π) [V' e_i → ZP]]]]]

A shell is necessarily compact because of the minimality requirement of the licensing condition. Any intervening phrase would destroy the relation of *minimal* c-command between the verb and YP, or between YP and the trace of the verb. In (24c), π would either disrupt minimal c-command between V and YP or between YP and the lower empty verb position. In each case mutual, minimal c-command is destroyed. Note that YP must minimally c-command the lower V-position. Mutual c-command is fulfilled if V minimally c-commands YP, and YP minimally c-commands a link of the chain of V. In (24c), the lower π would disrupt the minimal c-command of the lower, empty V position by YP. No mutual c-command between V and YP implies no licensing of YP by V.

In OV, the situation is different because the canonical directionality of licensing is congruent with the directionality of merger (25). Hence not only the head but any other node on the projection line can serve as a licensing node. As a consequence, there are no shells in OV. In sum, since minimal & mutual c-command is the core of the licensing relation; interveners are excluded in complex head-initial phrases but not in head-final ones. In (25a), V' as a projection of the head is a licit licenser for YP, but not in (24b) for the simple reason that the canonical directionality domain of V' includes YP in (25a), but not in (24b). Consequently, the OV structure tolerates interveners (26a), but the VO structure does not (26b,c).

(25) a. ... [... [YP ← [V' ZP ← V^o]]]

b. ... [... [YP ← [V' π ← [V' ZP ← [V' π V^o]]]]]

An adverb, for instance, as an adjunct to V' would not interfere with minimal & mutual c-command, since there is always a sister node for the next higher argument that is a projection of the head with the required directionality. This is the essential difference between head-initial ('VO') and head-final structures ('OV') with respect to compactness.

In OV, the projection nodes are licensing nodes, in VO they are not, because of the directionality mismatch. In VO, the only element that is able to provide the directionality requirement is the verbal head, and therefore it must be re-instantiated, whence the shell structure (24b) of complex head-initial phrases. As a consequence, head-initial structures are compact and do not leave room for scrambling.

(26) a. He would show (**unhesitatingly*) someone (**voluntarily*) his collection

- b. Er würde jemandem *ohne Zögern* seine Sammlung *freiwillig* zeigen
 he would someone without hesitation his collection voluntarily show

In (26b), the adverb '*unhesitatingly*' would destroy the minimal c-command relation between V and the object *everyone*. The adverb '*voluntarily*' disrupts minimal c-command between the object '*everyone*' and the lower, empty position of the verb. In each case, *minimal & mutual* c-command is violated. In (26b), on the other hand, there is always a sister node of the projection line with the required licensing directionality. Minimal & mutual c-command is a trivial property of sister constituents. The very same intervener status blocks scrambling. If an argument is scrambled, this means it is adjoined higher up. This turns the scrambled item into an intervener element. Hence, compactness and the ban against scrambling in VO are just two sides of the same medal.

Let us turn to the edge-effect. It singles out adjuncts that precede a head-initial phrase. What makes them special is the fact that these adjunct positions are not in the canonical directionality domain of the head-initial phrase. In head-final phrases, any preceding adjunct is in the directionality domain. Provided that the edge effect is an empirically correct generalization, it is a handy diagnostics for head-initial versus head-final structures, even if its grammatical source has not been fully understood yet²⁷ (see Haider (2013:144-45) for a proposal).

Let us proceed now to the hallmark of SVO languages, namely the *functional subject position*. The trigger of the 'EPP' property of SVO structures is a directionality mismatch. In SVO, the highest argument in the VP is not in the directionality domain of the verbal head (27a). The canonical directionality is to the right; the directionality of merger in phrases is to the left. Neither the verb nor a projection node can provide directional licensing for the VP-internal subject in (27a). Therefore, a functional head has to provide directional licensing (27b). The functional projection establishes the spec position that is typical for SVO languages (27b).

- (27) a. [VP XP_{Subj}. [V_i → [YP [V' e_i → ZP]]]]
 b. [FP XP_j [F' F° → [VP e_j [V_i → [YP [V' e_i → ZP]]]]]]

In OV, any argument of a verb is within the directionality domain of the verbal head, whence the absence of this functional projection in the OV-based clause structures. The functional spec position in (27b) is at the same time the position of *expletive subjects*. The absence of this functional layer in OV is the grammatical reason for the absence of expletives for this kind of subject position in OV. The EPP is a property of SVO languages. The functional spec position of the functional head that licenses the VP-internal subject is obligatorily lexicalized in SVO (except for pro-drop). The *that-t effect* and the restriction against *wh-subjects in situ* are immediately related to this functional subject position. Therefore it is not surprising that SVO languages differ from SOV languages precisely in these respects.

The explanation for *the ban against a wh-subject in situ in SVO* proposed in Haider (2010: 117) is based on this structural difference. In VO, an in-situ wh-subject is in a functional spec position. In OV, the in-situ wh-subject is in its VP-internal argument position. A wh-pronoun in a functional spec-position gains operator status. It cannot be interpreted as a dependent wh-

²⁷ Note that under the currently wide-spread assumption that adverbials are accommodated by specs of empty functional heads, there is no way of capturing the edge-effect. Phrases in Spec-positions would not be affected.

element anymore. This rules out patterns like (15a), discussed by Chomsky (1981:236). However, the acceptability of utterances with an in-situ wh-subject improves once the wh-subject is ‘endowed’ with operator properties, as Chomsky (1981:238) admits and Hornstein (1995:144) observes.

- (28) a. He wonders what *who* bought where?
b. What did *who*^{i/*j} reveal about *his*^{i/j} mother?

In Chomsky’s example (28a), ‘who’ is the binding operator for ‘where’, just like in Hornstein’s example (28b), ‘who’ is the binding quantifier for ‘his’. Once the wh-subject finds an element that may serve as provider for a variable to be bound, acceptability improves.

Let us return now to the rest of the OV-VO contrasts listed above. The distribution of *particles* and *result predicates* is a straightforward consequence of the licensing directionality of the verbal head. As predicates that are dependent on the verbal head, they are bound to occur in the canonical directionality domain. Furthermore, these elements are ‘inert’, that is, they do not scramble. If they end up distant from their verb, they must have been stranded. In all Germanic V2-languages, fronting of the finite verb strands the particle, but only in the stranding Germanic SVO languages, a stranded particle may appear in between two nominal objects. In the Germanic SOV languages, the particle immediately precedes the phrase final V position.

- (30) a. He sent them *up* a drink
b. The secretary sent the stockholders *out* a notice

The position of the particle in (30) is the lower, empty verb position in the VP shell. Thus, particle distribution is an immediate source of direct evidence for the shell structure of a head-initial VP, and the absence of a shell structure within head-final VPs.

Higher adverbials and *negation particles* semantically relate to the propositional content of the clause. Tense and mood situate the proposition. Tense and mood are coded functionally, represented formally on the finite verb. Hence, higher adverbials and sentence negation must c-command (a chain link of) the finite verb. In VO, this requirement can be met only in pre-VP positions. In OV, however, any position on the projection line preceding the VP-final head position c-commands the verb. Hence, VP-internal positions are suitable for sentence adverbials or sentence negation in OV but not in VO. Elevating the pattern of VO to the rank of a universal model would be inappropriate. It is not a universal property of negation or sentence adverbials that they have to precede the VP. This is merely a consequence for VO-languages.

Verb clustering is related to the right-branching constraint (21i.), too, but only indirectly. Stacking VPs in OV is bound to produce left-branching structures (31a). Cluster-formation avoids the recursive stacking of VPs in favor of recursively clustering the verbs within a single VP (31b). In VO, stacked VPs are right-branching (31c). Grammars are parser-friendly (as a result of cognitive evolution; see Haider 2013 ch.2 and Haider in press) and provide clustering (31b) as means of avoiding central embedding structures (31a)

- (31) a. *... [VP [VP [VP XP V₁] V₂] V₃]
b. ... [VP XP [[V₁ V₂]_{V°} V₃]_{V°-cluster}]
c. ... [VP V₁ [VP V₂ [VP V₃ XP]]]

A verb cluster is a complex structure that consists of head-to-head adjoined verbs. Since adjunction preserves the category, a head-to-head adjunction cluster is of the category V^o .²⁸ The cluster is a compact unit since head-to-head adjunction leaves no room for phrasal interveners.

5. Unspecified directionality? – Slavic as an exemplary case

Head-initial vs. head-final does not constitute an exhaustive classification. There are reasons for assuming a third setting, namely one with ‘variable’ head positioning. Previously, this option has been dismissed on a priori grounds, presumably in remembrance of the peripherality dogma of x-bar theory which postulates that lexical heads must be phrase-peripheral; see Kornai & Pullum (1990:35).²⁹

A representative example is the Slavic family. Slavic languages are customarily filed as SVO, mainly because subject-verb-object is a frequent word order in these languages.³⁰ However, an SVO ascription is questionable because typical syntactic characteristics of SVO languages are missing or contradicted. In fact, Slavic languages share more syntactic properties with OV than with VO languages as for instance scrambling, absence of the edge effect, order variation of auxiliaries, no ‘Superiority’ effect for in-situ wh-subjects, no ‘freezing’ of preverbal phrases (see Haider & Szucsich (in press) and the literature cited there).

Let me just exemplify the last property since it has not been introduced in the preceding section. In the German example (32a), the object of the scrambled, preverbal infinitival clause is fronted to the initial position of the main clause. In the Russian examples an attribute of a preverbal NP is questioned or topicalized and fronted.³¹ In an SVO language like English, extractions from these positions are ungrammatical.

- | | |
|--|------------------------------|
| <p>(32)a. Wen_i/Ihn_i hat [damit e_i zu konfrontieren]_j keiner e_j gewagt?
 whom/him has [with-it to confront] nobody dared
 ,Whom has nobody dared to confront with it?’</p> <p>b. Kakuju_i Ivan [-_i mašinu] <i>kupil</i> svoej žene?
 which_i Ivan [-_i car] <i>bought</i> his wife_{Dat}
 ‘Which car did Ivan buy for his wife?’</p> <p>b. Japonskuju_i Ivan [-_i mašinu] <i>kupil</i> svoej žene.
 Japanese_i Ivan [-_i car] <i>bought</i> his wife_{Dat}
 ‘A Japanese car, Ivan bought for his wife.’</p> | <p>German</p> <p>Russian</p> |
|--|------------------------------|

In German, a scrambled phrase remains within the directionality domain of the verb. In English, a preverbal phrase is outside the directionality domain. And what about the preverbal NP in (32b,c) in Russian? It behaves as if it was within the directionality domain, and it arguably

²⁸ Note that clusters are not *compounds*. Head-to-head adjunction structures are syntactic structures that are open for syntactic processes like movement within the cluster (see German or Dutch IPP constructions). A compound would be the result of a word formation process. Its internal structure is not accessible for syntactic processes like movement. This is contrary to Bierwisch’s (1990) attempt to derive clustering morphologically.

²⁹ Variable head-positioning has been considered by Hoeksema (1992:121), and by Santorini (1993) for Yiddish.

³⁰ “It is generally acknowledged that Russian is an SVO language in neutral contexts” John F. Bailyn (2002).

³¹ Note that this presupposes that the N-head is able to license to the left, which is predicted if the licensing directionality is ‘free’.

is indeed. Could a Russian verb be an ‘ambidextrously’ licensing head? In fact, this is what ‘underspecified directionality’ amounts to.

(21ii) refers to canonical directionality, which is parameterized. The values for the parameter are regressive or progressive, that is, to the left or to the right. This is not the whole picture, however. There is an implicit third option, namely the option that the value is not specified. In this case the ‘canonical’ directionality is free. Let us call this setting the ‘third type’ (T3). Languages of the third type may appear to be of a mixed type, with OV and VO properties. This impression is not wrong, but these languages are not ‘mixed’. Their grammars merely generate the properties that are consistent with the unspecified directionality of licensing. Here is an example. The patterns in (33) are genuine T3 patterns. At each point of merger, licensing may be implemented with either of the two directionality values.

- | | |
|---|-----------------|
| (33) a. $[X \leftarrow [Y \leftarrow [Z \leftarrow V^{\circ}]]]$ | OV-like pattern |
| b. $[X \leftarrow [Y \leftarrow [V^{\circ} \rightarrow Z]]]$ | T3-only pattern |
| c. $[X \leftarrow [V_i^{\circ} \rightarrow [Y [e_i \rightarrow Z]]]]$ | VO-like pattern |

Pattern (33c) is a frequent one, arguably for information-structure reasons, and provides the basis for misclassifying T3 languages as exceptional VO languages. They appear exceptional because of the patterns (33a,b) that are absent in unequivocal SVO languages. These patterns are customarily explained away as scrambling effects. Y and Z (in 33a) or Y in (33b) are said to have been scrambled across the head-verb. Independent evidence for this claim is wanting.

The T3 hypothesis correctly predicts the ‘deviations’ from the typical SVO properties. There is no edge effect, because pre-VP adjuncts are in the directionality domain in each case in (33). The subject is in the directionality domain. So there are no effects triggered by the functional subject position of the SVO languages, viz. no effect for in-situ wh-subjects (e.g. Toman 1981, Bošković 1997, Meyer 2004), no obligatory subject expletives in EPP contexts, and no freezing effects for extractions out of preverbal constituents (Stepanov 2007). The order variation of auxiliaries (Embick & Izvorski 1997) directly reflects the free choice of the directionality value.

6. Directionality in morphology

Directionality is an issue of word structure, too. Simple inspection of the cross-linguistic order patterns of morphemes tells that suffixing is more frequent than prefixing, both in inflection and derivation, and that right-headed compounding is more frequent than compounding with a left-headed base. This is a first indication that word structures show a strong preponderance of right-branching structures, just like syntactic structures. Williams (1981), who notes an asymmetry of head-placement in morphology, proposed a *Right-Hand Head Rule*.

The word structure of a sequence of morphemes that constitute a word consists of the base (‘head’) and the (structured) elements that are binarily merged with the base. The proper identification of the word structure crucially depends on the correct identification of the base. In (34a), the base is the verb, and it gets prefixed. In (34b) the base is the nominalizing suffix. The category change from verb to noun is a function of the category of the suffix. In (34c) it seems as if the prefix is category-changing which is a property of a base element. Does this

mean that the prefix ‘*en-*’ is the base and that the word structure of these English words is left-headed?

- (34) a. *decompose*, *reiterate* ...
b. *composer*, *iteration* ...
c. *enrich*, *encrust*, *enlighten*...
d. *ergrünen*, *verkrusten*, *belichten*... German
en+green+INF, en+crust+ INF, en+light+ INF

If the structure of a word like *encrust* consisted merely of ‘*en-*’ plus noun, the prefix would have to be the base that combines with a noun and yields a verb. But there is another, equally plausible analysis. Category conversion (viz. word class change) already applies to the noun (cf. *a brush*, *to brush*; *a maple-leave*, *to maple-leave down*) and yields a verb, which gets prefixed. Since English is the only Germanic language that lacks an infinitive suffix, conversion is less easy to identify than in other Germanic languages. The German examples (34d) clearly show an infinitival ending that apparently combines with an adjective or a noun. However, the prefix is a preverbal prefix that is combined with a verb (cf. *er-arbeiten_V*; *be-arbeiten_V*). The verb is generated by category conversion. In this case, the infinitival suffix is category changing inflection (see Haspelmath (1996) for a cross-linguistic investigation of word class changing inflection). In English, conversion cannot be signaled morphologically due to lack of infinitival morphology.

At this point of the discussion, let us compare word-structure (word syntax) and phrase structure (phrasal syntax) with respect to branching structure and directionality. In phrasal syntax, left-branching is ruled out for the projection line of a phrase. This amounts to a uniform branching structure, viz. a unidirectional, right-branching buildup of phrases, with the base position of the head at the end of the phrase which is simultaneously the bottom.

Head-final or head-initial is the reflex of the parameterized licensing relation. Licensing is directional and the directionality value is a lexically determined property of heads. Head-initial phrases turn out to be necessarily more complex than head-final ones, since the initial head is in a fronted position and in relation to its ‘trace’ in the base position. In other words, complex head-initial phrases employ a shell-structure.

Word syntax is organized in a parallel fashion, with the exception of shells. Left branching is ruled out and there is parameterized directionality of licensing. However, the counterpart of a shell-structure is necessarily unavailable for complex word structures. Word structure elements are not subject to word-internal displacement rules (viz. antecedent-trace configurations). As a consequence, complex left-headed word structures cannot be generated since this would presuppose a shell-structure. Left-headed word structures are restricted to the trivial format of a head and an element in the sister position (Haider 2001):

- (35) a. $[A_{\pi} X]_{\pi}$
b. $*[[A_{\pi} X]_{\pi} Y]_{\pi}$

If (35a) is the structure of a two member compound, branching direction does not come into play, but for (35b) it would. The result would be a left-branching structure, which is ruled out. Immediate confirmation comes from compounding in Romance languages. As mentioned in

the introductory section, nominal compounds are left-headed in these languages and recursive compounds are restricted to the format of (35a). Here are examples from Spanish and Catalan:

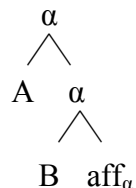
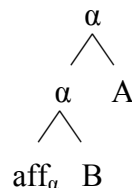
- (36) a. perro policía - *policía perro Spanish
 dog police (police dog)
 b. hombre aranya - *aranya hombre Catalan
 man spider (spider man)
 c. *perro policía mascota Spanish
 dog police pet (pet police dog)

Piera (1995) uses examples like (36c) as evidence for the ungrammaticality of recursive nominal compounding in Spanish. The same is true for the other Romance languages. In fact, no language with recursive nominal compounding is known whose compounds are head-initial.³² This contrasts sharply with the productivity of recursive head-final nominal compounds in other languages.

In the case of affixation, the ban against left-branching would arise only if an affix had more than one dependent or modifying element, but this is not the case. Neither a structure like (37a) nor (37b) is a recursive left-branching structure. In each case there are two separate affixes, each one combining with a single, possibly complex element.

- (37) a. [Aff_β [Aff_α X]_α]_β
 b. [[X Aff_α]_α Aff_β]_β

How would genuinely recursive structures with an *affix* as the base element look like? (38a) is a right-branching structure with a suffix as base. (38b) is the left-branching counterpart with a prefix as the base element.

- (38) a.  b. 

The word structure (38a) has a right-branching projection line, while (38b) has a left-branching one. Hence it is ruled out by the universal restriction against left-branching projection lines (see 21i). It is not clear, however, whether one of these structures is implemented at all in natural languages, with an affix as base. If B is a word that gets affixed by joining with the affix, then A is merged with a word, consisting of B plus affix. So A is not in the same relation to the affix as B is. It seems that nominal compounding with a noun as base element (in the place of the affix in (38) is the only case in which a recursive word structure is needed, and in these cases (38b) is correctly excluded.

Note that neither iterative prefixing nor iterative suffixing generates the type of right-branching word structures that are excluded, as (39) shows, except maybe for iterating *the*

³² Hoeksema (1992:126) discusses Vietnamese as a potential counterexample. He cites multi-member head-initial compounds but they are not *recursive* (see Haider 2013:196). They can be analyzed as two-member compounds whose members are (lexicalized) two-member compounds themselves.

very same suffix (39c). The excluded word structure is (35b). A single base element is merged consecutively with material on its right.

- (39) a. $[Aff_{\alpha} [Aff_{\beta} X]_{\beta}]_{\alpha}$ prefixes as base elements
b. $[[X Aff_{\beta}]_{\beta} Aff_{\alpha}]_{\alpha}$ suffixes as base elements
c. $*[[X Aff_{\alpha}]_{\alpha} Aff_{\alpha}]_{\alpha}$ iterated suffix as base element

In the case of (39c), the iteration of the suffix would produce kind of a recursive left-branching structure. However, it is safe to conjecture that the English prefix-iteration in (40) does not have a counterpart in any language in which the analogous affixes are suffixes.

- (40) a. anti-Semite, anti-anti-Semite, ...
b. pro-British, pro-pro-British, ...

A well-known case of suffix-iteration is the formation of diminutives. However, as expected, the iterated diminutive suffixes are different, as the examples from Spanish, Russian and German illustrate (41). Each suffix forms a separate diminutive word.

- (41) a. burr-o (donkey) > borr-ic-o > borr-iqu-it-o
b. Katya > Katyusha > Katyushka > Katyushichka > Katyushenichka
c. Esel (donkey) > Eselein > Eseleinchen (*Esel-lein-lein; *Esel-chen-chen)

In sum, there is directionality in word structure, too. A category-changing affix that serves as the base of a complex word may be a suffix or a prefix. This corresponds to the parametric directional licensing by lexical heads in phrasal syntax. *Recursive* left-branching structures are excluded in word structuring just like in phrase structuring. There is no counterpart of complex head-initial projections in word structure, however, since word syntax does not provide the means of forming head-initial shell structures.

7. Summary

Directionality-based conditions are operative in phrase structuring in two different respects. On the one hand, there is the cross-linguistically uniform directionality of ‘merger’. Merger is to the left; hence the resulting phrases are right-branching. Left-branching projections lines are not admitted. On the other hand there is the parameterized directionality of licensing. The parameter values are the alternative directions of licensing associated with lexical head classes, which yield head-initial and head-final structures, respectively.

The more complex organization of head-initial structures (‘shell-structure’) is the direct consequence of the mismatch between the universal branching direction and the licensing direction of the head in head-initial structures in an otherwise uniform mode of licensing (mutual, minimal c-command in the canonical direction).

If the lexical directionality value is not specified, phrase structuring allows for alternative positions of the head of a phrase. This is the case in a large sample of the languages of the world. The Slavic languages are representative of this setting.

Word structure follows the universal uniformity of branching. They are right-branching. Recursive expansions are always head-final. The head-initial counterparts of phrase structures do not come into being because word-structure does not provide the means for shell-structures.

Appendix: LCA (Kayne 1994) versus BBC (Haider 1992)

In the early nineties, Haider (1991, 1992) and Kayne (1992, 1994) proposed complementary solutions for the very same hypothesis. The hypothesis was this: phrase structures are universally asymmetric. They are right-branching.

In Haider (1991, 1992), the asymmetry is covered by the BBC ('basic branching constraint'), as an axiom of syntactic structuring. Kayne (1992, 1994) derives the asymmetry from an axiomatic *trans-derivational* constraint, viz. the LCA ('linear correspondence axiom'). This constraint requires that grammars be organized in such a way that the linearization of the terminals becomes a function, that is, a unique mapping, of structural asymmetries on linearization asymmetries.

Interestingly, the LCA can be met by alternative functions from structure to linearization, namely by (i) '*precede*' or by (ii) '*follow*'. If α asymmetrically c-commands β , then the terminals dominated by α will *precede* the terminals dominated by β in the implementation (i). However, in the implementation (ii), the terminals of α will *follow* the terminals of β . The two variants yield complementary grammar families but one family does not materialize in the set of human languages, as it seems. Grammars of natural languages do not employ (ii). So, in the LCA theory, the *mapping directionality* must be stipulated.

In fact, the LCA predicts *three* types of structuring. Type (i) would uniquely map c-command on '*precede*'. Type (ii) would uniquely map c-command on '*follow*'. Type (iii) would be the type with category-specific mapping relations ('*precede*' for one category, '*follow*' for another category) that produce differentiated linearization patterns, analogous to German and Dutch head-positioning in NP vs. VP. Followers of the LCA theory tacitly assume that languages are of type (i) only and overlook the other options. This tacit assumption is equivalent to the tacit assumption of a directionality requirement.

The LCA and the BBC theory handle head-initial versus head-final phrase structures in complementary ways, namely in terms of phrasal chains (LCA) versus head-chains (BBC). The LCA theory derives head-final structures from head initial ones by means of *phrasal movement*. Any postverbal item is fronted. In the evacuated phrase, the head ends up as the rightmost terminal of the phrase. In the BBC account, head-initial versus head-final is a property of the positioning of the *head* of a phrase. Complex head-initial phrases involve head-chains.

In the LCA system, head-final phrases are more complex than head-initial ones, since they are derivational continuations of head-initial phrases. In the BBC model, head-initial phrases are more complex than head-final ones since complex head-initial phrases consist of shell structures. This perfect complementarity opens numerous ways of empirically testing the relative explanatory success of the two models. These predictions are evaluated in detail in Haider (2013, chapter 9). It turns out that the LCA-based predictions are sans empirical support. The following table lists the major issues; for the details see Haider (2013, ch.9).

LCA predictions (empirically unsupported)	BBC predictions (empirically supported)
1. In OV, preverbal phrases are in Spec positions and therefore <i>opaque</i> for extractions like a preverbal VO subject. Hence in OV, all <i>preverbal arguments</i> are predicted to be opaque.	In OV, preverbal phrases are <i>transparent</i> for extraction since they are in the domain of their selecting head, and therefore transparent like postverbal objects in VO.
2. <i>Scrambling</i> places phrases into preverbal Spec-positions and thereby makes them opaque for extractions.	Scrambling is adjunction within the directionality domain of the head. Scrambled phrases remain transparent for extraction.
3. OV and VO languages are expected to display the ‘EPP-effect’.	The ‘EPP-effect’ is an SVO phenomenon and absent in SOV.
4. <i>Diachronically</i> , an SOV language develops out of an SVO language. SVO is the universal structure that gets masked by massive phrasal fronting in the change from VO to OV.	Diachronically, SOV and SVO are inverse settings. Typically, the OV or VO grammar develops out of a T3 grammar, by fixing the directionality of heads.
5. <i>Edge effect</i> is not admitted: <i>Adverbials</i> and modifier phrases (e.g. adnominal <i>attributes</i>) are phrases in Spec-positions and behave syntactically like phrases in spec positions.	Edge effect is admitted: Adverbial and attributive phrases are adjoined to the phrases they modify. The presence/absence of the edge effect immediately correlates with directionality.
6. ‘Immobile’ items remain in a <i>postverbal</i> position.	‘Immobile’ items remain at the directionally canonical side of the head they are dependent on, i.e. preverbal in OV, postverbal in VO.
7. <i>No adjunction</i> . Spec-head-complement is the universal building scheme. Adjunction structures are admitted, neither for phrase-to-phrase, nor for head-to-head adjoining. Multiple specs or multiple phrases per spec are required.	Adjunction is a standard structural configuration. Adverbial or attributive phrases are instances of phrase to phrase adjunction. Verbal clusters are instances of head-to-head adjunction. Multiple adjunctions are predicted.
8. No <i>lexically</i> specifiable <i>directionality</i> . Cross-categorial linearization differences are differences in terms of fronting items. <i>Example</i> : Clause-final complementizers are clause-initial, but their complement has been fronted.	Directionality values are lexically specifiable. Cross-categorial linearization differences may be differences in terms of the directionality value. <i>Example</i> : Lexical complementizers with the directionality value ‘regressive’ are clause-final.
9. <i>No nominalizations</i> of the series of verbal heads in either VO or OV, since they are heads of stacked VPs in each case.	Nominalization of verbal clusters in OV are nominalizations of the same series of verbal heads that project stacked VPs in VO.

Bibliography

- ANS (1984) see: Geerts et als. (1984)
- Bailyn, John (2002). Inversion, Dislocation and Optionality in Russian. In Gerhild Zybatow, U. Junghanns, G. Mehlhorn & L. Szucsich (eds), *Current Issues in Formal Slavic Linguistics*, Frankfurt/Main: Peter Lang. 280-293.
- Barss, Andrew & Lasnik, Howard (1986) 'A Note on Anaphora and Double Objects', *Linguistic Inquiry* 17: 347-354.
- Bayer, Josef (1998) 'Final complementizers in hybrid languages', *Journal of Linguistics* 35: 233–71.
- Bayer, Josef (2001) 'Two grammars in one: sentential complements and complementizers in Bengali and other South Asian languages', in Peri Bhaskararao and Karumuri V. Subbarao (eds.), *The yearbook of South Asian languages: Tokyo symposium on South Asian languages – contact, convergence and typology*, New Delhi: Sage Publications, pp. 11–36.
- Bech, Gunnar (1955) *Studien über das deutsche verbum infinitum*, Copenhagen: Munksgaard. (Second unchanged edition in 1983, introduced by Cathrine Fabricius-Hansen, Tübingen: Niemeyer)
- Biberauer, Theresa, Anders Holmberg & Ian Roberts (2007). Disharmonic word-order systems and the Final-over-Final-Constraint (FOFC). In Antonietta Bisetto & F. Barbieri (eds). *Proceedings of XXXIII Incontro di Grammatica Generativa*. p. 86-105. Bologna: Università di Bologna (http://amsacta.unibo.it/2397/1/PROCEEDINGS_IGG33.pdf)
- Bierwisch, Manfred. 1990 Verb cluster formation as a morphological process. In: Booij, G. & J. van Marle (eds.): *Yearbook of Morphology*. Dordrecht: Foris. 173-199.
- Bošković, Željko (1997) Fronting Wh-Phrases in Serbo-Croatian, in M. Lindseth & St. Franks (eds.) *Formal Approaches to Slavic Linguistics: The Indiana Meeting 1996*. Ann Arbor: Michigan Slavic Publications, 86-107.
- Chomsky, Noam (1981) *Lectures on Government and Binding*, Dordrecht: Foris
- Chomsky, Noam (1982) *Some concepts and consequences of the theory of government and binding*. Cambridge, Mass., MIT Press.
- Den Dikken, Marcel and Bernstein, Judy B. and Christina Tortora and Raffaella Zanuttini 2007. Data and grammar: Means and individuals. *Theoretical Linguistics* 33: 335–352
- Dixon, Robert M.W. (2011). *I am a linguist*. Leiden: Brill.
- Dryer, Matthew S. 2013. Prefixing vs. suffixing in inflectional morphology. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/26>, accessed 2014-03-18.)
- Dryer, Matthew S. (2009) 'The Branching Direction Theory of Word Order Correlations Revisited' in Sergio Scalise, Elisabetta Magni, Antonietta Bisetto (eds.), *Universals of Language Today*, Berlin: Springer. p.185-207.
- Embick, David and Roumyana Izvorski (1997). Participle-Auxiliary Word Orders in Slavic. In: W. Browne, N. Kondrashova and D. Zec (eds.), *Formal Approaches to Slavic Linguistics: The Cornell Meeting 1995*, 210-239. Ann Arbor: Michigan Slavic Publications.
- Engdahl, Elisabet (1986). *Constituent Questions*. Dordrecht: Reidel.
- Geerts, Guido, Walter Haeseryn, Jaap de Rooij and Maarten C. van den Toorn (1984) *Algemene Nederlandse Spraakkunst*, Groningen: Wolters-Noordhoff.

- Greenberg, Joseph H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. in: *Universals of language*. Joseph H. Greenberg (ed.), 73–113. Cambridge: MIT Press.
- Haider Hubert (in press). “Intelligent design” of grammars – a result of cognitive evolution. In: Aria Adli & Marco García García & Göz Kaufmann (eds.): *System, Usage, and Society*. Berlin/New York: de Gruyter.
- Haider, Hubert 2013. *Symmetry breaking in syntax*. Cambridge: Cambridge University Press.
- Haider, Hubert 2010. *The syntax of German*. Cambridge: Cambridge University Press.
- Haider, Hubert 2004. Pre-and postverbal adverbials in VO and OV. *Lingua* 114: 779-807.
- Haider, Hubert. 2001. Why are there no complex head-initial compounds? In: Chris Schaner-Wolles & John Rennison & Friedrich Neubart eds. *Naturally!* Torino: Rosenberg & Sellier (165-174).
- Haider, Hubert. 2000. Adverb placement – convergence of structure and licensing. *Theoretical Linguistics* 26: 95-134
- Haider, Hubert. 1992. Branching and Discharge. *Working Papers of the SFB 340*, #23. Univ. Stuttgart. [2000 in Peter Coopmans & Martin Everaert & Jane Grimshaw eds. *Lexical specification and insertion*. Amsterdam: Benjamins (p. 135-164)]
- Haider, Hubert 1991. On the structure of argument structure - and its projection. Lecture delivered at the *Conference on Lexical Specification and Lexical Insertion*, Univ. Utrecht December 1991 (published as Haider 1992 & 2000).
- Haider H. & Luka Szucsich (in press). Scrambling and V-positioning in Slavic languages – exceptionally VO or regular T3? In: *The German Middle Field in a Comparative and Diachronic Perspective*. Roland Hinterhölzl, Kristine Bentzen, Augustin Speyer & Luka Szucsich (eds.) Berlin: Mouton de Gruyter.
- Hankamer, Jorge (1974). On Wh-Indexing. *NELS* 5 (p. 61-76).
- Haspelmath, Martin (1996). "Word-class-changing inflection and morphological theory." In: Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1995*. Dordrecht: Kluwer, 43-66.
- Hoeksema, Jack (1992) ‘The head parameter in morphology and syntax’ in Dicky G. Gilbers and S. Looyenga (eds.) *Language and cognition 2*. Groningen: De Passage, p. 119-132.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*, Cambridge, Mass.: MIT Press
- Kayne, Richard S. 1992. Word Order. Keynote Lecture at the GLOW Colloquium, Lisbon, April 14, 1992.
- Kiparsky, Paul (1996) The Shift to Head-Initial VP in Germanic. In Höskuldur Thraínsson, Samuel Epstein, and Steve Peter (eds.), *Studies in Comparative Germanic Syntax II*. Dordrecht: Kluwer, pp. 140–79.
- Koopman, Hilda (1984). *The syntax of verbs*. Dordrecht: Foris Publications.
- Kornai, András & Geoffrey K. Pullum 1990. The X-bar theory of phrase structure. *Language* 66: 24-50.
- Lasnik, Howard. 2001. A Note on the EPP. *Linguistic Inquiry* 32 (2): 356-362.
- Larson, Richard (1988) ‘On the double object construction’, *Linguistic Inquiry* 19:335-391.
- Mallinson, Graham & Barry J. Blake, 1981. *Language typology: Cross-linguistic studies in syntax*. Amsterdam: North-Holland.

- Meyer, Roland (2004) *Syntax der Ergänzungsfrage. Empirische Untersuchungen am Russischen, Polnischen und Tschechischen*. München: Otto Sagner.
- Piera, Carlos 1995. "On Compounding in English and Spanish". In Héctor Campos, Paula Marie Kempchinsky (eds.). *Evolution and Revolution in Linguistic Theory*. Washington, D.C.: Georgetown University Press. pp. 302-315.
- Santorini, Beatrice (1993). The rate of phrase structure change in the history of Yiddish. *Language Variation and Change* 5: 257-283.
- Schallert, Oliver. 2014. *Zur Syntax der Ersatzinfinitivkonstruktion. Typologie und Variation*. (Studien zur Deutschen Grammatik 87). Tübingen: Stauffenburg.
- Singh, Uday N. (1980) 'Bole: an unresolved problem in Bengali syntax', *Indian Linguistics* 41: 188-195.
- Stepanov, Arthur (2007). The end of CED? Minimalism and extraction domains. *Syntax* 10: 80–126
- Szabolcsi, Anna and Frans Zwarts (1993) 'Weak islands and an algebraic semantics for scope-taking', *Natural Language Semantics* 1: 235–85.
- Toman, Jindřich (1981) Aspects of multiple wh-movement in Polish and Czech, in: Robert May & Jan Koster (eds.) *Levels of Syntactic Representation*. Dordrecht: Foris Publications, 293-302.
- Vikner, Sten (1995) *Verb Movement and Expletive Subjects in the Germanic Languages*, Oxford: Oxford University Press.
- Williams, Edwin. 1981. On the notions lexically related and head of a word. *Linguistic Inquiry* 12: 245–274.
- Zwart C. Jan-Wouter (1994) Dutch is head-initial. *The Linguistic Review* 11(3/4): 377-406.