A Construction-based Cross-linguistic
Analysis of V2 Word Order

Andrew C. Wetta

University at Buffalo
The State University of New York

Proceedings of the HPSG 2011 Conference

Department of Linguistics, University of Washington

Stefan Müller (Editor)

2011

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

Verb second (V2) word order is determined by considering the absolute position of clausal constituents. Previous accounts of such word order in HPSG have been developed for individual V2 languages (predominantly German) but are often not cross-linguistically applicable. I propose a set of generalized mechanisms in linearization-based SBCG which accounts for cross-linguistic V2 data by use of: (1) a simple two-valued feature rather than many-typed topological domains, (2) domain compaction, and (3) constructionally-determined domain positions. Not only does this analysis account for V2 placement, but it can also model verb third (V3) placement and other positionally-stipulated word orders.

1 Introduction

Verb second (V2) word ordering is defined by the appearance of the finite verb in the second position, determined by considering the absolute position of all clausal constituents. Such clauses exhibit a degree of flexible constituent order allowing a variety of elements, such as the subject or objects, to appear in the single position before the finite verb. Thus, it often becomes difficult to characterize such languages as SVO or OVS, as there are many possible permutations of syntactic elements, that is, there may be no dominant word order (cf. Dryer, 2011). This interplay between relatively free word order and a positionally-strict verbal position provides a challenge for syntactic analyses, particularly those based on phrase structure grammars.

The V2 phenomenon is most thoroughly examined and associated with German. However, there are other languages, including non-Indo-European ones, which also attest this type of word ordering. In order to provide a complete account of the phenomenon, these additional languages require equal examination so that a full characterization of V2 as a cross-linguistic phenomenon may be developed. As such, the syntactic structures of a genealogically broad sampling of V2 languages are considered, including Breton, German, Ingush, Karitiâna, Kashmiri, and Yiddish.

Using the insight from this language sampling, which is briefly summarized in this paper, it is possible to determine the syntactic structures which account for the occurrence of V2 word order and the degree to which these structures are shared among the languages, consequently enabling cross-linguistic generalizations of the phenomenon as a whole to emerge. These generalizations will be formalized in a linearization-based (Reape, 1994, 1996; Kathol and Pollard, 1995; Müller, 1996; Kathol, 2000) version of Sign-Based Construction Grammar (SBCG) (Sag, 2010;

[†]For helpful comments and suggestions I would like to thank Jeff Good, Jean-Pierre Koenig, Stefan Müller, and Ivan Sag. I owe a special thanks to Rui Chaves for his detailed comments and numerous discussions. Additionally, I am grateful to three anonymous reviewers and the audience of the HPSG 2011 conference. All remaining errors are solely my own.

Boas and Sag, to appear). This combination is particularly well suited to describe V2 languages because it both allows flexible constituent order via domains and linear precedence rules as well as the definition of constructions to restrict the positioning of clausal elements.

The analysis presented in this paper advocates the use of minimally-defined constructions which capture the constraints of this cross-linguistic word order phenomenon while remaining compatible with other language-dependent constructions and rules. This approach is in contrast to other analyses which utilize more restrictive mechanisms, such as topological fields, or extraction schemata and instead directly generalizes the structures attested in cross-linguistic data.

To begin, I will present a brief summary of the clause structures in the sampling of V2 languages and provide pertinent data in §2. In §3 previous analyses for V2 word order will be examined, and then in §4 I will describe a generalized construction-based analysis highlighting the mutually shared linearization mechanisms of the languages.

2 Verb Second Clause Structure

The constituent order of V2 languages is often relatively flexible, which allows many options for the linearization of elements. Naturally there are often pragmatic factors which control the order, but syntactically many variants are permissible. However, in all permutations, the finite verb is restricted to a particular position, such as the second position immediately after a single constituent, and may not be displaced like the other clausal elements. The example in (1) illustrates this interaction between flexible constituent order and the restriction of the finite verb to the second position, where the finite verb is shown in boldface.

(1) a. Peter **wollte** dem Jungen das Buch schenken.
Peter want.3SG.PST the.DAT boy.DAT the.ACC book give.INF

b. Dem Jungen **wollte** Peter das Buch schenken.

c. Das Buch **wollte** Peter dem Jungen schenken.

d. Schenken **wollte** Peter dem Jungen das Buch.

'Peter wanted to give the book to the boy.' *German* (Uszkoreit, 1987, 156)

In this particular German sentence, which characterizes V2 clause structure, the finite verb is consistently after exactly one constituent while all other elements may be flexibly placed with respect to syntactic constraints. Formally, following the definition by Anderson (2005, 179), a V2 clause is characterized by the verb with tense, mood, and agreement properties, if available, (i. e. the finite verb) appearing in the second position immediately after one constituent.

Although this V2 phenomenon is most cited with Germanic languages, most notably German but including Danish, Dutch, Icelandic, and Yiddish, among others, it also occurs in other non-Germanic languages such as Breton (Celtic), Ingush

(Nakh-Daghestanian), Karitiâna (Tupian), Kashmiri (Indic), and Romansch (Romance) as illustrated by examples (2)–(4). Additionally, the sentence in (4) shows how the first element may be an entire clause.

(2) akhbaar **por** laRkan raath
newspaper read boy yesterday
'It was the newspaper that the boy read yesterday' *Kashmiri* (Bhatt, 1999, 137)

(3) he boued **e** **tebr** Mona er gegin
her food PRT eat.3SG Mona in.the kitchen
'Mona eats her food in the kitchen' *Breton* (Press, 1986, 197)

(4) [boroja taso oky tykiri] **Ø-naka-hyryp-Ø** õwã
snake man kill PFV 3-PRT-cry-NFUT child
'When the man killed the snake, the child cried.' *Karitiâna* (Storto, 2003, 414)

2.1 Clause type asymmetries

Even though a language may employ V2 word order, it may not be applied to all clause types. That is, subordinate and question clauses, among others, may exhibit different finite verb placements than verb second positioning. For example, the Kashmiri sentence in (5) contains a subordinate clause which maintains V2 word order, not including the subordinator, however the Breton sentence in (6) attests a verb initial subordinate clause word order. The difference between verb placement in main and subordinate clauses is often called *root-subordinate asymmetry*, because each clause type exhibits different finite verb placements, but differences also extend beyond just these two clause types. Thus, the position of the finite verb is patterned by the clause type and is a necessary component of sentence structure for a V2 language.

(5) tem-is chu afsoos [ki yi kitaab **cha-yi**
he-DAT be.3SG.M regret.PRS.PTCP that this book be.F-2SG
tse par-mets]
you.F.SG.ERG read-PST.PTCP
'He regrets the fact that it is this book that you have read.' *Kashmiri* (Bhatt, 1999, 100)

(6) gwelout a reas Lenaig [**e** **save** an dour]
see.INF PRT do.PST.3SG Lenaig PRT rise.PST.3SG the water
'Lenaig saw the water was rising.' *Breton* (Stephens, 2002, 399)

2.2 Verbal elements

Although the finite verb must appear in the second position of a V2 clause, the non-finite verbs are realized in many different locations. For instance, Ingush, like German, places non-finite verbs at the end of a clause as in (7), but Breton commonly places the non-finite verbs either in the first position or immediately after the finite verb as in (8).

(7) Muusaa **vy** hwuona telefon *jettaxh*
Musa V.PROG you.SG.DAT telephone strike.CVB
'It's Musa on the phone for you.' *Ingush* (Nichols, 2009)

(8) E voued **en** **deus** *debret* Yann er wetur
his food 3SG.M have.PRS.3 eat.PST.PTCP Yann in.the car
'Yann has eaten his food in the car.' *Breton* (Press, 1986, 200)

Non-finite verbs have more flexibility in Yiddish and may appear in any position, that is, immediately after the finite verb, between arguments and adjuncts as in (9), or at the end of the clause. Additionally some V2 languages have constructions which allow non-finite verbs to be placed in the first position either alone or in groups such as a partial verb phrase like in example (10).

(9) m'**hot** durx ale fentster *arojssgehangn* weʃ
one=have.3SG through all windows out.hung.PST.PTCP laundry
'Out of all the windows one hung the laundry.' *Yiddish* (Weissberg, 1988, 153)

(10) [Das Buch *schenken*] **wollte** Peter dem Jungen.
the.ACC book give.INF want.PST.3SG Peter the.DAT boy
'Peter wanted to give the book to the boy.' *German* (Uszkoreit, 1987, 156)

2.3 Multiple first elements

In other instances, clause types may display a similar verb third (V3) order as with the sentences in (11) and (12) where the finite verb appears in the third position after two initial constituents. The German example presents an alternative word order from the usual V2 for main clauses. However, the Kashmiri content question clause must be V3 where a single constituent as well as the question word appear before the finite verb.

(11) [Zum zweiten Mal] [die Weltmeisterschaft] **errang** Clark 1995
the.DAT second time the world.championship win.1SG.PST Clark 1995
'Clark won the world championship for the second time in 1995.' *German* (Beneš, 1971) quoted from (Müller, 2005b)

(12) raath kyaa **dyut-na-y** rameshan
yesterday what.NOM give.PST.M.SG-3SG.ERG-2SG.DAT Ramesh.ERG
tse
you.DAT
'As for yesterday, what is it that Ramesh gave you?' *Kashmiri* (Bhatt, 1999, 107)

Additionally, even more elements could appear before the finite verb in certain contexts to form clause orders of V4, V5, and so forth (cf. Müller, 2003).

2.4 Summary

The data presented in the previous sections show the typical form of a V2 clause, that finite verb placement is dependent upon the clause type, and that the positioning of non-finite verbs varies in each language. Table 1 summarizes the possible finite verb placements by clause type for six V2 languages examined in an extensive typological survey[1] which I undertook. The new analysis of V2 word order presented in §4 generalizes the syntactic structures from this survey.

	Main: Affirmative	Negative	Subordinate: Content	Relative	Question: Content	Polar
Breton	V2	V_I(V2)	V_I	V_I(V2)	V2	V2(V3/V_I)
German	V2(V3)	V2(V3)	V_F(V_I)	V_F	V2	V_I
Ingush	V2(V3)	V2	V_F	V_F	V2	V2
Karitiâna	V2/V_I(V3)	V2/V_I	V_F	V_F	V2	V2/V_I
Kashmiri	V2	V2	V2	V_F	V3	V2(V_I/V3)
Yiddish	V2(V3)	V2	V2	V2(V_I)	V2	V_I(V2)

Table 1: Verb placement in various clause types. Non-basic alternative word orders appear in parentheses. (V_I = verb initial and V_F = verb final)

3 Previous Analyses

Previous analyses of V2 word order in HPSG (Pollard, 1996; Kathol, 2000; Borsley and Kathol, 2000; Richter and Sailer, 2001; Müller, 2002) generally fall somewhere on the spectrum between a purely linearization and extraction-based approach. The extraction-based approach accounts for flexible constituent order by motivating the movement or displacement of constituents to other locations in a clause by the application of additional phrase structure schemata. This most notably occurs with the movement of a single constituent to the first position immediately before the finite verb of a V2 clause. The linearization-based account posits the separation of syntactic structure and surface word order via *word order domains* (Reape, 1994, 1996). This separation allows the stipulation of a constituent's location without needing to motivate a parallel process in the syntactic structure. Thus, a single constituent's domain may be relegated to the first position without modifying the clause's phrase structure. This approach reflects the intuition that the same syntactic processes occur despite linear order.

[1]The sources for the typological survey include: Bhatt (1999), Borsley and Kathol (2000), Dudenredaktion (2005), Jacobs et al. (1994), Landin (1982), Landin (1984), Müller (2003), Nichols (2009), Press (1986), Stephens (2002), Storto (2003), Uszkoreit (1987), Wali and Koul (1997), Weissberg (1988)

Collectively these analyses all draw upon a common set of mechanisms to account for V2 clause structure: word order domains, linear precedence rules, an INVERTED feature, SLASH, topological fields, and constructions. Yet some of these mechanisms are redundant and perform similar functions. For example, the INVERTED feature and SLASH as well as domains and linear precedence (LP) rules both allow variety in the linear realization of elements. Similarly, topological fields and constructions both provide the means to constrain clausal elements in particular configurations.

The analysis I propose here follows a more strict linearization-based approach than previous analyses and requires no local extraction to the first position to account for V2 word order or a HEAD-FILLER SCHEMA to mark the topological field of the first element. Furthermore, in order to avoid the redundancies among many of the syntactic mechanisms and to provide an appropriately flexible yet succinct description which generalizes the linearization behavior of all V2 languages, I utilize only word order domains, LP rules, and constructions to stipulate clause-internal word order. This means I do not employ a topological field model or, in the case of a V2 clause, extraction via the HEAD-FILLER SCHEMA to the first position. I examine this selection of mechanisms in the next two sections.

3.1 Problems for topological fields

The topological field model, drawn from traditional grammar, provides a precise and accurate way in which to describe the word order of German. But this model becomes problematic when it is applied to other languages (cf. Kathol, 2000, 285) and increases the difficulty for cross-linguistic generalization. Consider the traditional order of topological fields for German cast into LP rules in (13) by Kathol (2000, 79), which describes the word order placement fields of a sentence.

(13) TOPOLOGICAL LP STATEMENT
Vorfeld $\prec$ *complementizer field* $\prec$ *Mittelfeld* $\prec$ *verb cluster* $\prec$ *Nachfeld*

This typological field schema presents the following problems when accounting for word order in the V2 languages reviewed in §2:

Competition between finite verb and complementizer In order to account for *root-subordinate asymmetries* (cf. §2.1), the finite verb and complementizer compete for the *complementizer field* (i. e. the second position): Only one of these elements may occupy the field and the complementizer takes precedence. If a complementizer is present, then the finite verb must appear in the only other verbal field, the *verb cluster* (i. e. clause final position). But in the case of a main clause, which has no complementizer, the finite verb is realized in the *complementizer field*. This competition describes clausal word order in asymmetric V2 languages, like German, but is inaccurate for symmetric languages like Yiddish and Kashmiri. In these languages the finite verb always appears in the *complementizer field* and the complementizer appears before the *Vorfeld* (i. e. the first position).

Non-finite verbs This model places all non-finite verbal elements in the *verb cluster*. But, Karitiâna and Breton often maintain a linearly contiguous verb phrase, in which case all the verbs, including non-finite ones, remain in the second position.

Post-verbal objects In some V2 languages, such as Yiddish, non-finite verbs may appear as a group among the non-verbal elements and not clause final. That is, non-verbal elements may appear both before and after the *verb cluster* thus effectively splitting the *Mittelfeld*. Because the *Nachfeld* in the topological field schema is for extraposed elements, there is no place to put objects after the non-finite verbs.

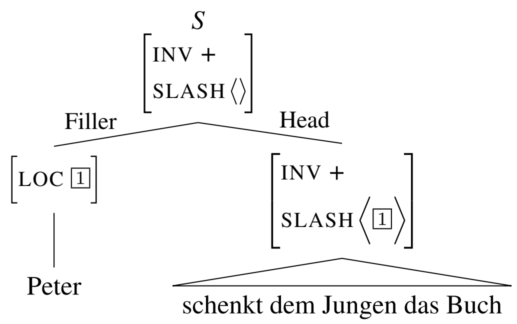
Thus, it is hard to extend this topological model, which was originally intended for German, to other V2 languages. Various modifications have been proposed to adapt the topological field model to other languages (Kathol, 2000; Borsley and Kathol, 2000), but no uniform and generalized model exists for all V2 languages. So, it is unclear if such a model may be used when describing a generalized V2 word order placement. Instead, I use constructions in my analysis to determine the clausal positions of constituents.

3.2 Problems for extraction

Many analyses utilize the HEAD-FILLER SCHEMA to front a constituent before a clause-initial finite verb to effectively produce V2 word order as a result of extraction,[2] which is illustrated in Figure 1. The INVERTED feature is also used in this example to displace the finite verb *schenkt* 'gives' from clause final position to clause initial. The HEAD-FILLER SCHEMA is typically associated with a class of constructions that link a filler to an arbitrarily embedded gap such as topicalization, relative clauses, and wh-interrogatives, all of which license otherwise impossible word orders, particularly in English. However, given the flexible constituent order of V2 languages and the ability of constituents to shuffle under normal circumstances as *word order domain* elements, it is possible to realize V2 word order without this schema.

The HEAD-FILLER SCHEMA subsumes a set of constructions which allow unbounded extraction, that is, the realization of arbitrarily embedded elements in an alternative location, usually clause initial. For example, non-subject wh-interrogatives in English are realized as a filler in the first position. However, subject wh-interrogatives are a type of SUBJECT-HEAD CONSTRUCTION (cf. Sag, 2010, 533) and do not require extraction to alter word order. Similarly, because word order domains allow any clausal element to appear in the first position via shuffling in V2 clauses, the HEAD-FILLER SCHEMA need not be employed to alter word order and realize the initial element.

[2] Although the analysis proposed by Kathol (2000) uses domains and LP rules to realize an element in the *Vorfeld*, this element is assigned to the *Vorfeld* by virtue of being the filler of the HEAD-FILLER SCHEMA (p. 85)



"Peter gives the boy the book."

Figure 1: Accounting for V2 with extraction.

There is cross-linguistic evidence which indicates that all wh-interrogatives, even subject ones, are reflected in the morphosyntax as extraction phenomena (Hukari and Levine, 1995; Bouma et al., 2001), which could indicate that first elements should be extracted. For instance, Yiddish verb inversion in embedded relative clauses (Diesing, 1990) is cited as part of this evidence, where the expletive *es* appears in the first position (after the wh-interrogative) before the finite verb in the absence of any other element. For instance, in (14a) an expletive is inserted to maintain the V2 word order of the subordinate clause, that is, the extracted wh-interrogative is unable to fill the first position as only local elements may satisfy the V2 word order requirements.

(14) a. Ikh veys nit [ver es **iz** gekumen].
I know.1SG not who.NOM EXPL be.3SG come.PST.PTCP
b. *Ikh veys nit ver iz gekumen.
'I don't know who came.' *Yiddish* (Diesing, 1990, 68)

(15) Ver **hot** gegesn dos broyt?
who.NOM have.3SG eat.PST.PTCP the bread
'Who ate the bread?' *Yiddish* (Diesing, 1990, 52)

However, this expletive is not used in content question clauses, which also utilize wh-interrogatives as in (15). Here the wh-interrogative is indeed able to fill the first position, reserved for local elements, thus suggesting that extraction is not used to license this clause. Thus, is seems that the obligatory extraction of a particular element is clause specific and should not be reflected in the basic mechanisms for the realization of V2 word order.

Finally, if the HEAD-FILLER SCHEMA were utilized to realize the first element of a V2 clause, it is unclear where this construction would appear in a FILLER-HEAD CONSTRUCTION hierarchy as illustrated in Figure 2. None of these constructions appropriately predict V2 in all of its instances. A TOPICALIZATION CONSTRUCTION would indeed allow V2 word order but also includes the corresponding prosodic and pragmatic information associated with topicalization, which are not appropriate for pragmatically focused constituents or elements with no emphasis, in particular, expletives. For instance, the expletive ‘es’ in German may appear locally in the first position as in (16a), however, it may not be topicalized to a matrix clause as shown in (16c).

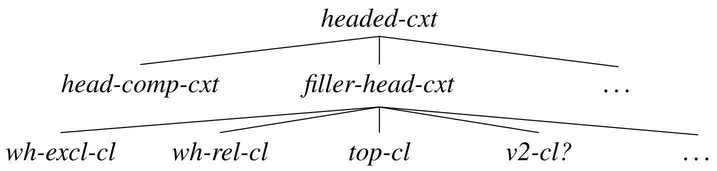


Figure 2: Placement of a V2-FILLER-HEAD-CONSTRUCTION.

(16) a. Es regnet in der Stadt.
EXPL rain.3SG.PRS in the city
‘It is raining in the city.’ *German*

b. [In der Stadt]$_i$ sagt er, dass es $\_i$ regnet.
in the city say.3SG.PRS he COMP EXPL rain.3SG.PRS
‘In the city, he said, that it’s raining.’ *German*

c. *Es$_i$ sagt er, dass $\_i$ in der Stadt regnet.

So, some V2-FILLER-HEAD CONSTRUCTION would need to be posited to allow V2 word order without any additional prosodic or pragmatic information. Additionally, because the first element must be realized clause internally (i.e. it may not appear in a higher matrix clause), this V2-FILLER-HEAD CONSTRUCTION would need to be constrained so that the filler could not cross clausal boundaries so that it would in fact be a bounded dependency. Such constraints are clearly very different than those of the TOPICALIZATION CONSTRUCTION. Thus, a HEAD-FILLER SCHEMA approach would require the definition of at least two nearly identical constructions.

The analysis I propose here avoids the over-generalization of extraction as well as the redundancies between HEAD-FILLER SCHEMAS and word order domains to realize V2, and instead captures the V2 word order by using only word order domains. Unbounded dependencies still exist under my analysis and are compatible with a V2 clause (cf. §4.2.1), but extraction is not necessary to realize V2.

4 Constructionally-Determined Word Order

Conceptually, the generalized analysis I propose here places all constituents of a clause into a *word order domain*. These domain elements are by default *flexible*, that is, able to shuffle, via Reape’s shuffle operator ‘◯’, and produce a variety of word orderings from a single set of domains. However, constructions may place positional restrictions on particular domain elements by specifying that they are *fixed* and stipulating their linear position within a clause. Linear precedence rules may only affect *flexible* domain elements and do not interact with *fixed* elements. In this way, free word order and strict positional stipulations may simultaneously exist within a single clause. Thus, a V2 construction would specify that the finite verb is *fixed* and must appear in the second position. All other *flexible* elements may then shuffle around this *fixed* verb, which is exempted from linear precedence constraints.

Formally, I describe this generalized analysis within the Sign-Based Construction Grammar (SBCG) framework (Sag et al., 2003; Sag, 2010; Boas and Sag, to appear). As such, I incorporate *domains* into the structure of a *sign*, like Reape (1994, 1996), via a DOM attribute which itself is a list of *signs*. Re-formulating the *Constituent Ordering Principle*, as shown in (17), a *sign's* FORM is then the concatenation of the FORM values of its domain elements.

(17)
$$\textit{sign} \Rightarrow \begin{bmatrix} \text{FORM } \mathrm{L}_1 \oplus \mathrm{L}_2 \oplus \ldots \oplus \mathrm{L}_n \\ \text{DOM} \left\langle \begin{bmatrix}\text{FORM } \mathrm{L}_1\end{bmatrix}, \begin{bmatrix}\text{FORM } \mathrm{L}_2\end{bmatrix}, \ldots, \begin{bmatrix}\text{FORM } \mathrm{L}_n\end{bmatrix} \right\rangle \end{bmatrix}$$

In §4.1 I will first describe the generalized mechanisms necessary for a construction-based analysis of V2 word order: (1) a simple two-valued feature rather than many-typed topological domain elements, (2) domain compaction, and (3) constructionally-determined domain positions. Then, in §4.2 I will outline grammar fragments to illustrate how these mechanisms license clause structure in V2 languages.

4.1 Generalized Mechanisms

4.1.1 Two-typed domain elements

In order to facilitate the division between *flexible* and *fixed* domain elements, I introduce a new attribute LIN with *linearization* values: *flexible* and *fixed*, as depicted in (18). This LIN attribute is part of a domain *sign* and has a default value of *flexible* defined by the constraint in (19). *Persistent Default Unification*, as described by Lascarides and Copestake (1999), is employed to ensure that the default value remains a part of the feature structure during unification and may be realized in a fully licensed construct when no other value overrides it, namely *fixed*. That is, unless otherwise specified, the *linearization* value of a domain element in a construct is *flexible*. The *fixed* value is only assigned by constructions to override the default *flexible* value. [3]

[3]It may be desirable to avoid using defaults, which could be done in two ways: (1) Some elements could be lexically marked *fixed* leaving all others underspecified. However, the same element may be *fixed* in one construction but *flexible* in another. Also, there should be no underspecified LIN attributes in a fully licensed clause so that the LP rules, which only affect *flexible* elements, behave properly. (2) Constructions could explicitly specify all potential elements as *fixed* or *flexible*, which means many clause constructions would stipulate lists of *flexible* elements to account for any other possible items. This ensures that all domain elements do not remain underspecified. But in order for the word order constructions defined here to appropriately interact with each other and correctly

(18) *linearization*

$$\text{linearization} \to \textit{flexible} \quad \textit{fixed}$$

(19) $$\textit{sign} \Rightarrow \left[\text{DOM } \textit{list}\left(\left[\text{LIN } /\textit{flexible}\right]\right)\right]$$

In this way, linear precedence rules may only affect domain elements with a LIN value of *flexible*, as illustrated by the sample LP rule in (20). This allows *fixed* domain elements to remain in a constructionally-determined position without affecting the placement of the other *flexible* elements.

(20) $$\begin{bmatrix}\text{LIN} & \textit{flexible}\\ \text{FOCUS} & -\end{bmatrix} \prec \begin{bmatrix}\text{LIN} & \textit{flexible}\\ \text{FOCUS} & +\end{bmatrix}$$

4.1.2 Domain compaction

Following Reape, there are two kinds of DOMAIN CONSTRUCTIONS: LIBERATING, which keeps the daughter domain elements of a construction independent in the mother, and COMPACTING, which, like Kathol and Pollard (1995) and Donohue and Sag (1999), creates a *single* new domain element in which all the daughter domain elements may still shuffle. Compaction allows LP rules to still affect the order of the domain elements in the mother's domain, but forces them to act as a single unit in any further construction. Thus, the compacting mechanism enables multiple elements, when appropriate and specified by language-specific constructions, to form a single domain element which may appear in a single constructionally-determined domain position.

(21) a. $$\textit{liberating-domain-cxt} \Rightarrow \begin{bmatrix}\text{MTR} & \left[\text{DOM } \mathrm{L}_1 \bigcirc \ldots \bigcirc \mathrm{L}_n\right]\\ \text{DTRS} & \left\langle \left[\text{DOM } \mathrm{L}_1\right], \ldots, \left[\text{DOM } \mathrm{L}_n\right] \right\rangle\end{bmatrix}$$

b. $$\textit{compacting-domain-cxt} \Rightarrow \begin{bmatrix}\text{MTR} & \left[\text{DOM } \left\langle \left[\text{DOM } \mathrm{L}_1 \bigcirc \ldots \bigcirc \mathrm{L}_n\right] \right\rangle\right]\\ \text{DTRS} & \left\langle \left[\text{DOM } \mathrm{L}_1\right], \ldots, \left[\text{DOM } \mathrm{L}_n\right] \right\rangle\end{bmatrix}$$

Compaction is vital for an analysis of flexible word order because it allows the definition of linear constituents which may not correspond to the phrase structure. This distinction is particularly salient with partial compaction (Kathol and Pollard, 1995; Yatabe, 1996), a mixture of the liberating and compacting domain constructions where only some of the daughter domains are compacted. This type of compaction is further explored in §4.2.1.

---

license clauses, they must remain silent about these other items. So, defaults seem to be necessary. Yet, it may still be possible to avoid defaults with method (2), which is compatible with the theory presented here. This is something which warrants further investigation.

4.1.3 Cross-linguistic constructions

Drawing from the constructional approach taken by Kathol (2000, Ch.7), my analysis similarly constrains clauses by a combination of *linear* and *sentence mode* constructions. Using the attested linearization patterns in V2 languages, I propose a general set of common clausal constructions for word order determination, provided in Figure 3, which describe the mutually occurring syntactic constraints in all V2 languages. The *sentence mode* constructions license various clause types such as declarative, relative, and interrogative. And as illustrated in §2.1, the clause type patterns the position of the finite verb in a clause, thus making the *sentence mode* a necessary component when specifying linear order. Each language independently stipulates the combination of *linear* and *sentence mode* constructions which license a complete clause.

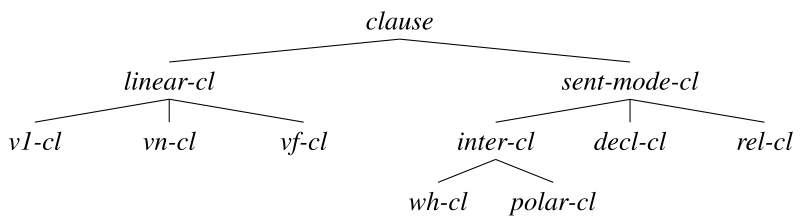


Figure 3: Hierarchy of clausal constraints common to all V2 languages.

The linear clause constraints are formally defined by the rules in (22)–(24). Each of these constructions explicitly states the location of the domain for the finite verb. The V1 and VF-CLAUSE CONSTRUCTIONS straight-forwardly stipulate that the domain element with the finite verb form must appear either clause initially or finally, respectively. Notice that the finite verb domain element is constructionally stipulated to be *fixed* and may be a phrase, that is, a complex predicate.

(22) a. In a verb initial clause, the domain element with the finite verb appears *before* all other domain elements.

b. $$\textit{v1-cl} \Rightarrow \left[\text{MTR}\left[\text{DOM}\left\langle\begin{bmatrix}\text{LIN} & \textit{fixed}\\ \text{SYN} & \left[\text{CAT}\left[\text{VFORM}\ \textit{finite}\right]\right]\end{bmatrix}\right\rangle \oplus \ldots\right]\right]$$

(23) a. In a verb final clause, the domain element with the finite verb appears *after* all other domain elements.

b. $$\textit{vf-cl} \Rightarrow \left[\text{MTR}\left[\text{DOM} \ldots \oplus \left\langle\begin{bmatrix}\text{LIN} & \textit{fixed}\\ \text{SYN} & \left[\text{CAT}\left[\text{VFORM}\ \textit{finite}\right]\right]\end{bmatrix}\right\rangle\right]\right]$$

The VN-CLAUSE CONSTRUCTION in (24) must not only specify the position of the finite verb domain element, but must also limit the number and types of elements that precede it so that V2 or V3 may be realized. In the absence of any other constructions to specify *fixed* domain elements before the finite verb, only one element appears before the verb, namely a *flexible* element, thus creating V2 word order. If there is an additional construction specifying *fixed* elements before the finite verb, it then becomes possible to define V3 word order or, for that matter, V4, V5, and so on. The VN-CLAUSE CONSTRUCTION is remarkable in that it licenses all placements of the finite verb in some nth position from the beginning of a clause in exactly the same way.

(24) a. In a clause which positions the finite verb domain element in the nth position from the beginning, the finite verb is preceded by exactly one *flexible* domain element and any number of *fixed* domain elements, in any order, and followed by all other domain elements.

b.

$$\textit{vn-cl} \Rightarrow \left[\textsc{mtr}\ \left[\textsc{dom}\ \left(\textit{list}\left(\left[\textsc{lin}\ \textit{fixed}\right]\right) \bigcirc \left\langle\left[\textsc{lin}\ \textit{flexible}\right]\right\rangle\right) \oplus \left\langle\begin{bmatrix}\textsc{lin} & \textit{fixed}\\ \textsc{syn} & \left[\textsc{cat}\ \left[\textsc{vform}\ \textit{finite}\right]\right]\end{bmatrix}\right\rangle \oplus \ldots\right]\right]$$

Finally, for all V2 languages which attest complementizers, these elements are not shuffled with a clause's word order domains and must instead be positionally stipulated by the COMPLEMENTIZER CONSTRUCTION as shown in (25), which is like the HEAD-FUNCTOR CONSTRUCTION. This construction concatenates a *fixed* complementizer domain to the beginning of a saturated clause's domain list. Here, SELECT indicates which expression the complementizer modifies, following Sag (to appear). Thus, the correct position of the complementizer is specified without interfering with a clause's word order. This separate COMPLEMENTIZER CONSTRUCTION is posited in order to avoid overgeneralizing the values of the LINEARIZATION features in other HEAD-FUNCTOR CONSTRUCTIONS.

(25)

$$\textit{complementizer-cxt} \Rightarrow \begin{bmatrix}\textsc{mtr} & \begin{bmatrix}\textsc{syn}\ X\\ \textsc{dom}\ L_1 \oplus L_2\end{bmatrix}\\ \textsc{dtrs} & \left\langle\begin{bmatrix}\textsc{syn} & \left[\textsc{cat}\ \begin{bmatrix}\textit{comp}\\ \textsc{select}\ H\end{bmatrix}\right]\\ \textsc{dom} & L_1{:}\left\langle\left[\textsc{lin}\ \textit{fixed}\right]\right\rangle\end{bmatrix}, H{:}\begin{bmatrix}\textsc{syn}\ X{:} & \begin{bmatrix}\textsc{cat}\ \left[\textsc{vform}\ \textit{finite}\right]\\ \textsc{val}\ \langle\rangle\end{bmatrix}\\ \textsc{dom}\ L_2 & \end{bmatrix}\right\rangle\\ \textsc{hd-dtr} & H\end{bmatrix}$$

4.2 Language-specific clause licensing

The use of the generalized mechanisms to describe the clause structure in a particular language may be illustrated by a fuller hierarchy of PHRASAL CONSTRUCTIONS in Figure 4. The HEADED CONSTRUCTIONS, adopted from SBCG (Sag, 2010, to appear), are not necessarily shared among V2 languages, but illustrate where they may exist in the phrasal hierarchy. Two variants of the HEAD-COMPLEMENT CONSTRUCTION are used: The PREDICATIONAL CONSTRUCTION combines a head with one or more of the items on its VALENCE list, but not all. Whereas the SATURATIONAL CONSTRUCTION combines a head with all of the remaining elements on its VALENCE list and licenses a complete clause. Thus, constructs may now be fully licensed by a combination of HEADED, DOMAIN, LINEAR-CLAUSE, and SENTENCE-MODE-CLAUSE CONSTRUCTIONS.

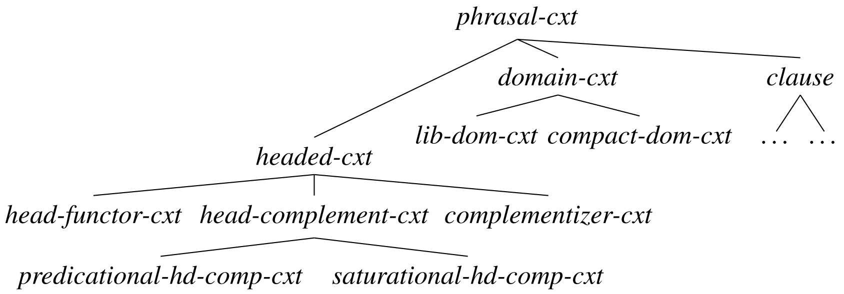


Figure 4: Partial hierarchy of phrasal constructs for V2 languages.

I will first briefly illustrate the use of the generalized mechanisms to license various word order phenomenon in German in §4.2.1, as this will enable an easy comparison to previous analyses. Then in §4.2.2 I will sketch out analyses in Kashmiri and Breton.

4.2.1 German

Consider the clausal hierarchy for German in Figure 5, which utilizes the common clausal constraints from Figure 3.[4] The bottom row in this hierarchy represents a sampling of complete clause constructs, which are a combination of the *linear* and *sentence-mode* types.

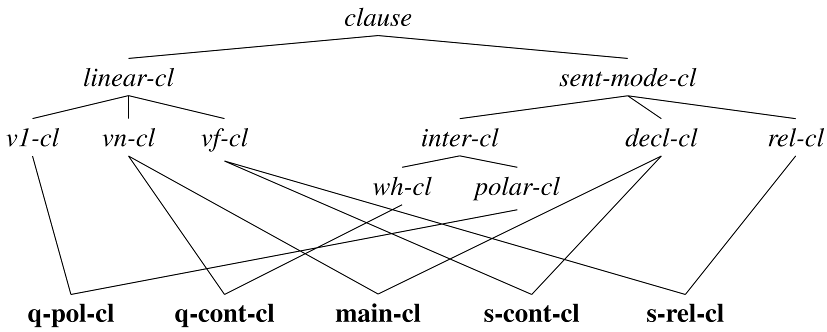


Figure 5: Partial hierarchy of clausal constructs for German.

[4]In all clausal hierarchies the following abbreviations are used to conserve space: s(ubordinate), cont(ent), rel(ative), q(uestion), pol(ar).

Given the language-specific phrasal constructions in (26), which also contain generalized constructions discussed in §4.1, it is possible to license a V2 main clause such as in example sentence (1b).

(26) Some PHRASAL CONSTRUCTIONS for German

a. *lib-pred-hd-comp-cxt* ⇒ *predicational-head-complement-cxt* ∧ *liberating-domain-cxt*

b. *main-lib-sat-hd-comp-cl* ⇒ *saturational-head-comp-cxt* ∧ *declarative-cl* ∧ *liberating-domain-cxt* ∧ *vn-cl*

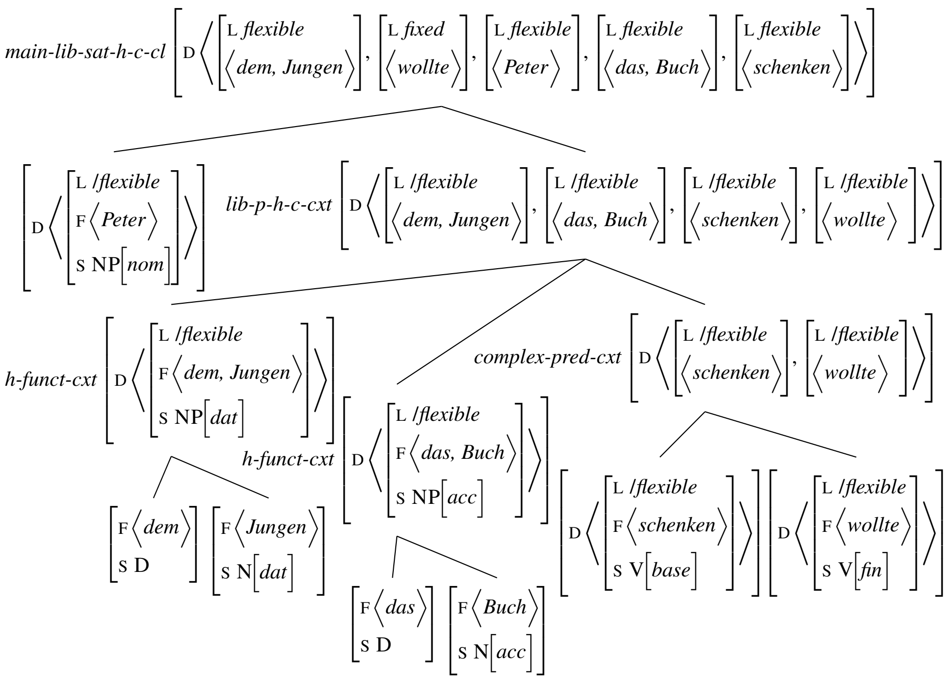


Figure 6: Clause structure for German V2 sentence.

The structure of this V2 sentence is illustrated in Figure 6.[5] Here, the MAIN-LIB-SAT-HD-COMP-CLAUSE CONSTRUCTION licenses the saturation of the finite verb's complement list while keeping all of the domain elements liberated and free to shuffle except for the finite verb itself, which is constructionally specified as *fixed* and relegated to the position after a single *flexible* domain, as according to the VN-CLAUSE CONSTRUCTION. A COMPLEX-PREDICATE CONSTRUCTION is used to create a verbal complex which combines all of the arguments from both verbs (cf. Hinrichs and Nakazawa, 1998, *inter alia*). Language-dependent LP rules determine the positions of the *flexible* elements, such as constraining the non-finite verb domain element to the end of the clause. Naturally, other constructions could

[5] Abbreviations will also be used in AVMs to conserve space: D(OM), L(IN), F(ORM), S(YN), C(AT), VF(ORM).

be defined to stipulate the non-finite verb domain element as *fixed* in a different position.

Similarly, the V3 sentence in (11) can be licensed by the same VN-CLAUSE CONSTRUCTION with further language-specific constraints, such as the DISCOURSE-PROMINENCE CONSTRUCTIONS in (27). These constructions utilize partial compaction, as mentioned in §4.1.2, which allow the first two elements before the finite verb to form a single domain element despite not forming a phrase structure constituent.

(27) DISCOURSE PROMINENCE CONSTRUCTIONS for German

a. $$\textit{doms}_{\bigcirc}\left(\left\langle \left[\text{DOM } \mathrm{X}_1\right], \ldots, \left[\text{DOM } \mathrm{X}_n\right] \right\rangle\right) \equiv \mathrm{X}_1 \bigcirc \ldots \bigcirc \mathrm{X}_n$$

b. *prom-part-compact-dom-cxt* ⇒

$$\begin{bmatrix} \text{MTR} & \begin{bmatrix} \text{DOM} \left\langle \begin{bmatrix} \text{PROM} + \\ \textit{doms}_{\bigcirc}\left(\mathrm{L}_1\right) \end{bmatrix} \right\rangle \bigcirc\ \textit{doms}_{\bigcirc}\left(\mathrm{L}_2\right) \end{bmatrix} \\ \text{DTRS} & \mathrm{L}_1{:}\textit{list}\left(\left[\text{PROM} +\right]\right) \bigcirc \mathrm{L}_2{:}\textit{list} \end{bmatrix}$$

c. *prom-main-cl* ⇒ *main-lib-sat-hd-comp-cl* ∧ *prom-part-compact-dom-cxt*

The PROMINENCE-PARTIAL-COMPACTION-DOMAIN CONSTRUCTION shown in (27b) appeals to a common discourse-oriented feature which compacts the prominent elements into a single domain. Here this discourse feature is represented by a binary PROM(INENCE) attribute. However, this construction and new feature are only used for illustrative purposes and do not necessarily reflect a pragmatic analysis, instead they only show how such an analysis is compatible with the other word order constraints proposed in this paper. So, using the new $\textit{doms}_{\bigcirc}$ function

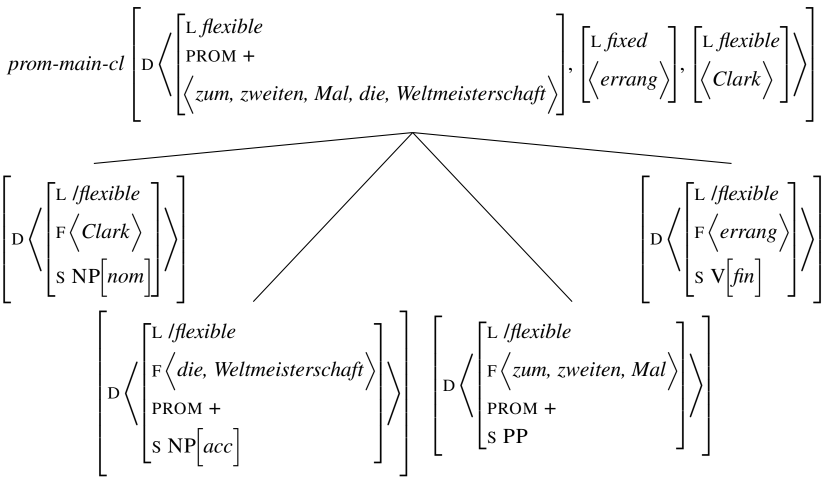


Figure 7: Clause structure for German V3 sentence.

defined in (27a), this construction stipulates that discourse prominent elements are compacted while all other elements remain liberated. Linear precedence rules subsequently cause the single prominent domain element to appear in the clause initial position. Figure 7 illustrates this clause structure for the V3 sentence in (11).

Finally, although the HEAD-FILLER SCHEMA is not used to realize the first element of a basic V2 clause, it still allows the non-local extraction of an embedded element and appropriately interacts with the VN-CLAUSE CONSTRUCTION to license a complete sentence. For instance, consider the sentence in (28) with an element extracted out of the subordinate clause into the first position of the matrix clause. Using the TOPICALIZATION CONSTRUCTION in (29) the clause structure for this complex sentence is illustrated in Figure 8.

(28) [Um zwei Millionen Mark]$_i$ versucht er [eine Versicherung $\_{i}$ zu
of two million Mark try.1SG.PRS he a insurance to
betrügen]
defraud.INF
'Of two million Marks, he is trying to defraud an insurance company.'
German (adapted from Müller, 2005a)

(29) TOPICALIZATION CONSTRUCTION for German
top-main-cl ⇒ *main-lib-sat-hd-comp-cl* ∧ *filler-head-cxt*

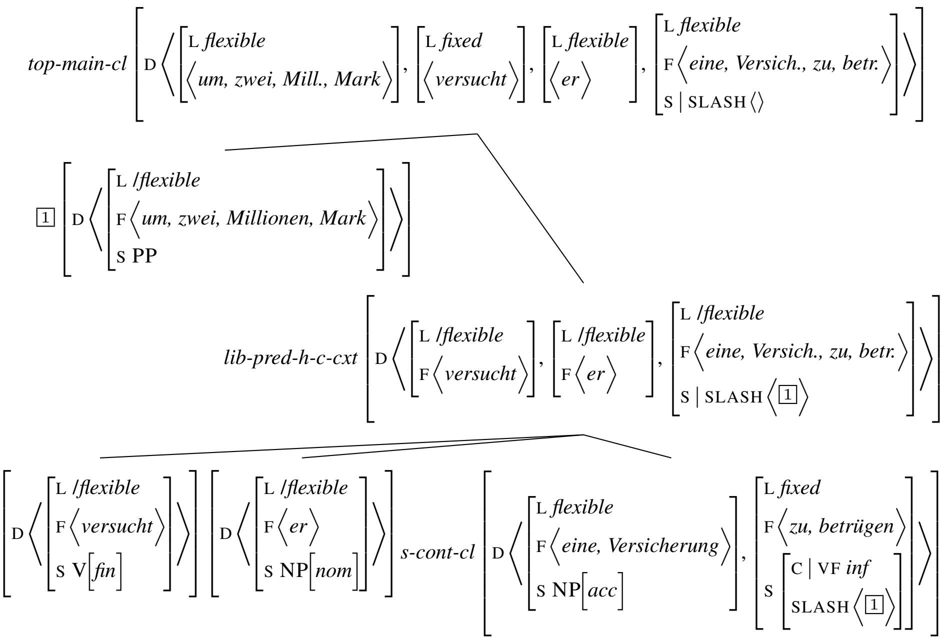


Figure 8: Clause structure for German sentence with topicalized element.

4.2.2 Other V2 Languages

The same generalized mechanisms extend to other V2 languages. For instance, the common clausal constraints are also used by Kashmiri in Figure 9 to define its clausal constructs. Notice that the linking of LINEAR and SENTENCE-MODE CLAUSAL CONSTRUCTS here are different than for German. Thus, a Kashmiri question clause, which is obligatorily V3 such as in example (12), may be licensed by the language-specific construction in (30) which also utilizes the common VN-CLAUSE CONSTRUCTION and has the resulting domain structure in (31). This construction uniquely specifies a *fixed* question word domain element which appears before the finite verb, thus allowing verb third word order.

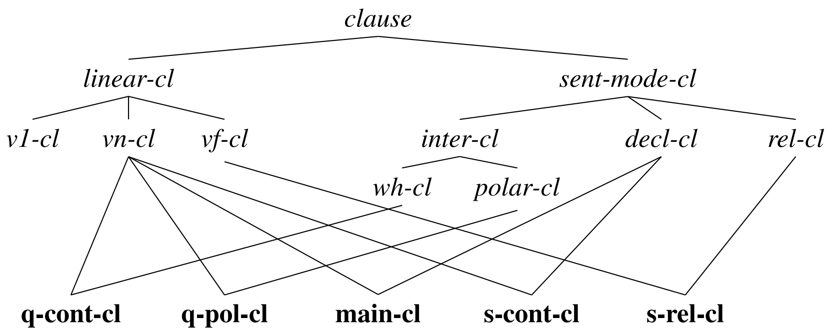


Figure 9: Partial hierarchy of clausal constructs for Kashmiri

(30) CONTENT QUESTION CONSTRUCTION for Kashmiri

$$\textit{cont-question-cl} \Rightarrow \textit{vn-cl} \wedge \textit{wh-cl} \wedge \left[\text{DOM}\left\langle \left[\text{LIN } \textit{flexible}\right], \begin{bmatrix}\text{LIN } \textit{fixed}\\ \text{SYN WH}\end{bmatrix}, \ldots \right\rangle\right]$$

(31)

$$\left[\text{D}\left\langle \begin{bmatrix}\text{L } \textit{flexible}\\ \text{F}\left\langle \textit{raath} \right\rangle\\ \text{S ADV}\end{bmatrix}, \begin{bmatrix}\text{L } \textit{fixed}\\ \text{F}\left\langle \textit{kyaa} \right\rangle\\ \text{S WH}\end{bmatrix}, \begin{bmatrix}\text{L } \textit{fixed}\\ \text{F}\left\langle \textit{dyutnay} \right\rangle\\ \text{S V}\left[\textit{fin}\right]\end{bmatrix}, \begin{bmatrix}\text{L } \textit{flexible}\\ \text{F}\left\langle \textit{rameshan} \right\rangle\\ \text{S NP}\end{bmatrix}, \begin{bmatrix}\text{L } \textit{flexible}\\ \text{F}\left\langle \textit{tse} \right\rangle\\ \text{S NP}\end{bmatrix} \right\rangle\right]$$

Additionally, domain compaction becomes important for the analysis of the Breton V2 clause in example sentence (8). When both of the finite and non-finite verbs are analyzed as a complex predicate, they may be compacted together to form a single domain element which is then correctly positioned by the VN-CLAUSE CONSTRUCTION as shown in (32).

(32)

$$\left[\text{D}\left\langle \begin{bmatrix}\text{L } \textit{flexible}\\ \text{F}\left\langle \textit{e, voued} \right\rangle\\ \text{S NP}\left[\textit{acc}\right]\end{bmatrix}, \begin{bmatrix}\text{L } \textit{fixed}\\ \text{F}\left\langle \textit{en, deus, debret} \right\rangle\\ \text{S VP}\left[\textit{fin}\right]\end{bmatrix}, \begin{bmatrix}\text{L } \textit{flexible}\\ \text{F}\left\langle \textit{Yann} \right\rangle\\ \text{S NP}\left[\textit{nom}\right]\end{bmatrix}, \begin{bmatrix}\text{L } \textit{flexible}\\ \text{F}\left\langle \textit{er, wetur} \right\rangle\\ \text{S PP}\end{bmatrix} \right\rangle\right]$$

5 Conclusion

By examining the mutually-shared characteristics of V2 languages it is possible to define the common mechanisms which accurately describe their word orders, namely: a shared set of LINEAR, SENTENCE-MODE, and DOMAIN CONSTRUCTIONS; *flexible* and *fixed* domain elements; language-specific constructions which specify *fixed* domain elements; domain compaction; and linear precedence rules which only affect *flexible* domain elements. In this paper I have shown that a linearization-based analysis can account for a variety of word ordering phenomena in V2 languages. Where traditional phrase structure rules are ill suited, using two-valued domain elements in combination with constructional stipulations, the interaction of flexible word order and strict positional constraints may be appropriately defined while remaining compatible with other phenomena such as non-local extraction.

References

Anderson, Stephen R. 2005. *Aspects of the Theory of Clitics*. Oxford: Oxford University.

Beneš, Eduard. 1971. Die Besetzung der ersten Position im deutschen Aussagesatz. In Hugo Moser (ed.), *Fragen der strukturellen Syntax und der kontrastiven Grammatik*, Sprache der Gegenwart — Schriften des IdS Mannheim, No. 17, pp. 160–182, Düsseldorf: Pädagogischer Verlag Schwann.

Bhatt, Rakesh Mohan. 1999. *Verb Movement and the Syntax of Kashmiri*. Dordrecht: Kluwer Academic Publishers.

Boas, Hans and Sag, Ivan A. (eds.). to appear. *Sign-Based Construction Grammar*. Stanford: CSLI Publications.

Borsley, Robert D. and Kathol, Andreas. 2000. Breton as a V2 language. *Linguistics* 38, 665–710.

Bouma, Gosse, Malouf, Robert and Sag, Ivan A. 2001. Satisfying Constraints on Extraction and Adjunction. In *Natural Language and Linguistic Theory*, vol. 19, pp. 1–65.

Bunt, Harry and van Horck, Arthur (eds.). 1996. *Discontinuous Constituency*. Berlin/New York: Mouton de Gruyer.

Diesing, Molly. 1990. Verb Movement and the Subject Position in Yiddish. *Natural Language and Linguistic Theory* 8(1), 41–79.

Donohue, Cathryn and Sag, Ivan A. 1999. Domains in Warlpiri. In *Sixth International Conference on HPSG–Abstracts*, pp. 101–106, Edinburgh.

Dryer, Matthew S. 2011. Order of Subject, Object and Verb. In Matthew S. Dryer and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*, Munich: Max Planck Digital Library.

Dudenredaktion. 2005. *Duden: Die Grammatik*, vol. 4. Mannheim: Dudenverlag.

Hinrichs, Erhard W. and Nakazawa, Tsuneko. 1998. Third Construction and VP Extraposition in German: An HPSG Analysis. In Erhard Hinrichs, Andreas Kathol and Tsuneko Nakazawa (eds.), *Complex Predicates in Nonderivational Syntax*, vol. 30 of *Syntax and Semantics*, pp. 115–157, San Diego, California: Academic Press.

Hukari, Thomas E. and Levine, Robert D. 1995. Adjunct Extraction. *Journal of Linguistics* 31(2), 195–226.

Jacobs, Neil G., Prince, Ellen F. and van der Auwera, Johan. 1994. Yiddish. pp. 388–419, London/New York: Routledge.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford: Oxford University.

Kathol, Andreas and Pollard, Carl. 1995. Extraposition via complex domain formation. In *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics*, pp. 174–180.

Landin, David J. 1984. An Outline of the Syntactic Structure of Karitiâna Sentences. In Robert Dooley (ed.), *Estudos Sobre Línguas Tupi do Brasil*, vol. 11, pp. 219–254, Brasilia: SIL Press.

Landin, Rachel M. 1982. Word Order Variation in Karitiâna. In *Arquivo Linguístico*, No. 149, Brasilia: SIL Press.

Lascarides, Alex and Copestake, Ann. 1999. Default representation in constraint-based frameworks. *Computational Linguistics* 25, 55–105.

Müller, Stefan. 1996. Yet Another Paper about Partial Verb Phrase Fronting in German. In *Proceedings of COLING 96*, pp. 800–805, Copenhagen.

Müller, Stefan. 2002. *Complex Predicates: Verbal Complexes, Resultative Constructions, and Particle Verbs in German*. Stanford, CA: CSLI Publications.

Müller, Stefan. 2003. Mehrfache Vorfeldbesetzung. *Deutsche Sprache* 31(1), 29–62.

Müller, Stefan. 2005a. Zur Analyse der deutschen Satzstruktur. *Linguistische Berichte* 201, 3–39.

Müller, Stefan. 2005b. Zur Analyse der scheinbar mehrfachen Vorfeldbesetzung. *Linguistische Berichte* 203, 297–330.

Nichols, Johanna. 2009. Germanic-like verb-second/verb-final word order in Nakh languages, handout at LSA Annual Meeting, San Francisco, Jan. 10.

Pollard, Carl J. 1996. On Head Non-movement. In Bunt and van Horck (1996), pp. 279–305.

Press, Ian. 1986. *A Grammar of Modern Breton*. Berlin: Mouton de Gruyter.

Reape, Mike. 1994. Domain Union and Word Order Variation in German. In John A. Nerbonne, Klaus Netter and Carl Jesse Pollard (eds.), *German in Head-Driven Phrase Structure Grammar*, pp. 151–197, Stanford, CA: CSLI Publications.

Reape, Mike. 1996. Getting Things in Order. In Bunt and van Horck (1996), pp. 209–254.

Richter, Frank and Sailer, Manfred. 2001. On the Left Periphery in German. In Detmar Meurers and Tibor Kiss (eds.), *Constraint Based Approaches to Germanic Syntax*, pp. 257–300, CSLI.

Sag, Ivan A. 2010. English Filler-Gap Constructions. *Language* 86(3), 486–545.

Sag, Ivan A. to appear. Sign-Based Construction Grammar: An informal synopsis. In Boas and Sag (to appear).

Sag, Ivan A., Wasow, Thomas and Bender, Emily M. 2003. *Syntactic Theory: A Formal Introduction*. Stanford: CSLI Publications.

Stephens, Janig. 2002. Breton. In Martin J. Ball and James Fife (eds.), *The Celtic Languages*, pp. 349–409, London/New York: Routledge.

Storto, Luciana R. 2003. Interactions between verb movement and agreement in Karitiana (Tupi stock). In *Revista Letras*, No. 60, pp. 411–433, Curitiba: Editora UFPR.

Uszkoreit, Hans. 1987. *Word Order and Constituent Structure in German*. Stanford: CSLI.

Wali, Kashi and Koul, Omkar N. 1997. *Kashmiri: A cognitive-descriptive grammar*. London/New York: Routledge.

Weissberg, Josef. 1988. *Jiddisch: Eine Einführung*. Bern: Peter Lang.

Yatabe, Shûichi. 1996. Long-Distance Scrambling via Partial Compaction. In M. Koizumi (ed.), *Formal Approaches to Japanese Linguistics 2*, pp. 303–317, Cambridge, MA.