6-2010

Verbal Complex Phenomena in West Central German: Empirical Domain and Multi-Causal Account

Shannon A. Dubenion-Smith
Western Washington University, shannon.dubenion-smith@wwu.edu

Follow this and additional works at: https://cedar.wwu.edu/mcl_facpubs
Part of the German Linguistics Commons, and the Modern Languages Commons

Recommended Citation
https://cedar.wwu.edu/mcl_facpubs/19

This Article is brought to you for free and open access by the Humanities at Western CEDAR. It has been accepted for inclusion in Modern & Classical Languages by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.
Verbal Complex Phenomena in West Central German: Empirical Domain and Multi-Causal Account

Shannon A. Dubenion-Smith
Western Washington University

This paper is a synchronic investigation of verbal complex phenomena in the West Central German dialects. The types of verbal complexes attested in 187 recordings of West Central German from the Zwirner Corpus are first classified and analyzed. A GoldVarb analysis reveals that in subordinate clause two-verb complexes, the factor groups syntagm and verbal prefix type have statistically significant effects on word order, while in main clause two-verb complexes the factor groups syntagm, verbal prefix type, and a grammatical correlate to focus have statistically significant effects. Taking as a point of departure Lötscher 1978 and Sapp 2007, a multi-causal account of verbal complex phenomena involving grammatical, functional, and performance factors is developed, with a close examination of the role of language processing in verbal complex formation. Following Hawkins’ (1994, 2004) notions of Early Immediate Constituents and Minimize Domains, it is argued that the type of verbal complex known as Verb Projection Raising can be linked in part to an advantage in language production. The paper concludes with a discussion of the speaker-hearer dichotomy as it relates to German clausal structure and the role of interfaces in the analysis of verbal complexes.*

1. Introduction.
A striking syntactic characteristic of the Continental West Germanic languages, including extraterritorial varieties such as Afrikaans and Pennsylvania Dutch, is the clustering of multiple verb forms in a VERBAL COMPLEX. Such complexes exhibit microparametric variation in the

* The present study is a condensed, revised version of portions of my doctoral dissertation (Dubenion-Smith 2008). I would like to thank Mark L. Louden, Shaw N. Gynan, and two anonymous reviewers for their helpful comments on earlier drafts of this work, as well as Petra Fiero for her assistance with several of the dialect examples. The usual disclaimers apply.

© Society for Germanic Linguistics
number, type, and ordering of elements that cluster, according to language, dialect, and idiolect (see den Besten & Edmondson 1983, Evers 1975, Haegeman & van Riemsdijk 1986, Schmid & Vogel 2004, and Wurmbraun 2006 among others). For example, the relative ordering of past participle, auxiliary, and modal in the verbal complex is fixed in standard German, as illustrated in 1.1

(1) Standard German

\[
\text{weil man nicht alles gelesen}
\begin{array}{c}
\text{haben}
\end{array}
\begin{array}{c}
\text{muss}_1
\end{array}
\begin{array}{c}
\text{3-2-1}
\end{array}
\] 

because one not everything read have must

‘because you don’t have to have read everything’ (Duden 2005:481)

By contrast, in standard Dutch and Bernese Swiss German more than one ordering is possible.

(2) Standard Dutch

a. dat hij niets gezien
\begin{array}{c}
\text{kan}_1
\end{array}
\begin{array}{c}
\text{hebben}_2
\end{array}
\begin{array}{c}
\text{3-1-2}
\end{array}

that he nothing seen can have

‘that he can’t have seen anything’ (ANS 1984:600)

b. dat hij niets kan
\begin{array}{c}
\text{gezien}_3
\end{array}
\begin{array}{c}
\text{hebben}_2
\end{array}
\begin{array}{c}
\text{1-3-2}
\end{array}

c. dat hij niets kan
\begin{array}{c}
\text{hebben}_2
\end{array}
\begin{array}{c}
\text{gezien}_3
\end{array}
\begin{array}{c}
\text{1-2-3}
\end{array}

(3) Bernese Swiss German

a. das er soll
\begin{array}{c}
\text{gseit}_3
\end{array}
\begin{array}{c}
\text{ha}_2
\end{array}
\begin{array}{c}
\text{1-3-2}
\end{array}

that he is supposed said have

‘that he is supposed to have said’ (Hodler 1969:689)

b. das er gseit
\begin{array}{c}
\text{ha}_2
\end{array}
\begin{array}{c}
\text{soll}_1
\end{array}
\begin{array}{c}
\text{3-2-1}
\end{array}

1 In this paper, I adopt the standard numbering system employed in the literature on verbal complexes. The subscripts on verb forms in the examples indicate their level of embeddedness at D-structure, assuming an OV base, and order of subcategorization. A higher number indicates a deeper level of embeddedness; \( V_x \) subcategorizes \( V_{x+1} \). Only those verbs in the complex itself are labeled with a subscript (that is, the finite verb that in main clauses has raised to C is not labeled).
The elements that appear in the verbal complex are not restricted to verb forms but may also include non-verbal material, such as adverbs, DPs, and PPs. Example 4 illustrates one such complex in which the PP *auf n Ruck* ‘onto the (his) back’ appears between the auxiliary *hâns* ‘have’ and past participle *brâcht* ‘brought’.

(4) Austrian German (Carinthia)

\[
\text{bis i n hâns} \quad \text{auf n Ruck} \quad \text{brâcht} \quad 1-\text{X-2}
\]

until I him have onto the back brought

‘until I got him onto his back’ (Patocka 1997:306)

Finally, a constituent may be extraposed from its canonical position in the inner field to the right of the verb cluster, as evidenced in 5.

(5) Austrian German (Upper Austria)

\[
\text{âft is-s eascht fai(n) putzt} \quad \text{wou(r)dn} \quad \text{mit de Messa} \quad 3-\text{2-X}
\]

then is-it first finely cleaned been with the knives

‘then it was first trimmed nicely with the knives’ (Patocka 1997:324)

Over the past several decades, verbal complex phenomena have attracted the attention of scholars from various research traditions. The results of this interest include, among other works, a thorough typology of the phenomena in the Bavarian dialects spoken in Austria (Patocka 1997), the treatment of verbal complex data in Dutch and Swiss German syntax atlases, in particular the *Syntactische Atlas van de Nederlandse Dialecten* (SAND 2009.2) and the *Syntaktischer Atlas der Deutschen Schweiz* (SADS), and numerous generative studies such as Haegeman 1992, a book-length investigation of West Flemish dedicated in large part to the verbal complex.

However, as Schmid & Vogel (2004:272) lament, “the number of (not only) generative explorations into the syntax of German dialects in Germany has been extremely small, and there is little hope that this situation will change soon.” Wurmbbrand (1999, 2000a, 2000b, 2002, 2004, 2006) has provided us with the most complete picture of verb clusters in varieties spoken across the German-speaking area, yet detailed

---

2 These elements are labeled “X” throughout the text.
data for the West Central German (WCG) dialects remain lacking. This state of affairs is reflected in the graphic from Patocka 1997:265 (see figure 1), which shows the distribution of the 2-3 vs. 3-2 order in what is known as the modal IPP construction.

![Figure 1](image-url)

Figure 1. Empirical gap in the distribution of the 2-3 vs. 3-2 order, modal IPP construction (Patocka 1997:265).

The question mark located squarely over the WCG dialect area (Westmitteldeutsch) clearly indicates an empirical lacuna to be filled.

What data we do have come mainly from subordinate clauses. With some exceptions, such as Patocka 1997, studies that examine verbal complexes exclude those from main clauses. Sapp (2006:5), for example, does so because “comparing verb clusters in main and subordinate clauses may be comparing unlike types.” As examples 4 and 5 above illustrate, verbal complexes in subordinate and main clauses differ

---

3 The WCG dialect region covers Luxembourg and the German states Hesse, Rhineland-Palatinate, and Saarland, portions of North-Rhine Westphalia, Thuringia, Baden-Württemberg, and Bavaria, as well as border areas in Belgium (Arlon) and France (Alsace, Lorraine).

4 In figure 1, the areas in which the 2-3 and 3-2 orders (for instance, *er hat müssen₂ arbeiten₃* vs. *er hat arbeiten₃ müssen₂* ‘he had to work’) predominate in base dialect are indicated by the vertical and horizontal lines, respectively.
crucially in that the former, but not the latter, include the finite verb. Nonetheless, data from verbal complexes in main clauses are rich, meriting proper documentation, and can give valuable insight into verbal complex formation, as this paper will demonstrate. Provided that main clauses are treated separately, their inclusion here should be unproblematic.

Previous research in generative frameworks has yielded a range of formal descriptive analyses that give important insight into how verbal complexes are derived. As Wurmbrand (2006) notes, however, what remains outstanding is an explanatory account of why verbal complexes arise. Beyond strictly formal syntactic analyses, there has been progress on this front: Lütscher (1978), on the basis of Zurich Swiss German data, has proposed a multi-causal approach to verbal complex phenomena involving grammatical, functional, and performance rules, and Sapp (2006, 2007) has made great strides in defining the role of functional rules with his investigation of focus in Early New High German (ENHG) verbal complexes. Yet the role of performance rules and the way in which the three rule types interact require further elucidation.

The purpose of this paper is twofold. First, it classifies and analyzes subordinate and main clause verbal complex phenomena in the WCG dialects spoken in the Federal Republic of Germany, attested in transcripts of 187 dialect recordings from the Zwierner Corpus.5 Second, it contributes to our theoretical understanding of verbal complexes, complementing Lütscher’s (1978) multi-causal account and Sapp’s (2007) investigation of the relationship between word order in the verbal complex and focus. Following Hawkins’ (1994, 2004) notions of EARLY IMMEDIATE CONSTITUENTS and MINIMIZE DOMAINS, this study links particular serializations to EXTRAPOSITION, a productive syntactic operation in German, and to processing constraints. I argue that the type of verbal complex exemplified in 4 above, known in the generative literature as VERB PROJECTION RAISING (VPR), can be linked in part to an advantage in language production.6 My analysis is compatible with

5 The WCG dialects spoken in Belgium, France, and Luxembourg have been excluded from this study.

6 Consider the WCG examples in (ii)a and (ii)b, derived from the underlying structures in (i)a and (i)b, respectively.
the concept of the KLAMMERNDEN VERFAHREN (bracketing) outlined in Ronneberger-Sibold 1994 and the so-called speaker-hearer dichotomy. In particular, I maintain that the verbal bracket, which according to Ronneberger-Sibold 1991 and 1994 serves to mark syntactic boundaries for the benefit of the hearer, can be disrupted through VPR partly because this entails a processing benefit for the speaker.

The remainder of the paper is structured as follows. In section 2, the WCG dialect area and corpus are introduced, followed by an empirical and statistical analysis of verbal complexes in sections 3 and 4. Sections 5, 6, and 7 outline a multi-causal account of these complexes with an emphasis on performance, building on Lötscher 1978 and Sapp 2007. The paper concludes with a summary in section 8.

2. Dialect Area and Corpus.
2.1. West Central German.
Since the 19th century, a number of scholars have made proposals for the division of the German dialect space (see, for example, Behaghel 1891,

(i) a. dass wir [Milch bringen$_{2}$] können$_{1}$
    that we milk bring can

    b. dass sie [da einen ordentlichen Korb voll Essen kochen$_{2}$] müssen$_{1}$
    that they there a decent basket full food cook must

(ii) a. dass wir Milch können$_{1}$ bringen$_{2}$
    that we milk can bring
    ‘that we can bring milk’

    b. dass sie da müssen$_{1}$
    that they there must

  einen ordentlichen Korb voll Essen kochen$_{2}$
  a decent basket full food cook
  ‘that they have to cook a decent basketful of food’

In (ii)a, an example of VERB RAISING (VR), only the verb bringen has raised from the embedded clause to the right of können, yielding a cluster comprising verbal elements alone, [können$_{1}$ bringen$_{2}$], with the 1-2 order. In (ii)b, an example of VPR, not only the verb kochen, but also part of its verbal projection, the object DP einen ordentlichen Korb voll Essen, have raised to the right of müssen, yielding the complex [müssen$_{1}$ einen ordentlichen Korb voll Essen kochen$_{2}$] with the 1-X-2 order.
Bernhardi 1844, Bremer 1892, Martin 1959, Mitzka 1943, and Wiesinger 1983). Niebaum & Macha (2005:217) consider Wiesinger’s (1983) division, which is based on a set of phonological and morphological features, the closest to linguistic reality since it clearly indicates the existence of transitional zones. As the authors point out, however, the majority of recent overviews of German dialects differentiate one dialect region from another on the basis of sound shift boundaries, resulting in a conception of dialect space in which one dialect area begins where another ends.\(^7\) The Institut für Deutsche Sprache in Mannheim, which maintains the material from which the data treated in this paper are drawn, organizes the sound recordings in its archives according to this type of division. Since the question of dialect differentiation is not central here, I therefore assume divisions of the WCG dialect area according to sound shift isoglosses, adopting those in Beckers 1980 as presented in figure 2 below.

WCG belongs to the central region of the High German dialect area bordering on Low German to the north, East Central German to the east, and Upper German to the south. The WCG region can be subdivided into two larger dialect areas along isogloss 6 in figure 2, the \textit{dat/das} line, namely Central Franconian and Rhine Franconian. The former comprises the smaller dialect areas Ripuarian and Moselle Franconian, the latter \textit{Pfälzisch} and Hessian.\(^8\) The dialects spoken in Lorraine (\textit{Lothringisch}) were not included in this study.

\(^7\) See Niebaum & Macha 2005:217, note 2 for an extensive list of works. Beckers (1980:468) points out one problem with such boundaries, namely that they are a system of division inherited from 19th-century historical linguistics to divide up the Old and Middle High German dialect space.

\(^8\) Since no one-word equivalent for the German dialect designation \textit{Pfälzisch} exists in English, I use the German term here. Note that Wiesinger (1983) uses the term \textit{Rheinfränkisch} to refer to the area labeled \textit{Pfälzisch} in figure 2.
Figure 2. West Central German dialect area (Beckers 1980:469).

2.2. The Corpus.

The corpus for the present study consists of 2,578 clauses from transcripts of 187 dialect recordings in the Datenbank Gesprochenes Deutsch (Spoken German Databank), maintained online by the Institut für Deutsche Sprache. The transcripts, classified as Vollmundart (base dialect), were taken from the Zwrinner Corpus compiled between 1955 and 1970 under the direction of Eberhard Zwrinner.9 These include all 73

9 Transcripts with the designation Vollmundart were chosen because they reflect the most base-dialectal form spoken in a given locality. According to Haas & Wagener (1992:4), Vollmundart is defined as “gut erhaltene, alte Ortsmundart, archaischer Laut- und Formenbestand” (well-maintained, old local dialect, archaic phonological and morphological inventory). Beyond this general description, however, no more specific criteria for this designation are given, nor
available transcripts from the Central Franconian dialect area and 114 transcripts from the Rhine Franconian dialect area. The number of transcripts from the Rhine Franconian dialect area totals roughly 175, yet only a subset of these was selected in order to strike a better balance between the number of transcripts from both major dialect regions. To this end, only one transcript from the oldest male and female speaker per Planquadrat (quadrant) was included for analysis.

![Figure 3. Quadrants represented in the corpus study (map from Haas & Wagener 1992:1611).](image)

is it clear how recordings in Vollmundart have been distinguished from those in Halbmundart (semi-dialect) and Regionalmundart (regional dialect), the other two categories. Note also that while the compilation of the Zwirner Corpus continued until 1970, the recordings on which this study is based were all made between 1955 and 1959.

10 Many recordings have not yet been transcribed and were therefore unavailable for analysis.

11 The Planquadrate are 256 km² quadrants into which the German dialect space can be divided for cataloging purposes. Each quadrant is numbered according to coordinates on both the X and Y axes, facilitating the identification of dialect by locality.
Figure 3 above indicates the quadrants from which at least one transcript was taken. The WCG dialect area as a whole is demarcated by the heavy black line. The smaller dialect areas, further distinguished by the medium-heavy lines, are labeled “a” (Ripuarian), “b” (Moselle Franconian), “c” (Hessian), and “d” (Pfälzisch).

The 2,578 total tokens include 1,450 subordinate and 1,128 main clauses that fall into four complex categories: subordinate clause two- and three-verb complexes and main clause two- and three-verb complexes. Only those tokens whose clausal status could be unambiguously determined were included in the corpus. Consider the examples in 6 and 7.

(6) dass die Leute sollen zufrieden sein mit uns    MC? 1-X-2-X?    ZWH55
that the people should be satisfied with us
‘that the people should be satisfied with us’

(7) wenn wir sollten in die Messe gehen    MC? 1-X-2?
whenever we should into the mass go
‘whenever we were supposed to go to mass’

Since no constituent appears between the subject die Leute ‘the people’ and the verb sollen ‘should’ in 6 or the pronominal subject wir ‘we’ and the verb sollten ‘should’ in 7, it is not possible to determine whether these are examples of main clauses with verb-second word order or subordinate clauses with the 1-X-2-(X) word order. Both clauses are introduced by a subordinating conjunction, dass and wenn, respectively; it is well known, however, that the subordinating conjunction weil may be followed by main clause word order in spoken German (see, for example, Gaumann 1983, Günther 1993, Louden 2000, Schvitalla 2003: 144–146, Wegener 1993). Thus, the type of conjunction alone is not sufficient to determine clausal status. For this reason, only clauses in which an element appears between the subject and first verb were included, as in 8 and 9.

---

12 While 28 subordinate clause and 21 main clause complexes involving zu + infinitive are attested in the corpus, these were excluded from the present analysis and will be treated in a separate study.
(8) dass man **so früh** musste\textsubscript{1} aufstehen\textsubscript{2} \hspace{1cm} 1-2
that one so early had to stand up \hspace{1cm} ZWI94
‘that you had to get up so early’

(9) dass wir **Schläge** hatten\textsubscript{1} gekriegt\textsubscript{2} \hspace{1cm} 1-2
that we blows had gotten \hspace{1cm} ZW9B8
‘that we had been spanked’

After excluding eight clauses that did not meet this criterion, all those remaining were checked against the audio recordings, entered into a database, and then coded and classified according to verb form and verb type for further analysis.

3. Results of the Empirical Analysis.
This section presents the results of an empirical analysis of verbal complexes attested in the corpus, namely, subordinate clause two- and three-verb complexes and main clause two- and three-verb complexes. Because of space limitations, it is not possible to include the three-verb complexes in the discussion of VPR and extraposition (see Dubenion-Smith 2008:95–96, 107 for an analysis). However, an analysis of the two-verb complexes is sufficient with respect to these phenomena for the theoretical discussion of verbal complexes below. For each of the four complex categories, sentences exemplifying the syntags attested in the corpus are presented, followed by tables summarizing their distribution. The syntags were determined by first sorting the clauses in the database by verb form (INF, IPP, PTC, Vfin) and verb type (AUX, LEX, MOD), and then further differentiating tokens with identical verb form-type combinations until all distinct constructions had been identified.\textsuperscript{13}

\textsuperscript{13} The abbreviations used throughout the text are as follows: AUX = auxiliary verb, INF = infinitive, IPP = *infinitivus pro participio*, LEX = lexical verb, MOD = modal verb, PTC = past participle, and Vfin = finite verb. For each verb form INF, IPP, PTC, or Vfin, the subscript indicates its type: AUX, LEX, or MOD. The order in which the abbreviations appear corresponds to the descending order of the verbs. If multiple word orders are possible with a given syntagm, an example with the most frequently occurring (canonical) order is provided. See the appendix for examples involving non-canonical word orders. Some of the descriptive labels for the subordinate clause syntags have been
3.1. Subordinate Clause Two-Verb Complexes.
The corpus contains 1,326 subordinate clause two-verb complexes that fall into the syntags illustrated in 10–16. The distribution of these syntags is presented in table 1.\(^{14}\)

(10) a. was der wieder gemacht, hat,\(^1\) PTC\(_{LEX}\)-Vfin\(_{AUX}\) perfect with \(haben\)^\(^{15}\) ZW1G6
what he again done has ‘what he did again’

b. weil er krank geworden, ist, PTC\(_{LEX}\)-Vfin\(_{AUX}\) perfect with \(sein\)^\(^{16}\) ZWP65
because he ill become is ‘because he became ill’

c. wenn die Leute das gewusst, hätten, PTC\(_{LEX}\)-Vfin\(_{AUX}\) subjunctive with \(haben\)^\(^{17}\) ZWI89
if the people that known had ‘if the people had known that’

d. wenn ich ein anständiger Kerl PTC\(_{LEX}\)-Vfin\(_{AUX}\)
if I a decent guy

gewesen, wäre, subjunctive with \(sein\) ZW9J8
been were ‘if I had been a decent guy’

borrowed from Sapp 2006, while those for the main clause syntags are entirely my own.

\(^{14}\) In this and the tables below, the percentages beside the values in the columns headed by the possible word orders are row percentages, while those in the column labeled ‘total tokens’ are column percentages.

\(^{15}\) The term “perfect” refers throughout the text to both the present perfect and pluperfect.

\(^{16}\) Clauses exhibiting the so-called Zustandspassiv (\(sein\)-passive, such as \(die\ Tür ist geöffnet\) ‘the door is open’) have not been included in this study.

\(^{17}\) No instances of the \(würde\) subjunctive, such as \(dass er den Acker bestellen würde\) ‘that he would plow the field’, are attested in the corpus.
e. dass die Leute auch angestellt werden
‘that the people also get hired’

(11) dass es gerade regnen muss
‘that it has to rain right now’

(12) wenn er da ernten tut
‘when he harvests there’

(13) wer sich die Augen zubinden lässt
‘who has his eyes blindfolded’

(14) die auch wirklich noch schaffen gehen
‘those who also really still go work’

(15) dass sie ihre Arbeit gemacht kriegen
‘that they get their work done’

(16) dass ich morgens … mit dem Onkel hinaus gemusst habe
‘that I had to go out with my uncle in the morning…’

\[18\] Lassen occurs in the corpus with the meanings ‘to cause’ and ‘to allow’ and is treated in this study as an auxiliary verb in the former and a lexical verb in the latter sense.

\[19\] The syntagms with kriegen occur in the corpus with agentive, causative, and passive semantics.
<table>
<thead>
<tr>
<th>Syntagm (group)</th>
<th>2-1</th>
<th>1-2</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTC + AUX</td>
<td>887 (92.4%)</td>
<td>73 (7.6%)</td>
<td>960 (72.4%)</td>
</tr>
<tr>
<td>(perfect, subjunctive, passive)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>INF + MOD</td>
<td>227 (74.9%)</td>
<td>76 (25.1%)</td>
<td>303 (22.9%)</td>
</tr>
<tr>
<td>INF + tun</td>
<td>27 (96.4%)</td>
<td>1 (3.6%)</td>
<td>28 (2.1%)</td>
</tr>
<tr>
<td>INF + LEX</td>
<td>13 (59.1%)</td>
<td>9 (40.9%)</td>
<td>22 (1.7%)</td>
</tr>
<tr>
<td>PTC + kriegen</td>
<td>8 (100%)</td>
<td>0 (0%)</td>
<td>8 (0.6%)</td>
</tr>
<tr>
<td>INF + lassen</td>
<td>4 (100%)</td>
<td>0 (0%)</td>
<td>4 (0.3%)</td>
</tr>
<tr>
<td>modal perfect</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (0.1%)</td>
</tr>
<tr>
<td>Totals</td>
<td>1,167 (88%)</td>
<td>159 (12%)</td>
<td>1,326 (100%)</td>
</tr>
</tbody>
</table>

Table 1. Distribution of subordinate clause two-verb complexes by syntagm (group).

As shown in table 1, the overall preferred order is 2-1, which occurs far more frequently than the 1-2 order, at a rate of 88% to 12%. The group of syntagms involving a past participle and an auxiliary verb, namely PTC + AUX (perfect, subjunctive, passive), and INF + MOD occur most frequently in the corpus, accounting for 1,263 out of 1,326 total tokens (95.2%), and display both the 2-1 and 1-2 orders. INF + tun and INF + LEX, which occur far less frequently, also display both the 2-1 and 1-2 orders, though the 1-2 order is rare with the former syntagm. Finally, the syntagms with the lowest frequencies, the PTC + kriegen, INF + lassen, and the modal perfect, display only the 2-1 order. The last syntagm is very rare in the corpus, occurring in a single token. Table 2

---

20 With the syntagm INF + tun, the 1-2 order occurs only with an intervening constituent (1-X-2).

21 These are the expected rates.

22 Unless otherwise noted, the past participle and infinitive forms are lexical verbs. Note too that the modal verbs include brauchen ‘need’, dürfen ‘may’, können ‘can’, müssen ‘must’, sollen ‘should’, and wollen ‘want’.

shows the breakdown of word order rates for PTC + AUX (perfect, subjunctive, passive).

<table>
<thead>
<tr>
<th>PTC + AUX (perfect, subjunctive, passive)</th>
<th>2-1</th>
<th>1-2</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfect with haben</td>
<td>514 (98.7%)</td>
<td>7 (1.3%)</td>
<td>521 (54.3%)</td>
</tr>
<tr>
<td>perfect with sein</td>
<td>224 (77.8%)</td>
<td>64 (22.2%)</td>
<td>288 (30.0%)</td>
</tr>
<tr>
<td>werden passive</td>
<td>110 (99.1%)</td>
<td>1 (0.9%)</td>
<td>111 (11.6%)</td>
</tr>
<tr>
<td>subjunctive with haben</td>
<td>22 (100%)</td>
<td>0 (0%)</td>
<td>22 (2.3%)</td>
</tr>
<tr>
<td>subjunctive with sein</td>
<td>17 (94.4%)</td>
<td>1 (5.6%)</td>
<td>18 (1.9%)</td>
</tr>
</tbody>
</table>

Table 2. Breakdown of word order rates for PTC + AUX.

With the perfect with haben, subjunctive with sein, and werden passive, the rate of the 1-2 order is very low, occurring in only 7 out of 521 tokens (1.3%), 1 out of 18 tokens (5.6%), and 1 out of 111 tokens (0.9%), respectively. The 1-2 order is far more frequent with the perfect with sein, occurring in 64 out of 288 (22.2%) tokens. Finally, the 1-2 order is unattested with the subjunctive with haben.

3.2. Subordinate Clause Three-Verb Complexes.
The corpus contains 124 subordinate clause three-verb complexes that fall into the syntagms illustrated in 17–30. The distribution of these syntagms is presented in table 3.

(17) a. wo das Stroh gebunden ist, das Stroh gebunden ist,
     where the straw was baled, the straw was baled
     "where the straw was baled"

b. dass ich fast erschlagen worden, dass ich fast erschlagen worden,
that I almost struck.dead were been, that I almost were been
"that I had almost been struck dead"

PTC_{LEX}^{PTC_{AUX}}V_{fin_{AUX}}

ZWG56

ZWB43
(18) a. was im Winter
    what in the winter
    geflickt$_3$ muss$_1$ werden$_2$
    mended must be
    ‘that has to be mended in the winter’

b. ob das … in Betrieb
    if that … in service
    genommen$_3$ könnte$_1$ werden$_2$
    taken could be
    ‘if that could be introduced into service’

(19) a. wo ich noch Kisten
    where I still crates
    machen$_3$ habe$_1$ müssen$_2$
    make have must
    ‘where I still had to make crates’

b. wenn ich noch zehn Minuten
    if I still ten minutes
    hätte$_1$ müssen$_2$ laufen$_3$
    had must run
    ‘if I had had to run another ten minutes’

(20) wie er sie zugemauert$_3$ gehabt$_2$ hat$_1$
    how he it bricked up had has
    ‘when he had bricked it up inside’

(21) wenn wir X machen$_3$ gegangen$_2$ sind$_1$
    when we X make gone are
    ‘whenever we went to do X’

---

$^{24}$ ACI = accusativus cum infinitivo. PTC = geblieben ‘stayed’, gegangen ‘gone’, gekommen ‘come’.
(22) was man geschickt3 hat1 gekriegt2 PTC\textsubscript{LEX}-PTC\textsubscript{LEX}-Vfin\textsubscript{AUX}
what one sent has gotten PTC + kriegen perfect
‘what you have been sent’ ZWI94

(23) die wo so das erste Mal INF\textsubscript{LEX}-INF\textsubscript{LEX}-Vfin\textsubscript{MOD}
who where PART the first time
танzen3 dürfen1 gehen2 double INF + MOD\textsuperscript{25}
twice may go ZWG92
‘who are allowed to go dancing for the first time’

(24) wenn man … mit zwölf PTC\textsubscript{LEX}-INF\textsubscript{AUX}-Vfin\textsubscript{MOD}
when one … with twelve
zweimal gefahren3 will1 haben2 perfect + MOD
twice driven wants have ZWQ71
‘if you want to have ridden twice by noon’

(25) dass man alles … INF\textsubscript{LEX}-INF\textsubscript{LEX}-Vfin\textsubscript{MOD}
that one everything …
baumeln3 lassen2 kann1 INF + lassen + MOD
dangle let can ZW630
‘that you can let everything dangle’

(26) wenn das erhalten3 geblieben2 wäre1 PTC\textsubscript{LEX}-PTC\textsubscript{LEX}-Vfin\textsubscript{AUX}
if that maintained stayed were PTC + subjunctive
‘if that had been maintained’ ZWA71

(27) wie man noch keine … Flieger INF\textsubscript{LEX}-PTC\textsubscript{LEX}-Vfin\textsubscript{AUX}
how one still no … airplanes
herumfliegen3 gesehen2 hat1 ACI perfect\textsuperscript{26}
fly.around seen has ZWP62
‘when you still didn’t see any airplanes flying around’

\textsuperscript{25} \textsc{Inf}_2 = \textit{gehen} ‘go’, \textit{bleiben} ‘stay’.

\textsuperscript{26} PTC = \textit{gesehen} ‘seen’, \textit{geblieben} ‘stayed’, \textit{gegangen} ‘gone’, \textit{gekommen} ‘come’.
(28) dass ihr mich heute
that you me today
mal gehen$_3$ gelassen$_2$ habt$_1$
MOD go let have
INF + lassen perfect
ZW194
‘that you let me go today’

(29) wenn wir es ja
if we it MOD
stehen$_3$ hätten$_1$ lassen$_2$
stand would have let
INF + lassen subjunctive IPP
ZW154
‘if we had let it stand’

(30) wo ich heute wieder soll$_1$
where I today again should
noch ein Stück abgemacht$_3$ kriegen$_2$
still a bit undone get
PTC + kriegen + MOD
ZWG98
mit dem Bulldog
with the tractor
‘where I should get another parcel of land harvested again today with the tractor’

As shown in table 3, the 3-1-2 order occurs most frequently overall, in 58 out of 124 total tokens (46.8%). However, the 3-2-1 order occurs nearly as frequently, in 50 out of 124 total tokens (40.3%). Thus, there is no robust predominance of a single word order as with the subordinate clause two-verb complexes. The most frequently occurring syntagms are the perfect of passive, werden passive + MOD, and modal IPP, representing 91 out of 124 total tokens (73.4%). Instances of the perfect of passive are nearly equally divided between the 3-2-1 and 3-1-2 orders, though the 3-2-1 order is slightly more frequent. With the werden passive + MOD and modal IPP, the 3-1-2 order occurs most frequently by a wide margin. The remaining syntagms are represented by 33 out of 124 total tokens (26.6%).
<table>
<thead>
<tr>
<th>Syntagm</th>
<th>3-2-1 (51.1%)</th>
<th>3-1-2 (48.9%)</th>
<th>1-3-2 (0%)</th>
<th>1-2-3 (0%)</th>
<th>Total tokens (37.9%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfect of passive</td>
<td>24</td>
<td>23</td>
<td>0</td>
<td>0</td>
<td>47</td>
</tr>
<tr>
<td><em>werden</em> passive + MOD</td>
<td>7 (28%)</td>
<td>15 (60%)</td>
<td>3 (12%)</td>
<td>0 (0%)</td>
<td>25 (20.2%)</td>
</tr>
<tr>
<td>modal IPP</td>
<td>3 (15.8%)</td>
<td>8 (42.1%)</td>
<td>4 (21.1%)</td>
<td>4 (21.1%)</td>
<td>19 (15.3%)</td>
</tr>
<tr>
<td>complex perfect</td>
<td>6 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>6 (4.8%)</td>
</tr>
<tr>
<td>subjunctive modal IPP</td>
<td>0 (0%)</td>
<td>2 (50%)</td>
<td>0 (0%)</td>
<td>2 (50%)</td>
<td>4 (3.2%)</td>
</tr>
<tr>
<td>INF + perfect</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>4 (3.2%)</td>
</tr>
<tr>
<td>PTC + <em>kriegen</em> perfect</td>
<td>1 (33.3%)</td>
<td>2 (66.7%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>3 (2.4%)</td>
</tr>
<tr>
<td>double INF + MOD</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td>1 (33.3%)</td>
<td>0 (0%)</td>
<td>3 (2.4%)</td>
</tr>
<tr>
<td><em>werden</em> passive + subjunctive modal</td>
<td>0 (0%)</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>subjunctive passive</td>
<td>1 (50%)</td>
<td>1 (50%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>perfect + MOD</td>
<td>0 (0%)</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>INF + <em>lassen</em> MOD</td>
<td>1 (50%)</td>
<td>0 (0%)</td>
<td>1 (50%)</td>
<td>0 (0%)</td>
<td>2 (1.6%)</td>
</tr>
<tr>
<td>PTC + subjunctive</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>ACI perfect</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>INF + <em>lassen</em> perfect</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>INF + <em>lassen</em> subjunctive IPP</td>
<td>0 (0%)</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>PTC + <em>kriegen</em> MOD</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (0.8%)</td>
</tr>
<tr>
<td>Totals</td>
<td>50 (40.3%)</td>
<td>58 (46.8%)</td>
<td>10 (8.1%)</td>
<td>6 (4.8%)</td>
<td>124 (100%)</td>
</tr>
</tbody>
</table>

Table 3. Distribution of subordinate clause three-verb complexes by syntagm.
Note that the orders 2-3-1 and 2-1-3 are unattested in the corpus. This fact is unsurprising: The 2-3-1 order is relatively uncommon, limited to modal constructions in Afrikaans, Swabian, and West Flemish, while the 2-1-3 order is conspicuously absent from the inventories of three-verb complexes cross-linguistically and has thus been treated extensively in the generative literature.\(^{27}\) (See Wurmbrand 2006:239–242 and Dubenion-Smith 2008:43–62 for discussions of these word orders.)

3.3. Main Clause Two-Verb Complexes.
The corpus contains 1,114 main clause two-verb complexes that fall into the syntags illustrated in 31–48. The distribution of these syntags is presented in table 4.\(^{28}\)

(31) die sind schön gelb
those are beautifully yellow

PTC\(_{\text{LEX}}\)–PTC\(_{\text{AUX}}\)

geröstet\(_{3}\), worden\(_{2}\) perfect of passive [S]
roasted been ZWB06

‘those were roasted a beautiful yellow’

(32) die Milchkannen müssen gewaschen\(_{3}\), werden\(_{2}\)
the milkjugs must washed passive [MOD]

PTC\(_{\text{LEX}}\)–INF\(_{\text{AUX}}\)

‘the milkjugs must be washed’ ZWW19

(33) a. wir hatten müssen\(_{2}\), lachen\(_{3}\)
we had must laugh modal IPP [H]

INF\(_{\text{LEX}}\)–IPP\(_{\text{MOD}}\)

‘we couldn’t help but laugh’ ZWH62

b. wie hätte ich ihn sollen\(_{2}\), rufen\(_{3}\)
how had I him should call subjunctive modal IPP [H]

INF\(_{\text{LEX}}\)–IPP\(_{\text{MOD}}\)

‘how should I have called him’ ZWI13

---

\(^{27}\) However, Heilmann (1999) does include 2-1-3 as a grammatical order in one variety of Swabian. In addition, Schmid & Vogel (2004) claim that the 2-1-3 order is possible under very specific stress conditions. Louden (1990, 2007, forthcoming) accounts for the 2-1-3 order in Pennsylvania Dutch by linking it to extraposition.

\(^{28}\) In the labels for the main clause two-verb syntags, the type of inflected verb located in C is given in brackets, where H = haben, MOD = modal verb, S = sein, and W = werden.
(34) da `bin ich auch arbeiten' gegangen
there am I also work gone
'I also went to work there'

(35) das muss interessant gewesen
that must interesting been be
'that must have been interesting'

(36) da `mussten wir Geld verdienen' gehen
there had.to we money earn go
then we had to go earn money'

(37) warum hast du den Bub schlagen lassen
why have you the boy hit let
'why did you let the boy get hit'

(38) ich habe noch gar nicht gesät gehabt
I have still at.all not sown had
'I still haven’t sown [any seeds]'

(39) wir wollten es nicht auswachsen lassen
we wanted it not grow.out let
'we didn’t want to let it reach full size'

(40) der hatte so einen Sitz gemacht gekriegt
he had such a seat made gotten
'he got a seat like that made for him'

(41) da `habe ich ihn ausstopfen' gelassen
then have I it stuffed let
'then I had it stuffed'

---


30 This syntagm includes constructions with both haben and sein.
(42) ich habe dich einmal
I have you once

einen anderen Walzer sehen_3 tanzen_3
one other waltz see dance

‘I saw you dancing another waltz once’

(43) mit der Maschine wird man
with the machine will one

nicht viel machen_3 können_2
not much make can

‘we won’t be able to do much with that machine’

(44) es hat sich wieder gelohnt_3 gemacht_2
it has REFL again worthy made

‘it was worth it again’

(45) dann wurde laufen_3 gegangen_2
then was run gone

‘then we went running’

(46) ich selbst habe noch gemusst_2 mähen_3
I myself have still had to reap

‘as for me, I still had to reap’

(47) bei der Dunkelheit kann
at the darkness can

allerhand getrieben_3 kommen_2
all sorts floating come

‘in darkness all sorts of things can come floating along’

31 IPP = helfen ‘help’, sehen ‘see’, lernen ‘learn’.
muss man ja auch mal
must one MOD also MOD
ausruhen$_3$ können$_2$
rest can
‘you have to be able to rest, you know’

<table>
<thead>
<tr>
<th>Syntagm</th>
<th>3-2</th>
<th>2-3</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfect of passive [S]</td>
<td>476</td>
<td>0</td>
<td>476 (42.7%)</td>
</tr>
<tr>
<td>passive [MOD]</td>
<td>224</td>
<td>0</td>
<td>224 (20.1%)</td>
</tr>
<tr>
<td>modal IPP [H]</td>
<td>75</td>
<td>103</td>
<td>178 (16%)</td>
</tr>
<tr>
<td>INF + perfect [H/S]</td>
<td>31</td>
<td>7</td>
<td>38 (3.4%)</td>
</tr>
<tr>
<td>perfect [MOD]</td>
<td>34</td>
<td>0</td>
<td>34 (3.0%)</td>
</tr>
<tr>
<td>double INF [MOD]</td>
<td>21</td>
<td>7</td>
<td>28 (2.5%)</td>
</tr>
<tr>
<td>INF + lassen IPP [H]</td>
<td>17</td>
<td>8</td>
<td>25 (2.2%)</td>
</tr>
<tr>
<td>complex perfect [H]</td>
<td>22</td>
<td>0</td>
<td>22 (2.0%)</td>
</tr>
<tr>
<td>INF + lassen [MOD]</td>
<td>14</td>
<td>2</td>
<td>16 (1.4%)</td>
</tr>
<tr>
<td>PTC + kriegen [H]</td>
<td>14</td>
<td>0</td>
<td>14 (1.3%)</td>
</tr>
<tr>
<td>INF + lassen perfect [H]</td>
<td>13</td>
<td>0</td>
<td>13 (1.2%)</td>
</tr>
<tr>
<td>subjunctive modal IPP [H]</td>
<td>4</td>
<td>8</td>
<td>12 (1.1%)</td>
</tr>
<tr>
<td>ACI IPP [H]</td>
<td>4</td>
<td>7</td>
<td>11 (1.0%)</td>
</tr>
<tr>
<td>modal future [W]</td>
<td>7</td>
<td>1</td>
<td>8 (0.7%)</td>
</tr>
<tr>
<td>PTC + perfect [H/S]</td>
<td>5</td>
<td>0</td>
<td>5 (0.4%)</td>
</tr>
<tr>
<td>INF + passive [W]</td>
<td>4</td>
<td>0</td>
<td>4 (0.4%)</td>
</tr>
<tr>
<td>INF + modal perfect [H]</td>
<td>0</td>
<td>2</td>
<td>2 (0.2%)</td>
</tr>
<tr>
<td>PTC + INF [MOD]</td>
<td>2</td>
<td>0</td>
<td>2 (0.2%)</td>
</tr>
<tr>
<td>INF + modal INF [MOD]</td>
<td>2</td>
<td>0</td>
<td>2 (0.2%)</td>
</tr>
<tr>
<td>Totals</td>
<td>969</td>
<td>145</td>
<td>1,114 (100%)</td>
</tr>
</tbody>
</table>

Table 4. Distribution of main clause two-verb complexes by syntagm.

As table 4 shows, the overall preferred order is 3-2, which occurs far more frequently than the 2-3 order, at a rate of 87% to 13%. Note that this distribution is very similar to that for the subordinate clause two-verb complexes (88% 2-1 to 12% 1-2). The syntagms perfect of passive
[S], passive [MOD], and modal IPP [H] occur most frequently in the corpus, accounting for 878 out of 1,114 total tokens (78.8%). Strikingly, the 2-3 order does not occur at all with the perfect of passive [S] and passive [MOD], both involving a past participle and a form of werden, while this is by far the preferred order with the modal IPP [H], at a rate of 57.9%. In section 4.1, we see that this distribution can be attributed in part to the combination of verb forms and types that comprise these syntagms. Among the remaining syntagms, both the 3-2 and 2-3 orders occur with seven, but the 2-3 order is preferred with only three: the subjunctive modal IPP [H], ACI IPP [H], and INF + modal perfect [H].

3.4. Main Clause Three-Verb Complexes.
The corpus contains 14 main clause three-verb complexes that fall into the syntagms illustrated in 49–55. The distribution of these syntagms is presented in table 5.

(49) das hat alles …
that has everything …
gespritzt, werden, müssen,
Sprayed be must
‘all that had to be sprayed …’

PTC\textsubscript{LEX-INF\textsubscript{AUX-IPP\textsubscript{MOD}}}

(50) die hatten wir müssen,
those had we must
pflastern, lassen,
pave let
‘we had to have those paved’

INF\textsubscript{LEX-LEX/AUX-IPP\textsubscript{MOD}}

(51) haben sie sollen, säen, kommen,
have they should sow come
double INF + MOD IPP [H]
‘they were supposed to come sow [seeds]’

INF\textsubscript{LEX-LEX-IPP\textsubscript{MOD}}

ZW0P4, ZWP64, ZW2K7
(52) das hätte der Fährmann  
that had the ferryman  

sollen₂ gewusst₄ haben₃  
should known have  

perfect + subjunctive MOD IPP [H]  

ZWP83

‘the ferryman should have known that’

(53) da sind wir vom Amerikaner  
there are we by the American  

gefangen₄ genommen₃ worden₂  
captured taken been  

PTC + perfect of passive [S]  

ZWA81

‘then we were taken capture by the American’

(54) ich soll es geschafft₄ gehabt₃ haben₂  
I should it managed had have  

complex perfect [MOD]  

ZWG44

‘I am supposed to have managed it’

(55) dann ist sie ein paar Tage  
then is it a few days  

liegen₄ gelassen₃ worden₂  
lie let been  

INF + lassen perfect of passive [S]  

ZWG09

‘then it was left to lie for a few days’

Compared to the other complex categories, the main clause three-verb complexes are relatively scarce, as table 5 shows. Nonetheless, 4 of the 6 possible word orders are attested in the corpus: 4-3-2, 4-2-3, 3-2-4, and 2-4-3. The 4-3-2 and 2-4-3 orders occur most frequently, each in 6 out of 14 total tokens (42.9%). Syntagms involving an IPP construction account for the majority of tokens, 11 out of 14 (78.6%). As with the subordinate clause three-verb complexes, there is no robust predominance of a single word order.
<table>
<thead>
<tr>
<th>Syntagm</th>
<th>4-3-2</th>
<th>4-2-3</th>
<th>3-2-4</th>
<th>2-4-3</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>werden passive + MOD IPP [H]</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>INF + lassen + MOD IPP [H]</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>double INF + MOD IPP [H]</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>perfect + subjunctive MOD IPP [H]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PTC + perfect of passive [S]</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>complex perfect [MOD]</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>INF + lassen perfect of passive [S]</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 5. Distribution of main clause three-verb complexes by syntagm.

3.5. Verb Projection Raising and Two-Verb Complexes.
This section deals with verb projection raising in subordinate clause two-verb complexes and main clause two-verb complexes. Recall from the previous section that 159 out of 1,326 subordinate clause two-verb complexes display the 1-2 order. Of these, 20 (12.6%) involve VPR with a constituent appearing between the members of the verb cluster. The intervening constituents fall into four types, illustrated in 56–59. Their distribution is presented in table 6.

(56) dass man so früh musste_1 **auf**stehen_2  separable prefix
  that one so early had to **stand** up
  ‘that you had to get up so early’
(57) die so früh wollten$_1$ **fertig** sein$_2$ adjective/adverb
who so early wanted finished be
‘who wanted to be finished so early’

(58) dass er jeden Augenblick musste$_1$
that he any moment had.to
**hinter eine Hecke laufen$_2$** PP
behind a hedge run
‘that he had to run behind a hedge at any moment’

(59) dass sie da müssen$_1$ **einen ordentlichen Korb**
that they there must a decent basket
**voll Essen kochen$_2$** indefinite NP
full food cook
‘that they have to cook a decent basketful of food’

<table>
<thead>
<tr>
<th>Intervening constituent</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>separable prefix</td>
<td>11 (55%)</td>
</tr>
<tr>
<td>adjective/adverb</td>
<td>4 (20%)</td>
</tr>
<tr>
<td>PP</td>
<td>3 (15%)</td>
</tr>
<tr>
<td>indefinite NP</td>
<td>2 (10%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>20 (100%)</td>
</tr>
</tbody>
</table>

Table 6. Distribution of VPR intervening constituents, subordinate clause two-verb complexes.

Table 6 shows that the majority of intervening constituents are separable prefixes. Note that unlike standard German (but like languages such as standard Dutch), WCG readily allows separable prefixes to be stranded. This is exemplified in the following example in which the particle **an** ‘on’ does not appear raised with **schneiden** ‘cut’.\(^\text{32}\)

\[^{32}\text{For further information on VPR constructions in a variety of Germanic languages, see Wurmbrand 2006:273–284. See Dubenion-Smith 2008:89–92, 102–105 for a discussion of separable prefix stranding in WCG.}\]
(60) wenn man das Thema da mal an will, schneiden, when one the topic there MOD on wants cut
‘if you want to broach the topic’

Let us now examine verb projection raising in main clause two-verb complexes. Compared to the subordinate clause two-verb complexes, the main clause two-verb complexes display VPR at a much higher rate. Of the 145 clauses with the 2-3 order (13%), 75 (51.7%) involve an intervening constituent. The attested VPR constructions fall into five types with a single intervening constituent and an additional 13 types with two or more intervening constituents. Table 7 shows the distribution of constructions with one intervening constituent, illustrated in 61–65.33

(61) ich habe schon zwei Mühlen
I have already two mills
geholfen, einrichten, separable prefix
helped set up
‘I have already helped set up two mills’

(62) da habe ich müssen mit ihm fahren, PP
there have I must with him go
‘then I had to go with him’

(63) da haben wir müssen schnell springen, adverb/adjective
there have we must quickly jump
‘then we had to run quickly’

(64) ich hatte müssen Schneider lernen, bare N
I had must tailor learn
‘I had to become a tailor’

(65) haben müssen, die Polizei alarmieren, definite NP
have must the police alarm
‘(we) had to notify the police’

33 See Dubenion-Smith 2008:103–104 for examples involving multiple intervening constituents.
Again, we see that the majority of intervening constituents are separable prefixes. A comparison of tables 6 and 7 reveals that for both the subordinate and main clause complexes, the three most common types of intervening constituent are the same (separable prefix, PP, adjective/adverb), followed by at least one nominal category (indefinite NP, bare N, definite NP).

<table>
<thead>
<tr>
<th>Intervening constituent</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>separable prefix</td>
<td>26 (34.7%)</td>
</tr>
<tr>
<td>PP</td>
<td>11 (14.7%)</td>
</tr>
<tr>
<td>adjective/adverb</td>
<td>9 (12%)</td>
</tr>
<tr>
<td>bare N</td>
<td>8 (10.7%)</td>
</tr>
<tr>
<td>definite NP</td>
<td>4 (5.3%)</td>
</tr>
<tr>
<td>multiple constituents</td>
<td>17 (22.7%)</td>
</tr>
<tr>
<td>Total</td>
<td>75 (100%)</td>
</tr>
</tbody>
</table>

Table 7. Distribution of VPR intervening constituents, main clause two-verb complexes.

3.6. Extrapolation and Two-Verb Complexes.

In this section, I discuss extrapolation in subordinate clause two-verb complexes and main clause two-verb complexes. As noted in section 1, verbal complexes may occur not only with elements appearing inside the cluster, but also with extraposed elements to the right of the cluster. The distribution of extrapolation in the subordinate clause two-verb complexes is presented in table 8.\(^\text{34}\)

\(^{34}\) Duden (2005:1223) distinguishes extrapolation from another type of expansion, Nachtrag (afterthought):

\begin{quote}
wie 'weit is des entfernt ↑* von port 'dixon ↑
how far is that away from Port Dixon

‘how far away is that, from Port Dixon?’
\end{quote}

The upward arrows indicate rising intonation, and the asterisk the division between two intonational phrases consisting of *wie weit is des entfernt* ‘how far away is that’ and *von port dixon* ‘from Port Dixon’.
<table>
<thead>
<tr>
<th>Extraposition</th>
<th>2-1</th>
<th>1-2</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>extraposed constituent</td>
<td>50 (90.9%)</td>
<td>5 (9.1%)</td>
<td>55 (4.1%)</td>
</tr>
<tr>
<td>nothing extraposed</td>
<td>1,117 (87.9%)</td>
<td>154 (12.1%)</td>
<td>1,271 (95.9%)</td>
</tr>
<tr>
<td>Total</td>
<td>1,167 (88%)</td>
<td>159 (12%)</td>
<td>1,326 (100%)</td>
</tr>
</tbody>
</table>

Table 8. Distribution of extraposition, subordinate clause two-verb complexes.

As table 8 shows, extraposition occurs in 55 out of 1,326 total tokens (4.1%). 50 out of 55 instances (90.9%) occur with the 2-1 order, while only 5 out of 55 (9.1%) occur with the 1-2 order. Thus the 1-2 order occurs less frequently than the expected rate (12%) in constructions that display extraposition.

Table 9 shows the distribution of extraposition by constituent type, illustrated in 66–72.\(^{35}\)

(66) wenn wir das Geld verdient\(_2\) haben\(_1\) in Amerika adjunct PP
when we the money earned have in America ZWP61

‘when we have earned the money in America’

Extraposition and afterthought can be differentiated according to whether the expansion is integrated into the intonational phrase of the preceding structure or whether it forms its own intonational phrase, as does von port dixon in the sentence above. Schwitalla (2003:117, note 12) notes that this differentiation can be difficult since prosodic integration is graded according to the number of prosodic factors present, such as a break or a stressed syllable in the expansion. However, the way of distinguishing the two types according to this criterion proved sufficient in the examination of recordings for the corpus study. See also Altmann 1981, Auer 1991, and Starke 1965 for discussion of postfield expansions.

\(^{35}\) I follow here Engel’s (1994) distinction between argument and adjunct PPs.
(67) was man schwarz konnte
what you black could

tauschen auf Lebensmittel argument PP
exchange for food ZW1G9

‘what you could exchange for food on the black market’

(68) wie sie gesagt haben früher
as they said have earlier ZWA97

‘as they used to say’

(69) wenn man Mist geladen hatte drei Tage
when one manure loaded had three days ZW1H8

‘when you had loaded manure for three days’

(70) die Gurken angebaut haben
who cucumbers cultivated have

hier in unserem Ort locative adverb +
here in our town adjunct PP

‘who have cultivated cucumbers here in our town’ ZW605

(71) wie man das Brot gebacken hat davon
as one the bread baked has from it ZWR43

‘when you baked the bread with it’

(72) ob er nun bang machen wollte den Heinrich object NP
if he MOD scared make wanted the Heinrich ZWQ44

‘if he wanted to scare Heinrich’

---

36 The corpus includes temporal, modal, and locative adverbs.
<table>
<thead>
<tr>
<th>Extrapoed constituent</th>
<th>2-1</th>
<th>1-2</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>adjunct PP</td>
<td>27 (90%)</td>
<td>3 (10%)</td>
<td>30 (54.5%)</td>
</tr>
<tr>
<td>argument PP</td>
<td>7 (87.5%)</td>
<td>1 (12.5%)</td>
<td>8 (14.5%)</td>
</tr>
<tr>
<td>adverb</td>
<td>7 (100%)</td>
<td>0 (0%)</td>
<td>7 (12.7%)</td>
</tr>
<tr>
<td>temporal NP</td>
<td>3 (75%)</td>
<td>1 (25%)</td>
<td>4 (7.3%)</td>
</tr>
<tr>
<td>locative adverb + adjunct PP</td>
<td>4 (100%)</td>
<td>0 (0%)</td>
<td>4 (7.3%)</td>
</tr>
<tr>
<td>da-compound</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (1.8%)</td>
</tr>
<tr>
<td>object NP</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (1.8%)</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>50 (90.9%)</strong></td>
<td><strong>5 (9.1%)</strong></td>
<td><strong>55 (100%)</strong></td>
</tr>
</tbody>
</table>

Table 9. Distribution of extrapoed constituents, subordinate clause two-verb complexes.

The majority of extrapoed constituents are adjunct PPs, accounting for 30 out of 55 tokens (54.5%). The adjunct PPs, in combination with the other adverbials and da-compound, comprise 47 out of 55 total tokens (85.5%), while argument PPs make up the remaining 8 out of 55 tokens (14.5%). The occurrence of only one instance of an extrapoed object NP falls in line with the increased unacceptability of object NP extraposition in the ENHG period to its general ungrammaticality in modern German (see Bies 1996).

Like the subordinate clause two-verb complexes, the main clause two-verb complexes display extraposition. The distribution of extraposition is presented in table 10.

<table>
<thead>
<tr>
<th>Extraposition</th>
<th>3-2</th>
<th>2-3</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>extrapoed constituent</td>
<td>96 (88.9%)</td>
<td>12 (11.1%)</td>
<td>108 (9.7%)</td>
</tr>
<tr>
<td>nothing extrapoed</td>
<td>873 (86.8%)</td>
<td>133 (13.2%)</td>
<td>1,006 (90.3%)</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>969 (87%)</strong></td>
<td><strong>145 (13%)</strong></td>
<td><strong>1,114 (100%)</strong></td>
</tr>
</tbody>
</table>

Table 10. Distribution of extraposition, main clause two-verb complexes.

As table 10 shows, extraposition occurs in 108 out of 1,114 total tokens (9.7%) at a rate significantly higher than in the subordinate clause two-verb complexes (4.1%). Ninety-six out of 108 instances (88.9%) occur with the 3-2 order, while only 12 out of 108 (11.1%) occur with the 2-3
order. Thus the 2-3 order occurs less frequently than the expected rate (13%) in constructions that display extraposition, as we saw above for the 1-2 order in the subordinate clause two-verb complexes.

Table 11 below shows the distribution of extraposition by constituent type, illustrated in 73–80.

(73) dä bin ich zweimal verwundet3 worden2 im Krieg adjunct PP
     there am I twice wounded been in.the war ZWA81
     ‘I was wounded twice in the war’

(74) ist erzählt3 worden2 von der Schule argument PP
     has told been about the school ZWW51
     ‘was told about the school’

(75) schwere Arbeit haben wir müssen2 machen3 dä adverb37
     difficult work have we must make there ZWW69
     ‘we had to do difficult work there’

(76) dä ist die Arbeit nicht gemacht3 worden2 darin dä-compound
     there is the work not done been in.the there ZWG05
     ‘the work wasn’t done in that area’

(77) die sind immer so geschnitten3 worden2
     they are always thus cut been
     in dem Wingert dä adjunct PP + adverb
     in the vineyard there ZWG50
     ‘they were always cut that way there in the vineyard’

(78) sind dann später
     are then later
     verkauft3 worden2 als Wurzelreben appositive
     sold been as rootvines ZWG50
     ‘(they) were sold later as rootvines’

37 The corpus includes temporal and locative adverbs.
(79) dann muss gemäß, werden jeden Morgen temporal NP
then must reaped be every morning ZWX10
‘then you have to reap every morning’

(80) da ist er geschmückt, worden mit preposition
there is he decorated been with ZWB39
‘he was decorated with it’

<table>
<thead>
<tr>
<th>Extrapoised constituent</th>
<th>3-2</th>
<th>2-3</th>
<th>Total tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>adjunct PP</td>
<td>55 (90.2%)</td>
<td>6 (9.8%)</td>
<td>61 (56.5%)</td>
</tr>
<tr>
<td>argument PP</td>
<td>19 (86.4%)</td>
<td>3 (13.6%)</td>
<td>22 (20.4%)</td>
</tr>
<tr>
<td>adverb</td>
<td>8 (72.7%)</td>
<td>3 (27.3%)</td>
<td>11 (10.2%)</td>
</tr>
<tr>
<td>da-compound</td>
<td>6 (100%)</td>
<td>0 (0%)</td>
<td>6 (5.6%)</td>
</tr>
<tr>
<td>adjunct PP + adverb</td>
<td>4 (100%)</td>
<td>0 (0%)</td>
<td>4 (3.7%)</td>
</tr>
<tr>
<td>appositive</td>
<td>2 (100%)</td>
<td>0 (0%)</td>
<td>2 (1.9%)</td>
</tr>
<tr>
<td>temporal NP</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (0.9%)</td>
</tr>
<tr>
<td>preposition</td>
<td>1 (100%)</td>
<td>0 (0%)</td>
<td>1 (0.9%)</td>
</tr>
<tr>
<td>Totals</td>
<td>96 (88.9%)</td>
<td>12 (11.1%)</td>
<td>108 (100%)</td>
</tr>
</tbody>
</table>

Table 11. Distribution of extrapoised constituents, main clause two-verb complexes.

As we saw above with the subordinate clause two-verb complexes, adjunct PPs account for the majority of extrapoised constituents in the main clause two-verb complexes as well, comprising 61 out of 108 total tokens (56.5%). These plus the adverbials, da-compound, appositives, and preposition make up 86 out of 108 total tokens (79.6%). Argument PPs account for the remaining 22 out of 108 tokens (20.4%). Extrapoised object NPs do not occur in the main clause two-verb complexes.

3.7. Summary of Main Findings.
The syntags and word orders attested in the Zwirner Corpus data were introduced in section 3.1–3.4. For the subordinate clause two- and three-verb complexes and the main clause two-verb complexes, these are presented in the summary tables 12–14 below.\footnote{I exclude the main clause three-verb complexes here because of the extremely low token count. Only the principal syntags are listed for the subordinate
### Table 12. Word orders in WCG, subordinate clause two-verb complexes.

<table>
<thead>
<tr>
<th>Syntagm</th>
<th>Word order</th>
</tr>
</thead>
<tbody>
<tr>
<td>INF + LEX</td>
<td>1-2 (40.9%)</td>
</tr>
<tr>
<td></td>
<td>2-1 (59.1%)</td>
</tr>
<tr>
<td>INF + MOD</td>
<td>1-2 (25.1%)</td>
</tr>
<tr>
<td></td>
<td>2-1 (74.9%)</td>
</tr>
<tr>
<td>perfect with <strong>sein</strong></td>
<td>1-2 (22.2%)</td>
</tr>
<tr>
<td></td>
<td>2-1 (77.8%)</td>
</tr>
<tr>
<td>subjunctive with <strong>sein</strong></td>
<td>2-1 (94.4%)</td>
</tr>
<tr>
<td></td>
<td>1-2 (5.6%)</td>
</tr>
<tr>
<td>INF + <strong>tun</strong></td>
<td>2-1 (96.4%)</td>
</tr>
<tr>
<td></td>
<td>1-2 (3.6%)</td>
</tr>
<tr>
<td>perfect with <strong>haben</strong></td>
<td>2-1 (98.7%)</td>
</tr>
<tr>
<td></td>
<td>1-2 (1.3%)</td>
</tr>
<tr>
<td><strong>werden</strong> passive</td>
<td>2-1 (99.1%)</td>
</tr>
<tr>
<td></td>
<td>1-2 (0.9%)</td>
</tr>
<tr>
<td>subjunctive with <strong>haben</strong></td>
<td>2-1 (100%)</td>
</tr>
<tr>
<td>**PTC + <strong>kriegen</strong></td>
<td>2-1 (100%)</td>
</tr>
<tr>
<td>**INF + <strong>lassen</strong></td>
<td>2-1 (100%)</td>
</tr>
<tr>
<td><strong>modal perfect</strong></td>
<td>2-1 (100%)</td>
</tr>
</tbody>
</table>

### Table 13. Word orders in WCG, subordinate clause three-verb complexes.

<table>
<thead>
<tr>
<th>Syntagm</th>
<th>Word order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>modal IPP</strong></td>
<td>1-2-3 (21.1%)</td>
</tr>
<tr>
<td></td>
<td>1-3-2 (21.1%)</td>
</tr>
<tr>
<td></td>
<td>3-1-2 (42.1%)</td>
</tr>
<tr>
<td></td>
<td>3-2-1 (15.8%)</td>
</tr>
<tr>
<td><strong>werden</strong> passive + <strong>MOD</strong></td>
<td>3-1-2 (60%)</td>
</tr>
<tr>
<td></td>
<td>1-3-2 (12%)</td>
</tr>
<tr>
<td></td>
<td>3-2-1 (28%)</td>
</tr>
<tr>
<td><strong>perfect of passive</strong></td>
<td>3-2-1 (51.1%)</td>
</tr>
<tr>
<td></td>
<td>3-1-2 (48.9%)</td>
</tr>
</tbody>
</table>

clause three-verb and main clause two-verb complexes. For each syntagm, the word order(s) that occur(s) more frequently than the expected rate is (are) listed first.
<table>
<thead>
<tr>
<th>Syntagm</th>
<th>Word order</th>
</tr>
</thead>
<tbody>
<tr>
<td>modal IPP [H]</td>
<td>2-3 (57.9%)</td>
</tr>
<tr>
<td></td>
<td>3-2 (42.1%)</td>
</tr>
<tr>
<td>perfect of passive [S]</td>
<td>3-2 (100%)</td>
</tr>
<tr>
<td>passive [MOD]</td>
<td>3-2 (100%)</td>
</tr>
</tbody>
</table>

Table 14. Word orders in WCG, main clause two-verb complexes.

In the subordinate clause two-verb complexes, both the 2-1 and 1-2 orders occur, at a rate of 88% and 12%, respectively. The syntagms INF + LEX, INF + MOD, and perfect with sein display the non-canonical 1-2 order at rates higher than the expected rate. Four word orders are attested in the subordinate clause three-verb complexes, 3-2-1, 3-1-2, 1-3-2, and 1-2-3, occurring at respective rates of 40.3%, 46.8%, 8.1%, and 4.8%. The modal IPP displays the non-canonical word orders 1-2-3 and 1-3-2, the werden passive + MOD the order 1-3-2, and the perfect of passive the order 3-2-1 more frequently than the expected rates. The main clause two-verb complexes display both the 3-2 and 2-3 orders, at rates nearly identical to those in the subordinate clause two-verb complexes: 87% and 13%. The syntagm modal IPP [H] displays the non-canonical order 2-3 at a rate far higher than the expected rate. Finally, the four word orders 4-3-2, 4-2-3, 3-2-4, and 2-4-3 are attested in the main clause three-verb complexes, at respective rates of 42.9%, 7.1%, 7.1%, and 42.9%.

Sections 3.5 and 3.6 presented the distribution of elements in VPR and extraposition constructions, limited to the two-verb complexes. The analysis of VPR showed that the rate of occurrence is far higher in the main clause than in the subordinate clause complexes (51.7% vs. 12.6%) and that separable prefixes comprise the largest group of intervening constituents. Similarly, the rate of extraposition is higher in the main clause than in the subordinate clause complexes (9.7% to 4.1%). For both complex categories, the rate of the non-canonical word order is lower than the expected rate when an element is extraposed, and adjunct PPs comprise the largest group of extraposed constituents.

4. Results of the Statistical Analysis of the Zwirner Corpus Data.
This section presents the results of a statistical analysis of the Zwirner Corpus data. I take as the point of departure Sapp 2006, a study of
subordinate clause verbal complexes from ENHG to modern German. Based on a corpus of ENHG texts and experimental data, Sapp aims to determine the effect of various factors on word order, many of which have been shown in previous studies to influence serialization in the verbal complex (see, for example, Ebert 1981). The linguistic and rhythmic factors Sapp tests are class or phonological weight of the word/phrase preceding the verb cluster, VP coordination, clause type, syntagm, verbal prefix type, and extraposition. The non-linguistic factors are time period, dialect, occupation, and genre. Finally, Sapp investigates the influence of the pragmatic factor focus (see Sapp 2006:6–8).

Sapp (2006) determines that syntagm, verbal prefix type, extraposition, time period, dialect, occupation, and genre have statistically significant effects on word order in the two-verb complexes, while in the three-verb complexes only syntagm and verbal prefix type show clear effects. Sapp (2006, 2007) also demonstrates that in both the two- and three-verb complexes, variation in word order can be linked to focus.

The results of the empirical analysis presented in section 3 suggested that word order variation in the verbal complex may be linked to several factors, namely syntagm, verbal prefix type, and extraposition—factors shown in Sapp’s (2006) study to have statistically significant effects on word order. While it is possible that the factors that have no statistically significant effect in Sapp’s study might have such an effect in the Zwirner Corpus data, here I limit the statistical analysis to syntagm, verbal prefix type, extraposition, and one additional factor linked to word order in Sapp’s (2006, 2007) work: object focus. As we see below, both syntagm and verbal prefix type have a statistically significant effect on word order in the subordinate clause two-verb complexes. In the main clause two-verb complexes, syntagm, verbal prefix type, and a grammatical correlate to focus have statistically significant effects. Because of statistical insignificance, the results for the subordinate clause three-verb complexes are not discussed.

---

39 Sapp also codes his data for sex and education but cannot carry out an analysis according to these variables because of a lack of information.

The Zwirner Corpus data were analyzed with the statistics package GoldVarb X.\footnote{For information on GoldVarb, see Paolillo 2002, Robinson et al. 2001, Tagliamonte 2006, and Young & Bayley 1996.} GoldVarb is a multivariate analysis application originally developed for use in sociolinguistics that allows the researcher to determine the effect of various independent variables or FACTORS, organized into FACTOR GROUPS, on the occurrence of a single dependent variable. For each factor, the program outputs a PROBABILISTIC WEIGHT or FACTOR WEIGHT, a value ranging from 0 to 1.0. A value below 0.5 indicates that the factor has a disfavoring effect on the occurrence of the dependent variable; a value above 0.5—a favoring effect. Values very close to 0.5 indicate that a factor has neither a disfavoring nor a favoring effect on the occurrence of the dependent variable. Within a particular factor group, only factors for which all possible word orders are attested may be included in the statistical analysis as an independent factor, otherwise GoldVarb will not run the analysis. These so-called *KNOCKOUT* factors must either be combined with other factors in logical groupings or be eliminated from the analysis altogether. In the present study, the dependent variable is word order, and examples of the factor groups tested are syntagm and focus.

Recall from section 3.1 that in the subordinate clause two-verb complexes, the 1-2 order occurs at an overall rate of 12\%, the expected rate. Generally speaking, factors for which the rate of 1-2 is higher than the expected rate will be favoring factors, while factors for which the rate of 1-2 is lower than the expected rate will be disfavoring factors. This principle holds for the main clause two-verb complexes as well. Note, however, that because of the mechanics of the program, the rates of occurrence and factor weights do not always correspond exactly. In the tables below, factor weights are presented from highest to lowest. Following common practice, the results of the tests to follow will be considered statistically significant when $p \leq 0.05$.

4.1. Syntagm and Two-Verb Complexes.

In this section, I consider the effect of syntagm on word order in subordinate and main clause verb complexes. I begin with subordinate clause two-verb complexes. In the corpus, syntagms with an infinitive display the 1-2 order at a far higher rate than those with a past participle.
Table 15 shows that the former favor, and the latter disfavor, the 1-2 order.

<table>
<thead>
<tr>
<th>Syntagm</th>
<th>2-1</th>
<th>1-2</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>INF + LEX, MOD, <em>tun, lassen</em></td>
<td>271 (75.9%)</td>
<td>86 (24.1%)</td>
<td>0.713</td>
</tr>
<tr>
<td>PTC + AUX (perfect, passive, subjunctive)</td>
<td>896 (92.5%)</td>
<td>73 (7.5%)</td>
<td>0.417</td>
</tr>
<tr>
<td>Total</td>
<td>1167 (88%)</td>
<td>159 (12%)</td>
<td></td>
</tr>
<tr>
<td><em>p &lt; 0.001</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 15. Effect of general syntagm on word order, subordinate clause complexes.

Table 16 presents the factor weights for these syntagms broken down into specific types.  

<table>
<thead>
<tr>
<th>Syntagm</th>
<th>2-1</th>
<th>1-2</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>INF + LEX</td>
<td>13 (59.1%)</td>
<td>9 (40.9%)</td>
<td>0.921</td>
</tr>
<tr>
<td>INF + MOD</td>
<td>227 (74.9%)</td>
<td>76 (25.1%)</td>
<td>0.849</td>
</tr>
<tr>
<td>perfect with <em>sein</em></td>
<td>224 (77.8%)</td>
<td>64 (22.2%)</td>
<td>0.828</td>
</tr>
<tr>
<td>INF + other (<em>tun, lassen</em>)</td>
<td>31 (96.9%)</td>
<td>1 (3.1%)</td>
<td>0.352</td>
</tr>
<tr>
<td>subjunctive (<em>haben, sein</em>)</td>
<td>39 (97.5%)</td>
<td>1 (2.5%)</td>
<td>0.301</td>
</tr>
<tr>
<td>perfect with <em>haben</em></td>
<td>514 (98.7%)</td>
<td>7 (1.3%)</td>
<td>0.186</td>
</tr>
<tr>
<td><em>werden</em> passive</td>
<td>110 (99.1%)</td>
<td>1 (0.9%)</td>
<td>0.133</td>
</tr>
<tr>
<td>Total</td>
<td>1158 (87.9%)</td>
<td>159 (12.1%)</td>
<td></td>
</tr>
<tr>
<td><em>p &lt; 0.001</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 16. Effect of specific syntagm on word order, subordinate clause complexes.

42 INF + *lassen*, the subjunctive with *haben*, PTC + *kriegen*, and the modal perfect only occur with the 2-1 order and are therefore *knockout* factors. In table 16, INF + *lassen* and the subjunctive with *haben* have thus been collapsed with INF + *tun* and the subjunctive with *sein*, respectively, while PTC + *kriegen* and the modal perfect have been excluded.
As the table illustrates, the syntagms INF + LEX, INF + MOD, and perfect with sein favor the 1-2 order, while this order is disfavored by INF + other (tun, lassen), the subjunctive (haben, sein), the perfect with haben, and the werden passive. The syntagms can be placed on a hierarchy according to factor weight: INF + LEX > INF + MOD > perfect with sein > perfect with haben > werden passive. This hierarchy is similar to those found in previous studies of ENHG. Compare the hierarchies in 81a–d.43

(81) a. INF + LEX > INF + MOD > perfect with sein
   > perfect with haben > werden passive (Zwirner Corpus)

   b. INF + MOD > perfect with sein > perfect with haben
   > werden passive (Bies 1996:58)

   c. perfect with sein > INF + MOD > perfect with haben
   > werden passive (Ebert 1992:5)

   d. INF + LEX, AUX > INF + MOD > perfect with sein
   > perfect with haben > werden passive (Sapp 2006:56)

The four hierarchies are nearly identical, with the exception of the relative ordering of the constructions perfect with sein and INF + MOD on Ebert’s hierarchy. Despite this deviation, the order perfect with sein > perfect with haben > werden passive holds in each case. Thus, it may be concluded that the relationship between syntagm and word order in the subordinate clause two-verb complexes has not changed from Early New High German to modern German, as far as the WCG dialects are concerned.

Let us now consider main clause two-verb complexes. Since 11 of the main clause two-verb syntagms, including two of the three principal syntagms, are unattested with either the 3-2 or 2-3 order and are therefore *KNOCKOUT* factors, it was not possible to conduct a statistical analysis with each separate syntagm. However, it was possible

---

43 For purposes of comparison, the sein passive and future constructions have been excluded in the hierarchies from Bies 1996, Ebert 1992, and Sapp 2006. Constructions are listed in the descending order of preference for the 1-2 order.
to analyze the syntagms in three larger groupings, namely syntagms involving an IPP, an infinitive, or a past participle. Recall that subordinate clause two-verb complexes with an infinitive favor, and those with a past participle disfavor, the 1-2 order. This generalization holds for the main clause two-verb complexes as well. As table 17 shows, syntagms with two infinitives favor the 2-3 order, as do the syntagms with an IPP. The syntagms with a past participle, on the other hand, disfavor the 2-3 order.

<table>
<thead>
<tr>
<th>Verb form</th>
<th>3-2</th>
<th>2-3</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPP</td>
<td>100 (44.2%)</td>
<td>126 (55.8%)</td>
<td>0.977</td>
</tr>
<tr>
<td>infinitive</td>
<td>44 (81.5%)</td>
<td>10 (18.5%)</td>
<td>0.815</td>
</tr>
<tr>
<td>past participle</td>
<td>825 (98.9%)</td>
<td>9 (1.1%)</td>
<td>0.247</td>
</tr>
<tr>
<td>Total</td>
<td>969 (87%)</td>
<td>145 (13%)</td>
<td></td>
</tr>
<tr>
<td>p &lt; 0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 17. Effect of verb form on word order, main clause complexes.

The syntagms may also be categorized according to whether the lexical verb occurs with another lexical verb, a modal, or an auxiliary. Table 18 shows that when a lexical verb is combined with a modal or another lexical verb, the 2-3 order is favored. In contrast, combinations of a lexical and an auxiliary verb disfavor the 2-3 order.

---

44 IPP = modal IPP [H], INF + lassen IPP [H], subjunctive modal IPP [H], and ACI IPP [H]; infinitive = double INF [MOD], INF + lassen [MOD], modal future [W], and INF + modal INF [MOD]; past participle = perfect of passive [S], passive [MOD], INF + perfect [H/S], perfect [MOD], complex perfect [H], PTC + kriegen [H], INF + lassen perfect [H], PTC + perfect [H/S], INF + passive [W], INF + modal perfect [H], and PTC + INF [MOD]. Since a participle can occur both with another participle or with an infinitive, only those syntagms with two infinitives are included in the infinitive grouping.

45 For the syntagms with lassen, the tokens have been assigned to either the lexical verb or auxiliary grouping, depending on the function of lassen in the sentence.
<table>
<thead>
<tr>
<th>Verb types</th>
<th>3-2</th>
<th>2-3</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEX + MOD</td>
<td>88 (43.6%)</td>
<td>114 (56.4%)</td>
<td>0.980</td>
</tr>
<tr>
<td>LEX + LEX</td>
<td>107 (80.5%)</td>
<td>26 (19.5%)</td>
<td>0.903</td>
</tr>
<tr>
<td>LEX + AUX</td>
<td>774 (99.4%)</td>
<td>5 (0.6%)</td>
<td>0.199</td>
</tr>
<tr>
<td>Total</td>
<td>969 (87%)</td>
<td>145 (13%)</td>
<td></td>
</tr>
</tbody>
</table>

\( p < 0.001 \)

Table 18. Effect of verb types on word order, main clause complexes.

In conjunction, the hierarchies IPP > infinitive > past participle and LEX + MOD > LEX + LEX > LEX + AUX correspond roughly to the hierarchy for the subordinate clause two-verb complexes, namely INF + LEX > INF + MOD > perfect with sein > perfect with haben > werden passive, whereby complexes involving a lexical or modal verb and an infinitive more greatly favor the 1-2 order than those involving an auxiliary verb and a past participle. These hierarchies are also reflected in the three most frequently occurring main clause two-verb syntagms, the perfect of passive [S], passive [MOD], and modal IPP [H], presented in table 19.

<table>
<thead>
<tr>
<th>Syntag</th>
<th>3-2</th>
<th>2-3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>perfect of passive [S]</td>
<td>476 (100%)</td>
<td>0 (0%)</td>
<td>476 (42.7%)</td>
</tr>
<tr>
<td>passive [MOD]</td>
<td>224 (100%)</td>
<td>0 (0%)</td>
<td>224 (20.1%)</td>
</tr>
<tr>
<td>modal IPP [H]</td>
<td>79 (41.6%)</td>
<td>111 (58.4%)</td>
<td>190 (17.1%)</td>
</tr>
</tbody>
</table>

Table 19. Distribution of the three most frequently occurring syntagms, main clause complexes.

Here the additive effect of verb form and verb type is particularly striking. With the perfect of passive [S] and passive [MOD], which involve a past participle and the auxiliary werden, both at one end of their respective hierarchies, 0 out of a total 700 tokens (0%) occur with the 2-3 order. The modal IPP [H], on the other hand, which involves verbal elements at the other end of the hierarchies, an IPP and a modal, displays the 2-3 order at a rate of 58.4%, far higher than the expected rate of 13%.
4.2. Verbal Prefix Type and Two-Verb Complexes.
This section deals with the effect of verbal prefix type on word order in subordinate and main clause complexes. I begin with the subordinate clause two-verb complexes. The results in table 20 reveal that the 1-2 order is favored by the presence of a stressed separable prefix on the non-finite verb in the complex and disfavored by the presence of an unstressed prefix or the absence of any prefix.

<table>
<thead>
<tr>
<th>Prefix type</th>
<th>2-1 (80.1%)</th>
<th>1-2 (19.9%)</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>stressed</td>
<td>302</td>
<td>75</td>
<td>0.662</td>
</tr>
<tr>
<td>unstressed</td>
<td>670 (94.6%)</td>
<td>38 (5.4%)</td>
<td>0.438</td>
</tr>
<tr>
<td>no prefix</td>
<td>195 (80.9%)</td>
<td>46 (19.1%)</td>
<td>0.421</td>
</tr>
<tr>
<td>Total</td>
<td>1167 (88%)</td>
<td>159 (12%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 20. Effect of verbal prefix type on word order, subordinate clause complexes.

Sapp (2006:61–62) points out that the correlation between stressed separable prefix and the 1-2 order might be attributed to an interaction between verbal prefix type and syntagm. Past participles typically begin with the unstressed prefix ge-, and as we saw in section 4.1, syntagms involving a past participle disfavor the 1-2 order. Sapp tests for this interaction using GoldVarb’s cross-tabulation function. The results of the cross-tabulation for the Zwirner Corpus data are presented in table 21 below.

These results show that the factor groups syntagm and verbal prefix type are, in fact, independent of each other. For both syntagm groupings, that is, syntagms involving either a participle or an infinitive, the rate of 1-2 is above the expected rates (8% and 24%) when the prefix on the non-finite verb is stressed and below the expected rates for the other two prefix types. Note, however, that for all three prefix types the rate of 1-2 is higher for the infinitive grouping than for the participle grouping. It may be concluded, then, that while syntagm and verbal prefix type are independent factor groups, their effect on word order is additive.
<table>
<thead>
<tr>
<th>Verbal prefix type</th>
<th>Word order</th>
<th>Participle</th>
<th>Infinitive</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>stressed</td>
<td>1-2</td>
<td>38 (13%)</td>
<td>37 (40%)</td>
<td>75 (20%)</td>
</tr>
<tr>
<td></td>
<td>2-1</td>
<td>246 (87%)</td>
<td>56 (60%)</td>
<td>302 (80%)</td>
</tr>
<tr>
<td>stressed</td>
<td>1-2</td>
<td>35 (5%)</td>
<td>3 (8%)</td>
<td>38 (5%)</td>
</tr>
<tr>
<td></td>
<td>2-1</td>
<td>637 (95%)</td>
<td>33 (92%)</td>
<td>670 (95%)</td>
</tr>
<tr>
<td>no prefix</td>
<td>1-2</td>
<td>0 (0%)</td>
<td>46 (20%)</td>
<td>46 (19%)</td>
</tr>
<tr>
<td></td>
<td>2-1</td>
<td>13 (100%)</td>
<td>182 (80%)</td>
<td>195 (81%)</td>
</tr>
<tr>
<td>Totals</td>
<td>1-2</td>
<td>73 (8%)</td>
<td>86 (24%)</td>
<td>159 (12%)</td>
</tr>
<tr>
<td></td>
<td>2-1</td>
<td>896 (92%)</td>
<td>271 (76%)</td>
<td>1167 (88%)</td>
</tr>
</tbody>
</table>

Table 21. Cross-tabulation of verbal prefix type with syntagm, subordinate clause complexes.

Like the subordinate clause two-verb complexes, the main clause two-verb complexes exhibit an effect of verbal prefix type on word order, illustrated in table 22.

<table>
<thead>
<tr>
<th>Prefix type</th>
<th>3-2</th>
<th>2-3</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>stressed</td>
<td>287 (86.7%)</td>
<td>44 (13.3%)</td>
<td>0.654</td>
</tr>
<tr>
<td>no prefix</td>
<td>155 (62.8%)</td>
<td>92 (37.2%)</td>
<td>0.641</td>
</tr>
<tr>
<td>stressed</td>
<td>527 (98.3%)</td>
<td>9 (1.7%)</td>
<td>0.341</td>
</tr>
<tr>
<td>Total</td>
<td>969 (87%)</td>
<td>145 (13%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 22. Effect of verbal prefix type on word order, main clause complexes.

The results show that a stressed separable prefix or no prefix on the main lexical verb in the complex favors the 2-3 order. Yet unlike the subordinate clause two-verb complexes in which the presence of an unstressed prefix or the absence of any prefix has a disfavoring effect on the 1-2 order, only unstressed prefixes have a disfavoring effect in the main clause two-verb complexes overall.

As with the subordinate clause two-verb complexes, an interaction between verbal prefix type and syntagm can be tested for using
GoldVarb’s cross-tabulation function, the results of which are presented in table 23.

<table>
<thead>
<tr>
<th>Verbal prefix type</th>
<th>Word order</th>
<th>Infinitive</th>
<th>Participle</th>
<th>IPP</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>stressed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>2 (25%)</td>
<td>2 (1%)</td>
<td>40 (60%)</td>
<td>44  (13%)</td>
<td></td>
</tr>
<tr>
<td>3-2</td>
<td>6 (75%)</td>
<td>254 (99%)</td>
<td>27 (40%)</td>
<td>287 (87%)</td>
<td></td>
</tr>
<tr>
<td>no prefix</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>8 (19%)</td>
<td>7 (11%)</td>
<td>77 (55%)</td>
<td>92  (37%)</td>
<td></td>
</tr>
<tr>
<td>3-2</td>
<td>35 (81%)</td>
<td>58 (89%)</td>
<td>62 (45%)</td>
<td>155 (63%)</td>
<td></td>
</tr>
<tr>
<td>unstressed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>9 (45%)</td>
<td>9   (2%)</td>
<td></td>
</tr>
<tr>
<td>3-2</td>
<td>3 (100%)</td>
<td>513 (100%)</td>
<td>11 (55%)</td>
<td>527 (98%)</td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-3</td>
<td>10 (19%)</td>
<td>9 (1%)</td>
<td>126 (56%)</td>
<td>145 (13%)</td>
<td></td>
</tr>
<tr>
<td>3-2</td>
<td>44 (81%)</td>
<td>825 (99%)</td>
<td>100 (44%)</td>
<td>969 (87%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 23. Cross-tabulation of verbal prefix type with syntagm, main clause complexes.

The results of the cross-tabulation suggest that the effect of verbal prefix type on word order is not independent of syntagm. For syntagms involving two infinitives or an IPP, the rate of 2-3 is above the expected rates (19% and 56%) when the prefix on the main lexical verb is stressed and at or around the expected rates when there is no prefix. For syntagms involving a participle, on the other hand, the rate of 2-3 is at the expected rate (1%) when the prefix is stressed but far above the expected rate when there is no prefix. The fact that the verbal prefix types stressed and no prefix favor the 2-3 order with different syntagms indicates an interaction between the factor groups verbal prefix type and syntagm. Table 24 shows the effect of verbal prefix type on the groupings infinitive/IPP and participle separately.
Table 24. Effect of verbal prefix type on word order, main clause complexes, infinitive/IPP, and participle.

The favoring effect of a stressed separable prefix on the 2–3 order in the syntagms with two infinitives or an IPP appears less robust than in table 22, while the favoring effect of the absence of a verbal prefix on syntagms with a participle is even clearer.46

4.3. Focus and Two-Verb Complexes.
As discussed in section 4, Sapp (2006, 2007) investigates the relationship between word order in the verbal complex and object focus, in particular contrastive and new information focus, and, by extension, old information focus.47 Sapp (2006:58–59, 2007:302–303) finds that contrastive

---

46 The unstressed prefixes cannot be tested with the participle grouping because tokens with this prefix type do not occur with the 2-3 order.

47 Consider the sentences in (i).

(i) a. What did Petra do?
   Petra [Foc READ].

b. Who read?
   [Foc PETRA] read.
focus strongly favors the 1-2 order in his ENHG corpus. New information focus also favors the 1-2 order, though less so than contrastive focus. Old information focus, on the other hand, somewhat disfavors the 1-2 order. Sapp (2006:81, 2007:304) also determines that new information and contrastive focus favor every attested order except 3-2-1 (1-3-2/1-2-3/3-1-2 in combination) in the subordinate clause three-verb complexes.

Sapp (2006:37, 2007:303) concedes the difficulties inherent in ascertaining the information status of clauses in a written language. Without the intonation cues present in spoken language it is not always possible to correctly determine focus. For this reason, Sapp (2006:58–61) seeks to corroborate his findings based on information structure with focus-related linguistic correlates.

One such correlate is the status of the object NP preceding the verb cluster. Definite NPs, including pronouns, tend to represent old information, while indefinite NPs tend to represent new information and, by extension, new information focus. Sapp also examines the intervening NPs (including pronouns and QPs) in VPR constructions. He assumes

If we assume that information can be old or new, then the old information in (i)a is *Petra* and in (i)b the fact that someone read. Accordingly, the new information focus in (i)a and (i)b is *read* and *Petra*, respectively.

Consider now the sentences in (ii).

(ii)  
   a. Did Cornelius read?
       Yes, [foc SANDRA] read.
   b. Did Sandra read or did Cornelius read?
       [foc SANDRA] read.

In these examples, the focus on *Sandra*, which can either be new (iia) or old (iib) to the discourse, contrasts it with something that has already been mentioned, namely *Cornelius*.

Sapp determined the focus status of each clause in his corpus as follows. If an argument could be viewed as a contrast to another argument, then the clause was marked “contrastive.” Clauses containing an argument that had not yet been mentioned within the present section of text, or within the last one or two pages if the text was not divided into sections, were marked “new.” Clauses not marked “contrastive” or “new” were marked “old.”
that an NP that remains inside the VP and is raised, regardless of its definiteness, is focused.\footnote{Sapp (2006:59) points out that new but not contrastive information focus can be linked to the status of the precluster NP because both indefinite and definite NPs can be focused contrastively.}

A second piece of evidence that Sapp considers is scrambling, which has been shown to correlate with focus in German (see Haider & Rosengren 2003). If an object appears to the left of either an adverbial immediately preceding the verb cluster or a negator, it is scrambled and represents old information, and if it appears to the right, it is not scrambled and represents new or contrastive information.

In the present study, I chose to preliminarily limit the examination of focus to grammatical correlates. Since the occurrence of objects with adverbials or negators is too rare in the Zwirner Corpus data to determine whether objects have been scrambled, I analyzed only the information status of preceding/intervening object NPs.\footnote{An anonymous reviewer warned against the automatic equation of definiteness and indefiniteness with old and new information, respectively. Indeed, this association should be viewed as a strong tendency, one which holds in the Zwirner Corpus data. I plan to investigate focus directly in a future study.}

The preceding NPs in the subordinate clause two-verb complexes are assignable to four groups: definite, indefinite, non-NP, and zero. The definite NPs include NPs preceded by a definite article, possessive pronoun, or NumP, as well as proper names and pronouns.\footnote{An anonymous reviewer pointed out that an NP preceded by an NumP is not necessarily definite (for example, es waren (einmal) zwej Könige ‘once upon a time there were two kings’). In the Zwirner Corpus data, all such NPs are, in fact, definite.} The indefinite NPs include bare Ns, NPs preceded by an indefinite article, a QP (jeder ‘every’, kein ‘no’, alle ‘all’), or a quantificational adverb (viel ‘much’, ein bisschen ‘a bit’), as well as etwas ‘something’, etwas + adjective, alles ‘everything’, and nichts ‘nothing’. Finally, adjectives, adverbs, nicht ‘not’, modal particles, and da-compounds comprise the group of non-NPs. Since the syntagms werden passive, perfect with sein, and subjunctive with sein may not occur with object NPs, tokens with these constructions were excluded from the analysis.
The effect of the information status of the preceding object NP on word order is inconclusive. Table 25 shows factor weights for definites and no constituent above and below 0.5, respectively, and for non-NPs and indefinites around 0.5. However, with \( p = 0.168 \), the results for the factor group are statistically insignificant.

<table>
<thead>
<tr>
<th>Preceding constituent</th>
<th>2-1</th>
<th>1-2</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>definites</td>
<td>191 (88.8%)</td>
<td>24 (11.2%)</td>
<td>0.589</td>
</tr>
<tr>
<td>non-NPs</td>
<td>290 (87.6%)</td>
<td>41 (12.4%)</td>
<td>0.514</td>
</tr>
<tr>
<td>indefinites</td>
<td>132 (91%)</td>
<td>13 (9%)</td>
<td>0.497</td>
</tr>
<tr>
<td>none</td>
<td>183 (93.4%)</td>
<td>13 (6.6%)</td>
<td>0.382</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>796 (89.7%)</td>
<td>91 (10.3%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 25. Effect of the information status of the preceding constituent on word order, subordinate clauses.

The addition of intervening NPs in the VPR constructions to the indefinite NPs results in a factor weight for this factor above 0.5, but this has little effect on the factor weights overall. As can be seen in table 26, the values are even closer to 0.5 than in table 25, and the \( p \)-value is higher.

<table>
<thead>
<tr>
<th>Preceding (or intervening) constituent</th>
<th>2-1</th>
<th>1-2</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>indefinites plus intervening NPs</td>
<td>132 (88%)</td>
<td>18 (12%)</td>
<td>0.571</td>
</tr>
<tr>
<td>non-NPs</td>
<td>290 (88.4%)</td>
<td>38 (11.6%)</td>
<td>0.529</td>
</tr>
<tr>
<td>definites</td>
<td>191 (89.7%)</td>
<td>22 (10.3%)</td>
<td>0.499</td>
</tr>
<tr>
<td>none</td>
<td>183 (93.4%)</td>
<td>13 (6.6%)</td>
<td>0.400</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>796 (89.7%)</td>
<td>91 (10.3%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 26. Effect of the information status of the preceding/intervening constituent on word order, subordinate clauses.
Thus, there is a possible effect of object focus on word order in the subordinate clause two-verb complexes as indicated by the occurrence of precluster indefinite NPs and intervening NPs in the VPR constructions. Because of statistical insignificance, however, no firm conclusions may be drawn.

To test the effect of focus in the main clause two-verb complexes, I once again examined the occurrence of precluster NPs. These NPs are assignable to the same four groups as for the subordinate clause complexes: definite, indefinite, non-NP, and zero. The definite NPs include NPs preceded by a definite article, possessive pronoun, or NumP, as well as proper names and pronouns. Included in the group of indefinite NPs are bare Ns, NPs preceded by an indefinite article, a QP (jeder ‘every’, kein ‘no’, alle ‘all’), or a quantificational adverb (viel ‘much’, ein bisschen ‘a bit’), as well as etwas ‘something’, etwas + adjective, manch ‘many a’ + adjective, alles ‘everything’, and nichts ‘nothing’. Finally, adjectives, adverbs, nicht ‘not’, modal particles, and da-compounds comprise the group of non-NPs. Since the majority of tokens involving a participle are passive constructions (perfect of passive [S], passive [MOD]), which may not occur with an object NP, only the groupings infinitive and IPP were included in the analysis.

As table 27 shows, only when no constituent precedes the verb cluster is the 2-3 order favored. In contrast, the occurrence of any constituent, whether a definite or indefinite NP or non-NP, disfavors the 2-3 order.

<table>
<thead>
<tr>
<th>Preceding constituent</th>
<th>3-2</th>
<th>2-3</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>14 (22.2%)</td>
<td>49 (77.8%)</td>
<td>0.800</td>
</tr>
<tr>
<td>non-NPs</td>
<td>62 (55.4%)</td>
<td>50 (44.6%)</td>
<td>0.439</td>
</tr>
<tr>
<td>definites</td>
<td>33 (57.9%)</td>
<td>24 (42.1%)</td>
<td>0.391</td>
</tr>
<tr>
<td>indefinites</td>
<td>27 (69.2%)</td>
<td>12 (30.8%)</td>
<td>0.291</td>
</tr>
<tr>
<td>Total</td>
<td>136 (50.2%)</td>
<td>135 (49.8%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 27. Effect of the information status of the preceding constituent on word order, main clauses (infinitive and IPP).

When the intervening NPs in the VPR constructions are added to the indefinite NPs, however, the 2-3 order is favored by the absence of a
constituent and by indefinites plus intervening NPs, and disfavored by
definites and non-NPs. These results are shown in table 28.

<table>
<thead>
<tr>
<th>Preceding (or intervening) constituent</th>
<th>3-2</th>
<th>2-3</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>none</td>
<td>14 (33.3%)</td>
<td>28 (66.7%)</td>
<td>0.721</td>
</tr>
<tr>
<td>indefinites plus intervening NPs</td>
<td>27 (37.5%)</td>
<td>45 (62.5%)</td>
<td>0.617</td>
</tr>
<tr>
<td>definites</td>
<td>33 (58.9%)</td>
<td>23 (41.1%)</td>
<td>0.384</td>
</tr>
<tr>
<td>non-NPs</td>
<td>62 (61.4%)</td>
<td>39 (38.6%)</td>
<td>0.384</td>
</tr>
<tr>
<td>Total</td>
<td>136 (50.2%)</td>
<td>135 (49.8%)</td>
<td></td>
</tr>
</tbody>
</table>

Table 28. Effect of the information status of the preceding/intervening constituent on word order, main clauses (infinitive and IPP).

Thus, it may be concluded for the main clause two-verb complexes that object focus, as indicated by the occurrence of precluster indefinite NPs and intervening NPs in the VPR constructions, has an effect on word order in the Zwirner Corpus data.

4.4. Extraposition and Two-Verb Complexes.

The final factor group tested in the statistical analysis is extraposition. Sapp (2006:57–58) demonstrates that the rate of 1-2 in his corpus is far higher than the expected rate when some constituent is extraposed, but very close to the expected rate when no extraposed constituent appears to the right of the verb cluster. However, Sapp’s (2007:303–304) analysis reveals that when extraposed adjunct PPs and argument NPs/PPs are tested separately, the latter but not the former favor the 1-2 order. Since Bies (1996) shows that extraposed arguments in her ENHG corpus tend to be heavy or focused, Sapp (2007:304) concludes that the correlation in his data between the 1-2 order and argument extraposition can be attributed to focus, not to extraposition in general. He takes these findings as a further piece of indirect evidence for the effect of focus on word order.

The effect of extraposition on the 1-2 order in the subordinate clause complexes is shown in table 29.
<table>
<thead>
<tr>
<th>Extraposition</th>
<th>2-1</th>
<th>1-2</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>other extrapolated constituent</td>
<td>11 (91.7%)</td>
<td>1 (8.3%)</td>
<td>0.552</td>
</tr>
<tr>
<td>nothing extrapolated</td>
<td>1117 (87.9%)</td>
<td>154 (12.1%)</td>
<td>0.502</td>
</tr>
<tr>
<td>extraposed argument</td>
<td>8 (88.9%)</td>
<td>1 (11.1%)</td>
<td>0.439</td>
</tr>
<tr>
<td>extraposed adjunct PP</td>
<td>31 (91.2%)</td>
<td>3 (8.8%)</td>
<td>0.417</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1167 (88%)</strong></td>
<td><strong>159 (12%)</strong></td>
<td></td>
</tr>
<tr>
<td><em>p = 0.908</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 29. Effect of extraposition on word order, subordinate clause complexes.

Factor weights for extraposed arguments and adjuncts PPs indicate that these have a disfavouring effect on the occurrence of the 1-2 order, while the factor weight for other extraposed constituents points to a slight favoring effect. No firm conclusions may be drawn, however, since the results are statistically insignificant with _p = 0.908_.

The results of the analysis of extraposition in the main clause complexes are presented in table 30. Recall from section 3.6 that no argument NPs are attested in these data.

<table>
<thead>
<tr>
<th>Extraposition</th>
<th>3-2</th>
<th>2-3</th>
<th>Factor weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>other extrapolated constituent</td>
<td>18 (85.7%)</td>
<td>3 (14.3%)</td>
<td>0.762</td>
</tr>
<tr>
<td>extraposed argument</td>
<td>19 (86.4%)</td>
<td>3 (13.6%)</td>
<td>0.695</td>
</tr>
<tr>
<td>nothing extrapolated</td>
<td>872 (86.8%)</td>
<td>133 (13.2%)</td>
<td>0.494</td>
</tr>
<tr>
<td>extraposed adjunct PP</td>
<td>59 (90.8%)</td>
<td>6 (9.2%)</td>
<td>0.423</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>968 (87%)</strong></td>
<td><strong>145 (13%)</strong></td>
<td></td>
</tr>
<tr>
<td><em>p = 0.806</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 30. Effect of extraposition on word order, main clause complexes.

In contrast to extraposed adjunct PPs, extraposed argument PPs and other extraposed constituents have factor weights that indicate a favoring
effect on the 2-3 order. However, no firm conclusions may be drawn since the results are statistically insignificant with \( p = 0.806 \).

4.5. Summary of Main Findings.

Four factor groups (syntagm, verbal prefix type, focus, and extraposition) were tested with the statistical analysis program GoldVarb to determine their effect on word order in the subordinate and main clause two-verb complexes. The factor groups syntagm and verbal prefix were shown to have statistically significant effects on word order in both the subordinate and main clause complexes. Focus, as indicated by the occurrence of pre-cluster indefinite NPs and intervening NPs in the VPR constructions, has a statistically significant effect on word order only in the main clause complexes with the groupings infinitive and IPP. Finally, extraposition has no statistically significant effect on word order in either the subordinate or main clause complexes.

A glance at the corpus and the factor groups that have a statistically significant effect on word order in the subordinate and main clause two-verb complexes reveals that 156 out of 159 complexes with the 1-2 order and all 145 out of 145 complexes with the 2-3 order can be linked to one or more of these factor groups. One-hundred forty-nine out of 159 subordinate clause two-verb complexes and 136 out of 145 main clause two-verb complexes involve a syntagm or verb form that favors the 1-2 or 2-3 order, respectively. As for the factor group verbal prefix type, the rates are 75 out of 159 subordinate clause two-verb complexes and 49 out of 145 main clause two-verb complexes. Finally, the factor group focus can be linked to the 2-3 order in 45 out of 135 main clause two-verb complexes in the groupings infinitive and IPP.

Recall from section 1 that Lötscher (1978) proposed the interplay of grammatical, functional, and performance rules to account for word order variation in the verbal complex. In sections 4.1–4.3, we saw how word order in the verbal complex is influenced by the linguistic (grammatical) factor groups syntagm and verbal prefix type as well as the functional factor group focus. While performance has not yet come to the fore, one important finding from section 3.5 was the high rate of VPR in the main clause two-verb complexes. In section 6, I argue that the occurrence of VPR, alone and in conjunction with extraposition, which in this section was not conclusively shown to affect word order, can be attributed in part to the third set of rules in Lötscher’s account, namely
performance rules. First, however, let us turn to a discussion of the multi-causal account of word order in the verbal complex proposed in Lötscher 1978.

5. A Multi-Causal Account of Word Order in the Verbal Complex.

In recent decades, a substantial amount of generative literature on verbal complexes has appeared. As Wurmbrand (2006:249) points out in her empirical and theoretical overview, most analyses rest on the notion that verbal complexes arise through syntactic movement (see, for example, den Dikken & Hoekstra 1997, Haegeman 1995, 1998a, 1998b, Koopman & Szabolcsi 2000, Robbers 1997, Zwart 1996).51 In her evaluation of these analyses, Wurmbrand (2006:285–286) maintains that an explanatory syntactic analysis must be able to predict the attested and unattested word orders in verbal complexes cross-dialectally, on the basis of independent principles found elsewhere in the grammar. She notes, however, that the rules formulated to derive verbal complexes are often no more than stipulated parameters that provide only a formal description of the word order facts. She concludes the following:

While it is undoubtedly the case that descriptive accounts [...] are unattractive for their lack of explanatory value, it is not clear at the current stage of the research on verb clusters that it is trivial (possible at all) to come up with an alternative that does not suffer from these very same inadequacies (Wurmbrand 2006:294).

Wurmbrand (2006:234) sums up below what seems to be the current state of affairs, ending with a challenge for further research:

Many interesting accounts have been suggested addressing the question of how verb clusters are derived. However, what still appears to be an open question is that of why the elements of a verb cluster are inverted in certain languages and constructions. An answer to the question of what the (deep) motivation of verb-cluster formation is and why this phenomenon exists only in certain languages is still outstanding, and any new insight into this question would be highly beneficial not only

51 For alternative approaches, see Haegeman & Van Riemsdijk 1986 (reanalysis and inversion), Hinrichs & Nakezawa 1994 (HPSG), Steedman 1985 (categorical grammar), and Williams 2004 (CAT), among others.
to the study of verb clusters and restructuring, but also to our general understanding of clause structure and the architecture of grammar.\textsuperscript{52}

It would thus appear that an explanation for verbal complex phenomena perhaps cannot be adequately captured solely by a formal syntactic account.\textsuperscript{53} The question arises, then, whether the bulk of the explanation for variation in the verbal complex might not instead lie outside the core syntax. Indeed, when we look beyond the core syntax, we discover important insights into why such variation exists.

To my knowledge, L"otscher (1978) is the first to discuss the multifaceted nature of verbal complex phenomena. In his study of verb placement in Zurich German and other varieties, he writes the following:

Es ist also die Frage zu beantworten, ob es, wenn schon keine einfache Einzelregel formulierbar ist, doch wenigstens eine kleine Anzahl von generellen Regeln gibt, die die Vielfalt systematisch geordnet darstellbar machen. Dabei muß die zusätzliche Komplikation berücksichtigt werden, daß für die Erklärung der Wortstellung wohl kaum ein einziger Regeltyp vorausgesetzt werden kann. Vielmehr lassen sich mindestens drei interagierende, aber primär voneinander unabhängige Arten von Regeln ansetzen: Erstens grammatisch bedingte Regeln—Regeln wie die der Adjektivstellung bei Nomen, der Verbbestimmung usw., die mehr oder weniger willkürlich eine Abfolge bestimmen, die aber auch anders aussehen könnten; performanzbedingte Regeln—Regeln, welche es ermöglichen, performanzmäßig allzu komplexe Strukturen in leichter zu verarbeitende Strukturen umzubilden, wie etwa die verschiedenen Extrapositionsregeln in vielen Sprachen; endlich funktional bedingte Regeln—Regeln, deren Zweck die Ermöglichung des Ausdrucks von bestimmten funktionellen Satzverhältnissen im Sinne der Thema-Rhema-Unterscheidung ist. Alle drei Regeltypen wirken bei der Formation eines bestimmten Satzes natürlich mit, und so sind ihre Interaktionen oft nicht leicht zu erfassen. Dennoch würden wohl die

\textsuperscript{52} Schmid & Vogel (2004:242) set out a similar challenge: “What is needed is a systematic account not just of how orders are derived, but why dialects choose which subsets of the possible orders under which circumstances, thereby using a minimum of stipulative assumptions.”

\textsuperscript{53} One word order fact about which formal syntactic accounts have provided interesting insights is the rare cross-linguistic occurrence of the 2-1-3 order discussed in section 3.2. See Wurmbrand 2006 for details.
Gegebenheiten allzu sehr simplifiziert, wollte man sämtliche Stellungspfänomen von einem einzigen Gesichtspunkt her darstellen (Lötscher 1978:10–11).

In the remainder of section 5, I examine the components of Lötscher’s catholic approach to verbal complex formation—grammar, function, and performance—in succession, taking into account Sapp’s (2007) work and the results of the empirical and statistical analyses of the Zwirner Corpus data presented in sections 3 and 4.

5.1. Grammar.
Lötscher (1978:12) formulates a grammatical rule in terms of the position of the verbal elements in a complex in relation to one another. In effect, the instantiation of this rule arranges the verbal elements in their order of subcategorization from right to left, as in standard German for the majority of syntagms, to yield a left-branching structure. The basic

---

54 “If no single simple rule can be formulated, the question remains, then, whether there are at least a small number of general rules that can provide some order to the variation. Here the additional complicating factor, that surely no single type of rule can be assumed to explain word order, must be taken into account. Rather, at least three interacting, but fundamentally independent, types of rules may be applied: the first of these is grammatical rules—rules such as the placement of adjectives with nouns, verb-second, etc., which determine an order more or less arbitrarily but could also take a different form; performance rules—rules that enable structures, which in terms of performance are too complex, to be transformed into structures that are easier to process, such as the different extraposition rules in many languages; and, finally, functional rules—rules whose purpose is to enable the expression of certain functional sentential relationships along the lines of the theme-rheme differentiation. All three types of rules play a role in the formation of a particular sentence, and for this reason, their interactions are often not easily captured. Nonetheless, it would be an oversimplification if one tried to explain all of these word order phenomena from a single point of view” (my translation, SADS).

55 With the modal IPP/subjunctive modal IPP (as in was sie dennoch nicht hätte sagen sollen ‘what she nonetheless should not have said’), the obligatory word order in standard German is 1-3-2. The syntagms causative/ACI IPP (as in warum du mich hast kommen lassen ‘why you had me come’), future MOD + INF (as in dass sie bald wird ausziehen müssen ‘that she will have to move out soon’), and INF + modal INF + MOD (as in weil du den Kindern müsstest helfen
orderings derived through this rule may then be altered through inversion in interaction with a performance rule and are subject to functional rules (Lötscher 1978:12–13, 17–18).

I propose here that Lötscher’s (1978) grammatical rule be modified and expanded to encompass a rule and factors that interact and feed verbal complex formation. First, a grammatical rule defines, for a particular grammar, the range of word order variation in the verbal complex.\(^{56}\) This moves us already beyond the scope of a purely core-syntactic investigation concerned with exactly how the orders permitted and not permitted may be predicted, motivated, and constrained. In contrast to such an investigation, I take for granted the range of variation for a given grammar.\(^{57}\) A set of grammatical factors then constrains the orders defined by the grammatical rule. In sections 4.1 and 4.2, we saw that in the Zwisner Corpus data, the grammatical factor groups syntagm and verbal prefix type have statistically significant effects on word order in subordinate and main clause two-verb complexes. Given a set of all possible word orders in a particular grammar, there will be a number of subsets of these possible word orders, linked to different syntagms.\(^{58}\) The factor group verbal prefix type may then influence the occurrence of a particular word order within these subsets of word orders.

In the Zwisner Corpus data, for example, two word orders are possible in the subordinate clause two-verb complexes: 2-1 and 1-2. Not \(könnten\) ‘because you would have to help the children’) allow both the 1-3-2 and 3-2-1 orders. (See Dubenion-Smith 2008:43–62 for a discussion of these word orders.)

\(^{56}\) This set is subject to cross-linguistic, inter- and intraregional, as well as inter- and intraspeaker variation (see Dubenion-Smith 2008:29–70, 108–113). A grammar may be defined at any of these levels.

\(^{57}\) This proposal is supported by the fact that the range in variation in standard German, for example, is an artifact of leveling through the process of standardization and cannot, to a large extent, be derived from grammar-internal principles. Thus it seems reasonable to take as the starting point different sets of possible word orders associated with different grammars.

\(^{58}\) The constraints placed by syntagm on word order can also take the form of preference. For example, the full set of possible word orders defined by the grammatical rule may be possible for two particular syntagms, but the preference between these orders may vary.
every syntagm allows both these orders, however; while both are attested with INF + MOD, for example, only the 2-1 order is attested with PTC + *kriegen*.\(^{59}\) Which of the two word orders, 2-1 or 1-2, is realized in complexes with the syntagm INF + MOD is further determined in part by the type of verbal prefix on the lexical verb.

5.2. *Function.*

Lötscher suggests that functional rules also determine word order in the verbal complex. Lötscher (1978:4–6, 17–18) discusses a correspondence between certain word orders, specifically in main clause verbal complexes involving a modal and lexical verb plus a non-verbal element, and the position of both sentential accent and rhematic/thematic material.

Sapp (2007) more closely examines the relationship between word order and function—focus in particular. As discussed in section 4.3, Sapp found that in his ENHG corpus, new information and contrastive object focus favor the 1-2 order in subordinate clause two-verb complexes, while in subordinate clause three-verb complexes, focus favors all attested orders but 3-2-1 when taken together. Sapp supports this finding with evidence from scrambling, extraposition, and experiments testing the effect of focus on word order in modern German (see Sapp 2007:305–315). As for the Zwirner Corpus data, it was shown in section 4.3 that there is a statistically significant link between focus, as indicated by the occurrence of precluster indefinite NPs and intervening NPs in the VPR constructions, and the 2-3 order in 45 out of 135 main clause two-verb complexes in the groupings infinitive and IPP.

5.3. *Performance.*

The last of Lötscher’s (1978) rules deals with the connection between word order and performance. Recall that this rule allows structures that in terms of performance are too complex to be restructured to facilitate processing. On the basis of psycholinguistic evidence indicating that left-branching structures are more difficult to parse than right-branching ones, Lötscher (1978:12–17) proposes a rule to invert the elements in the strictly left-branching verbal complexes generated by the grammatical rule described in section 5.1 above. This inversion rule applies not only

---

\(^{59}\) For purposes of illustration, I assume that the unattested word order is un-grammatical.
to verbal elements, but non-verbal elements as well, allowing for the derivation of VPR constructions. Similarly, Kaufmann (2007:157–158, 198), following Haider 2003:91, 119–123, associates V(P)R structures with lower costs in parsing. He concludes that “the raising of VPs may be the speaker’s attempt to reduce the hearer’s processing load by producing less complex, right-branching structures” (p. 198).

The assumption that right branching structures are easier to parse than left-branching ones would seem to rest on the notion that the structural and semantic/thematic analysis of arguments is postponed in left-branching phrases, where the occurrence of the head is delayed. This delay places a load on working memory that can be alleviated through mechanisms that enable earlier analysis, for example, through V(P)R. These assumptions are compatible with head-driven approaches to parsing, as advocated by Pritchett (1991). However, it has been argued, on the basis of more recent psycholinguistic evidence from German (Bader & Lasser 1994) and Korean (Koh 1997) and confirmed, for example, in later studies of Japanese (Aoshima et al. 2003, Kamide & Mitchell 1999 among others), that the structural analysis of arguments occurs before the appearance of the verbal head, which belies a head-driven approach. Furthermore, the semantic/thematic interpretation of arguments has been shown to proceed incrementally (see, for example, Bornkessel et al. 2003 and Schlesewsky et al. 2003). If an incremental pre-head attachment account is assumed to be correct, then arguments for the advantage of right-branching V(P)R constructions for the hearer are weakened. In addition, an account that assumes the production of V(P)R structures for the benefit of the hearer entails SPEAKER ALTRUIISM, the notion that the speaker “packages” utterances in ways easiest for the hearer to parse, which has been shown in the literature to be problematic (see Kirby 1999:60–62, 135–136).60

Alternatively, the occurrence of V(P)R constructions may be associated with lower derivational costs of some sort for the speaker, as Kaufmann (2007:157–158) also suggests. In section 6, I first outline a speaker-oriented account of VPR based on principles of language

---

60 This is not to say that there is no advantage of V(P)R constructions for parsing; Bach et al. (1986), for example, demonstrate the advantage of crossed dependencies in Dutch over nested dependencies in German, associated with VR constructions.
processing proposed in Hawkins 1994 and 2004, linking this process to a productive syntactic operation in spoken German (including the Zwirner Corpus data), namely extraposition.\textsuperscript{61} I then place VPR in the broader context of speaker-hearer dichotomies, specifically those linked to bracketing, a salient characteristic of German.

Building on earlier groundbreaking work, Hawkins (2004) aims to define the relationship between word order preferences in performance and fixed word order conventions in grammars. Noting the commonly held view, dating back to Chomsky 1965, that performance has not significantly shaped grammars, Hawkins maintains that grammars are, in fact, inextricably linked to performance preferences. He defines this linkage in the Performance-Grammar Correspondence Hypothesis (PGCH):

(82) Performance-Grammar Correspondence Hypothesis (PGCH)
Grammars have conventionalized syntactic structures in proportion to their degree of preference in performance, as evidenced by patterns of selection in corpora and by ease of processing in psycholinguistic experiments (Hawkins 2004:3).

To test this hypothesis, Hawkins (2003, 2004) develops a set of general principles of efficiency and complexity that underlie word order preferences in performance and grammars: MINIMIZE DOMAINS, MINIMIZE FORMS, and MAXIMIZE ON-LINE PROCESSING. Here I limit the discussion to Minimize Domains since this is the key principle of the analysis of VPR presented in section 6.1.

Hawkins (2004:25) captures the main idea of Minimize Domains with the slogan “Express the most with least.” This principle is defined in 83.

(83) Minimize Domains (MiD)
The human processor prefers to minimize the connected sequences of linguistic forms and their conventionally associated syntactic and semantic properties in which relations of combination and/or depen-

\textsuperscript{61} In the following, I use the term “processing” in its broad sense to refer to both parsing and production.
dency are processed. The degree of this preference is propor-
tional to the number of relations whose domains can be minimized in
competing sequences or structures, and to the extent of the minimiz-
ation difference in each domain (Hawkins 2004:31).62

This principle follows from the observation that in languages that display
variation in the size of the domain in which relations of combination
and/or dependency are processed, there are clear preferences in per-
formance for minimized domains. When domain sizes have been fixed in
a grammar, this same preference is reflected in the grammatical
convention, according to the PGCH.

One instantiation of MiD is the principle EARLY IMMEDIATE
CONSTITUENTS (EIC), developed in Hawkins 1990 and 1994, which
states that words and constituents occur in the orders they do so that
syntactic groupings called PHRASAL COMBINATION DOMAINS (PCDs),
consisting of a phrasal mother node M and its IMMEDIATE
CONSTITUENTS (ICs), can be produced and recognized in performance as
rapidly and efficiently as possible.63 The formal definitions of PCD and
EIC are given in 84 and 85.

---

62 Hawkins (2004:31–32) gives the following definitions for the terms
COMBINATION, DEPENDENCY, and DOMAIN.
Combination: Two categories A and B are in a relation of combination iff they
occur within the same mother phrase and maximal projections (phrasal com-
bination), or if they occur within the same lexical co-occurrence frame (lexical
combination).
Dependency: Two categories A and B are in a relation of dependency iff the
parsing of B requires access to A for the assignment of syntactic or semantic
properties to B with respect to which B is zero-specified or ambiguously or
polysemously specified.
Domain: A combinatorial or dependency domain consists of the smallest
connected sequence of terminal elements and their associated syntactic and
semantic properties that must be processed for the production and/or recognition
of the combinatorial or dependency relation in question.

63 Hawkins (1994:73) defines ICs as “all and only the phonetically realized
lexical, functional, and phrasal categories most immediately dominated by M in
surface structure, excluding any phrasal categories that are non-maximal
projections of the head.” Here he makes the conventional distinction between
(84) Phrasal Combination Domain (PCD)
The PCD for a mother node M and its I(mmediate) C(onstituent)s consists of the smallest string of terminal elements (plus all M-dominated non-terminals over the terminals) on the basis of which the processor can construct M and its ICs (Hawkins 2004:107).

(85) Early Immediate Constituents (EIC)
The human processor prefers linear orders that minimize PCDs (by maximizing their IC-to-nonIC [or IC-to-word] ratios), in proportion to the minimization difference between competing orders (Hawkins 2004:107).

I now illustrate these principles and the mechanics of the metric involved in measuring IC-to-word ratios, beginning with the following sentences, which exemplify a performance rule of English commonly known as HEAVY NP SHIFT:

(86) a. Mary [\text{VP} \text{gave} [\text{NP} \text{the book that she had been searching for}]
  \begin{array}{ccccccccc}
   1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\
   \text{since last Christmas} & \text{[PP to Bill]} & & & & & & & \\
   10 & 11 & 12 & 13 & & & & & \\
  \end{array}

b. Mary [\text{VP g} \text{ave} \text{[PP to Bill]}\text{NP the book that she had been searching for since last Christmas}][\text{NP the book that she had been searching for since last Christmas}]
   \begin{array}{cccc}
   1 & 2 & 3 & 4 \\
   \text{since last Christmas} & \text{[PP to Bill]} & & & \\
   & & & & \\
   & & & & \\
   & & & & \\
   & & & & \\
   & & & & \\
   & & & & \\
   & & & & \\
   & & & & \\
   \end{array}

(Hawkins 2004:26)

In these sentences, the VP consists of three ICs: V, NP, and PP. In 86a, these ICs can be constructed within a thirteen-word PCD, extending from the word that constructs the first IC (V), \text{gave}, to the word that constructs the last IC (PP), to. (NP is constructed at its left edge at D, the.\textsuperscript{64}) The IC-to-word ratio, which measures the words that must be processed to construct the ICs of a phrase, is calculated by dividing the lexical categories, such as N, V, Adj, and P, and functional categories, such as C, I, and D.

\textsuperscript{64} VP, PP, and NP are constructed here according to what Hawkins (1994:342–359) terms “Mother Node Construction.”
number of ICs by the number of words in the PCD. The higher the ratio, the more minimal the PCD. This follows since a higher ratio is achieved when a given number of ICs can be constructed within fewer words.\textsuperscript{65} Thus, the ratio for 86a is $3/13 = 23\%$. In 86b, on the other hand, the same three ICs can be constructed on the basis of a much smaller PCD consisting of only four words, yielding a much higher IC-to-word ratio of $3/4 = 75\%$.

Given this difference in IC-to-word ratios, we would thus predict a preference in performance for 86a over 86b, a prediction supported by the results of a corpus study presented in Hawkins 1994:182–184.\textsuperscript{66} Hawkins (1994:83–95, 121–249; 2004:103–123) provides compelling cross-linguistic evidence that certain patterns of word order preference in performance can be derived from EIC and, more generally, MiD. I forego further discussion of the evidence here since the main purpose of this section is to introduce the mechanics of the EIC calculations.

The sentences in 86 above involve a single phrase, VP, with three ICs. In cases where more than one phrase or PCD is involved, it is necessary to determine the IC-to-word ratio for each PCD and calculate an aggregate score.\textsuperscript{67} I now illustrate such a calculation on the basis of obligatory CP extraposition in German. Consider the following sentences:

\textsuperscript{65} As Hawkins (2004:32, footnote 3) points out, the IC-to-word ratio is a shortcut for the IC-to-non-IC ratio, which more precisely captures how much non-IC material must be processed in IC construction. While the metrics yield different ratios, what is important here is the relative difference between these ratios. Crucially, Wasow (2002) has shown that the differences in the predictions of these two metrics are statistically insignificant. For a discussion of the clear practical advantages of counting words over non-ICs, see Hawkins 1994:74–76.

\textsuperscript{66} Hawkins (2004:111) notes that the heaviness of NP and PP is not the only factor that determines their relative ordering. Wasow (1997) shows that the types of collocation involved in sequences of [V PP] play an important role. Thus the ordering patterns for Heavy NP Shift depend on multiple preferences; see Hawkins 2004:111–117 for discussion.

\textsuperscript{67} Hawkins (1994, 2004) presents aggregate scores as percentages, adding up individual percentages for each PCD and dividing the result by the total number of PCDs. I choose here to express aggregates as the sum of percentages since it magnifies differences, making them clearer.
(87) a. *er \[CP_1 \text{ hatte} [VP_{CP_2} \text{ dass er nicht lange leben würde}]\]

he had that he not long

gewusst]]

live would known

‘he had known that he did not have long to live’

(Hawkins 1994:302)

b. er \[CP_1 \text{ hatte} [VP \text{ gewusst }_{CP_2} \text{ dass er nicht lange leben würde}]\]

Let us assume that C_1 \text{(hatte)} constructs CP_1, \text{V (gewusst)} VP, and C_2 \text{(dass)} \text{CP}_2. Let us further assume that \text{CP}_2 is attached to \text{VP} when extraposed. The IC-to-word ratios for 87a and 87b can then be calculated, according to Hawkins (1994:305), as follows:

(88) a. *er \[CP_1 \text{ hatte} [VP_{CP_2} \text{ dass er nicht lange leben würde}] \text{gewusst}]]

---------------------------------------------------------------

\text{CP}_1: 2/8 = 25% 

---------------------------------------------------------------

\text{VP}: 2/7 = 29%

\text{AGR} = 54

b. er \[CP_1 \text{ hatte} [VP \text{ gewusst }_{CP_2} \text{ dass er nicht lange leben würde}]\]

---------------

\text{CP}_1: 2/2 = 100% 

---------------

\text{VP}: 2/2 = 100%

\text{AGR} = 200

We see that the aggregate score for 88b is far higher than the score for 88a, indeed an optimal score of 200, which is reflected in the grammaticality of 88a vis-à-vis 88b. Notice that the results support the PGCH since the preference pattern in performance has been grammaticalized. Hawkins (1994:95–102, 250–341; 2004:123–136) discusses EIC predictions for preferences in grammar in detail; again I leave a discussion of these details aside.

The benefits of domain minimization and EIC for the hearer seem clear since the ICs in a smaller domain can be parsed more quickly. Indeed, in its initial formulation in Hawkins 1990 and 1994, the PCD is
referred to as the “Constituent Recognition Domain” (CRD), the
definition of which entails the parsing of terminal and non-terminal
nodes. Likewise, the original definition of EIC makes reference to the
“human parser.” In Hawkins 2004, on the other hand, the author
formulates his principles more neutrally in order to capture the benefits
for both the hearer and speaker, using the terminology “construct” and
experiments (see Stallings 1998) that confirm the benefits for the speaker
and argues that in fact, the metric involved in measuring EIC as outlined
above is more plausible from the speaker’s perspective than from the
hearer’s.68 Hawkins (2004:271) readily admits that the benefits for the
speaker and hearer, as they relate to his principles, must be more clearly
defined. I assume here, in light of issues raised in section 5.3 above,
benefits for the speaker in domain minimization, an assumption sup-
ported by independent evidence, as we see in section 6.2 below.69

With the tools of the IC-to-word ratio calculation in hand, let us now
examine how Hawkins’ principles of MiD and EIC can account, in part,
for constructions involving VPR formation and extraposition attested in
the Zwirner Corpus data. For the derivation of the sample sentences

68 Consider the following sentence:
The man [\text{VP waited} \text{[PP for his son]} \text{[PP in the cold but not unpleasant wind]}]

According to the procedure outlined in the main text, the PCD for this sentence
stretches across five words, from V waited to P. This implies that PP will be
constructed at the occurrence of in. Note, however, that the attachment of PP is
potentially ambiguous: if PP were [PP2 in distress], then it would be attached
lower in the tree to DP. As Hawkins (2004:106) points out, the speaker knows
the structure of PP in advance since he/she is producing the utterance and can
accordingly construct PP in the correct place upon reaching in. In contrast, the
hearer does not have access to this information and must (potentially) hear more
of the phrase until the correct place of attachment can be determined.

69 For discussion of the speaker-hearer dichotomy, see, for example, Ferreira &
below, I assume Haegeman’s (1992) analysis, which treats VPR as rightward adjunction of the VP with its arguments to a maximal projection, preceded by optional scrambling. Furthermore, I adopt Frey & Pittner’s (1998) analysis of adverbials, which, as they argue, are generated in different base positions according to semantic class and may undergo scrambling. Full underlying and surface representations for the sample sentences are provided in the footnotes; however, to facilitate EIC calculations, simplified structures appear in the main text.  

I begin with a main clause involving both extraposition and VPR:

(89) habe sollen, in die Stadt gehen, mit der Wäsche

‘(I) was supposed to go into town with the laundry’

Based on this derivation, I posit the clausal structure in 90 as the “basic” order associated with 89.

(90) \[ \text{CP have [VP\text{\textsubscript{2}}[VP\text{\textsubscript{1}} mit der Wäsche][PP\text{\textsubscript{2}} in die Stadt] gehen] sollen]} \]

This sentence, comprising the projections CP, VP\text{\textsubscript{2}}, and VP\text{\textsubscript{3}}, consists of the following ICS:

1. CP = C (habe), VP\text{\textsubscript{2}}
2. VP\text{\textsubscript{2}} = V\text{\textsubscript{2}} (sollen), VP\text{\textsubscript{3}}
3. VP\text{\textsubscript{3}} = V\text{\textsubscript{3}} (gehen), PP\text{\textsubscript{1}}, PP\text{\textsubscript{2}}

---

\[ 70 \] For full syntactic tree structures, see Dubenion-Smith 2008:215–227.

\[ 71 \] Full underlying and surface representations for 89 in the main text are as follows:

a. \[ \text{CP} \text{\textsubscript{C}} [\text{AgrP[DP ich]}][\text{TP[VP\text{\textsubscript{1}}[VP\text{\textsubscript{2}}[VP\text{\textsubscript{3}} mit der Wäsche][PP\text{\textsubscript{2}} in die Stadt] gehen] sollen]}[\text{haben}][\text{T}\text{\textsubscript{Agr}}]] \]

b. \[ \text{CP[DP ich]}_b \text{\textsubscript{b}} \text{\textsubscript{habe}}_i [\text{AgrP t}_b [\text{TP[VP\text{\textsubscript{1}}[VP\text{\textsubscript{2}} t\text{\textsubscript{s}} sollen]][VP\text{\textsubscript{3}} t\text{\textsubscript{e}} [PP\text{\textsubscript{2}} in die Stadt] gehen]]}_a t_1 t_\text{\textsubscript{1}} t_\text{\textsubscript{1}} [PP\text{\textsubscript{1}} mit der Wäsche]_c] \]

\[ 72 \] The phrases that do not contain lexical material—AgrP, TP, and VP\text{\textsubscript{1}}—have been omitted here, in keeping with common practice in performing EIC calculations.
I assume that C constructs CP, V₂ VP₂, V₃ VP₃, P₁ (mit) PP₁, and P₂ (in) PP₂. Given these ICS and their constructors, the IC-to-word ratios and aggregate score for the basic sentence in 90 are as follows:

(91) [CP habe [VP₂ [VP₃PP₁ mit der Wäsche] [PP₂ in die Stadt] gehen] sollen]

CP: 2/9 = 22%

---

VP₂: 2/2 = 100%

---

VP₃: 3/7 = 43%
AGR = 165

I now calculate the ratios and aggregate score for a theoretical sentence with extraposed PP₁ that has been attached to VP₂. Following Hawkins 1994 and 2004, I assume discontinuous constituency.⁷³

(92) [CP habe [VP₂ [VP₃PP₂ in die Stadt] gehen] sollen] [VP₃PP₁ mit der Wäsche]]

CP: 2/6 = 33%

---

VP₂: 2/2 = 100%

---

VP₃: 3/6 = 50%
AGR = 183

If we compare the aggregate scores in 91 and 92, we see that the extraposition of PP₁ results in an improvement of the aggregate score over the score for the competing basic clause, like CP extraposition in 88 above.

---

⁷³ In the derivation of the attested sentence in note 71, PP₁ is attached to AgrP. Whether PP₁ is attached to VP₂ or AgrP is inconsequential for the calculations here and those to follow; for purposes of presentation it is more convenient to show PP₁ attached to VP₂. The same holds for the extraposed PP in 103 of the main text. Note also that here and in subsequent structures, the origin of discontinuous constituents is indicated by the bracket notation. In 92 of the main text, for example, we see that the origin of extraposed PP₁ is VP₃.
The aggregate score calculations for a second theoretical sentence with VPR (raising of VP₃) alone and the attested sentence with both VPR and extraposition are as follows:⁷⁴

(93) \[ [\text{VP₃} \text{mit der Wäsche}] [\text{VP₃ in die Stadt}] \text{gehen}] \]

\[
\begin{array}{l}
\text{CP: } 2/2 = 100\% \\
\text{VP₃: } 2/8 = 25\% \\
\text{AGR = 168}
\end{array}
\]

(94) \[ [\text{VP₃} \text{mit der Wäsche}] [\text{VP₃ in die Stadt}] \text{gehen}] \]

\[
\begin{array}{l}
\text{CP: } 2/2 = 100\% \\
\text{VP₃: } 2/5 = 40\% \\
\text{AGR = 200}
\end{array}
\]

We see that in 93, VPR alone results in an aggregate score only slightly higher than the score for the basic sentence, but in 94, VPR and extraposition together show an additive effect with a greater improvement in aggregate score than with either operation alone.

Next, I calculate the aggregate scores for a pair of attested main clauses with similar S-structures that display VPR alone.⁷⁵

---

⁷⁴ Hubert Haider (p.c.), as cited in Heilman 1999:96, views the succession of verbs such as habeollen in 93 of the main text as a special case of verb-second, whereby the whole cluster has raised to C. Weiß (1998:54), on the basis of cliticization data from Bavarian, argues against this analysis. Consider the following sentence:

\[
\text{na han’e miassn}_2 \text{ de Duam-uar av-zoing}_3
\]

then has.he must the tower-clock wind

‘then he had to wind the clock in the tower’ \hspace{1cm} \text{(Weiß 1998:54)}

A clitic may only appear on the finite verb, as does the subject clitic e ‘he’ in the sentence above. If a complex verb han miassn were raised to C, then the appearance of e would be unexpected. Thus Haider’s analysis may be ruled out.
(95) **hatte müssen** __tief in den Sack__ greifen  
‘(he) had to reach deep into the bag’ (i.e., shell out a lot of money)  

(96) **hat müssen** __dem Lehrer den Stecken__ schneiden  
‘(one) had to cut the stick for the teacher’

I assume that 95 and 96 are related to the basic sentences in 97 and 98, respectively.

(97) \( [\text{CP} \text{hatte} [\text{VP}_2[\text{VP}_3[\text{PP tief in den Sack} \text{greifen}] \text{müssen}]]] \)

(98) \( [\text{CP} \text{hat} [\text{VP}_2[\text{VP}_3[\text{DP}_1 \text{dem Lehrer}[\text{VP}_2 \text{den Stecken} \text{schneiden}] \text{müssen}]]] \)

I begin with 97, which consists of three projections: CP, VP\(_2\), and VP\(_3\). For present purposes, we need be concerned with only CP and VP\(_2\). The ICs of these phrases are as follows:

1. CP = C (hatte), VP\(_2\)
2. VP\(_2\) = V\(_2\) (müssen), VP\(_3\)

I assume that C constructs CP, V\(_2\) VP\(_2\), and V\(_3\) (greifen) VP\(_3\).

Below are the calculations of the aggregate scores for the basic 99 and attested 100 sentence with VPR.

---

75 Full underlying and surface representations for 95 and 96 in the main text are as follows:

(95') a. \([\text{CP} [\text{AgP}[\text{DP (er)}]_b [\text{TP}_1[\text{VP}_1[\text{VP}_2[\text{VP}_3[\text{PP tief in den Sack} \text{greifen}] \text{müssen}]]] \text{haben}]_T[Agr]]\)
   b. \([\text{CP}_2 [\text{DP (er)}]_b \text{hatte}_b [\text{AgP}_b [\text{TP}_1[\text{VP}_1[\text{VP}_2[\text{VP}_3[\text{PP tief in den Sack} \text{greifen}]_t_1]_t_1]_t_1]]] \text{müssen}]]\)

(96') a. \([\text{CP} [\text{AgP}[\text{DP (der)}]_b [\text{TP}_1[\text{VP}_1[\text{VP}_2[\text{VP}_3[\text{DP}_1 \text{dem Lehrer}[\text{PP dem Stecken} \text{schneiden}] \text{müssen}]]] \text{haben}]_T[Agr]]\)
   b. \([\text{CP}_2 [\text{DP (man)}]_b \text{hat}_b [\text{AgP}_b [\text{TP}_1[\text{VP}_1[\text{VP}_2[\text{VP}_3[\text{PP dem Lehrer}[\text{DP}_2 \text{den Stecken} \text{schneiden}]_t_1]_t_1]_t_1]]] \text{müssen}]]\)
(99) [CP hatte [VP2[VP3[PP tief in den Sack] greifen] müssen]]

CP: 2/7 = 29%

--------

VP2: 2/2 = 100%
AGR = 129

(100) [CP hatte [VP2 müssen][VP2[VP3[PP tief in den Sack] greifen]]]

--------

CP: 2/2 = 100%

--------

VP2: 2/6 = 33%
AGR = 133

As in 93, in 100 the occurrence of VPR alone, involving a VP with an embedded PP, improves the aggregate score only slightly.

For 98, I assume the same projections and ICs as for 97.

1. CP = C (hat), VP2
2. VP2 = V2 (müssen), VP3

However, there is a difference between these sentences in the constituent structure of VP3 and the way in which this phrase is constructed. While in 97 the ICs of VP3 are V3 and PP, and V3 constructs VP3, in 98 the ICs of VP3 are V3, DP1, and DP2. Furthermore, as Hawkins (1994:393–397) argues, an accusative- or dative-marked determiner can construct VP.\(^{76}\) Thus, in 98 above, VP3 is constructed upon the occurrence of the dative-marked D1 (dem) of DP1. Based on these assumptions, then, the following are the aggregate score calculations for the basic sentence 101 and the attested sentence with VPR 102.

---

\(^{76}\) VP can be constructed by an accusative- or dative-marked determiner according to what Hawkins (1994:361–366) terms “Grandmother Node Construction.”
(101) \[ \text{CP hat } \text{VP2}\{\text{DP1 dem Lehrer}\{\text{DP2 den Stecken} schneiden} \text{ müssen}\} \]

\[ \text{CP: } 2/7 = 29\% \]

\[ \text{VP2: } 2/6 = 33\% \]

\[ \text{AGR = 62} \]

(102) \[ \text{CP hat } \text{VP2 müssen}\{\text{VP3}\{\text{DP1 dem Lehrer}\{\text{DP2 den Stecken} schneiden}\} \]

\[ \text{CP: } 2/2 = 100\% \]

\[ \text{VP2: } 2/2 = 100\% \]

\[ \text{AGR = 200} \]

We see that VPR with left-edge construction of VP₂ results in a far improved aggregate score.

I now calculate aggregate scores for the following subordinate clause, which displays both extraposition and VPR.\(^{77}\)

(103) \(\text{wo ich heute wieder soll}_{1} \text{ noch ein Stück}\)
where I today again should still a bit

\(\text{abgemacht}_{3} \text{ kriegen}_{2} \text{ mit dem Bulldog}\)
undone get with the tractor

‘where I should get another parcel of land harvested again today with the tractor’

ZWG98

Based on this derivation, I posit the clausal structure in 104, in which AgrP and TP have been conflated as IP for expository reasons, as the basic order associated with 103.

\(^{77}\) Full underlying and surface representations for 103 in the main text are as follows:

a. \[ \text{CP, wo } \text{[AgrP[DP ich][TP[VP1[VP2[VP3\{Ad, heute\{Ad, wieder\}v_{3}\{pp mit dem Bulldog\}{DP, noch ein Stück\} abgemacht\{kriegen\{sollen\}T\{Agr\}\}]]}} \]

b. \[ \text{CP, wo } \text{[AgrP[DP ich][TP[TP[VP1[VP1\{Ad, heute\}{Ad, wieder\}t_{d}t_{d}t_{d}t_{d}\}}v_{3}\{t_{c}\{DP, noch ein Stück\} abgemacht\{kriegen\{sollen\}T\}}]}]} \]
(104)  \[\text{CP wo [IP [DP1 ich] [VP2 [VP3 [Adv1 heute] [Adv2 wieder]]] [PP mit dem Bulldog] [DP2 noch ein Stück] abgemacht] kriegen soll]}\]

As we can see, the sentence in 104 consists of four projections, CP, IP, VP₂, and VP₃, with the following ICs:

1. CP = C (wo), IP
2. IP = I (soll), DP₁, VP₂
3. VP₂ = V₂ (kriegen), VP₃
4. VP₃ = V₃ (abgemacht), Adv₁, Adv₂, PP, DP₂

I assume that C constructs CP, V₂ VP₂, V₃ VP₃, P (mit) PP, and D₂ (ein) DP₂. The construction of IP, on the other hand, is less straightforward. As the phrasal head, I can construct IP. Additionally, Hawkins (1994: 384–385, 398) proposes that in languages such as German, S (IP) can be constructed by the complementizer C or by a nominative-marked NP (DP) at the left edge of S. Therefore, both wo and ich can also construct IP in 104.

Below, I present the aggregate score calculations for four sentences: the basic sentence 105, theoretical versions with extraposition 106 and VPR 107 alone, and the attested sentence 108 with both extraposition and VPR. To simplify the calculations, I do not include the ratios for CP and VP₂ since these remain constant, and I assume that DP₁ constructs IP.

---


79 In each of the sentences in 105–108 of the main text, the PCDs for CP and VP₂ consist of two words, wo ich and abgemacht kriegen, respectively. If it is assumed that C constructs IP here, it is not clear how the IC-to-word ratios for IP would be calculated since C is not an IC of IP, yet the PCD for IP must begin at the leftmost IC that constructs it.
(105) \[\text{[CP wo [IP [DP1 ich] [VP2 [VP3 [Adv1 heute] [Adv2 wieder]] [PP mit dem Bulldog]]]}}\]

\[\text{[DP2 noch ein Stück] abgemacht] kriegen] soll]\]

IP: 3/12 = 25%  
VP3: 5/9 = 56%  
AGR = 81

(106) \[\text{[CP wo [IP [DP1 ich] [VP2 [VP3 [Adv1 heute] [Adv2 wieder]] [DP2 noch ein Stück]]]}}\]

\[\text{abgemacht] kriegen] soll][[VP3[PP mit dem Bulldog]]]\]

IP: 3/9 = 33%  
VP3: 5/9 = 56%  
AGR = 89

(107) \[\text{[CP wo [IP [DP1 ich] [VP1 [VP3 [Adv1 heute] [Adv2 wieder]]]] soll]}\]

\[\text{[VP2[VP3[PP mit dem Bulldog][DP2 noch ein Stück] abgemacht] kriegen]]}}\]

IP: 3/12 = 25%  
VP3: 5/10 = 50%  
AGR = 75
(108) \[\text{CP} \text{wo} \text{[IP} \text{[DP} \text{ich}]\text{[VP}^1 \text{[VP}^3 \text{[Adv}_1 \text{heute}]\text{[Adv}_2 \text{wieder}]mdp]} \text{soll}\]

\[\text{VP}_3\]

\[\text{VP}_2\text{[VP}_3\text{[DP}_2 \text{noch ein Stück} \text{abgemacht} \text{kriegen}]\text{[VP}_3\text{[IP} \text{mit dem Bulldog}]mdp}\]

As we can see, extraposition alone results in an aggregate score higher than that of the basic sentence, but VPR alone results in a lower score. Additionally, we see that the score for the sentence with extraposition and VPR together is identical to that for the sentence with extraposition alone.

For cases in which aggregate scores for competing structures are non-optimal and need to be further distinguished since they are very close or identical, Hawkins (1994:82–83) proposes an on-line EIC metric that takes into account not only the size of the PCD but also the position of the ICs within the domain. The idea behind this metric is the preference for a leftward skewing of ICs within a domain.\(^{80}\)

On-line

\(^{80}\) Hawkins’ (1994:82–83) procedure for calculating on-line EIC ratios is as follows:

Calculating left-to-right IC-to-non-IC ratios: The L-to-R IC-to-non-IC ratio for a non-optimal CRD [PCD] is measured by first counting the ICs in the domain from left to right (starting from 1), and then counting the non-ICs (or words alone) in the domain (again starting from 1). The first IC is then divided by the total number of non-ICs that it dominates (for example, 1/2); the second IC is divided by the highest total for the non-ICs that it dominates (for example, if this IC dominates the third through seventh non-IC in the domain, then 2/7 is the ratio for the second IC); and so on for all subsequent ICs. The ratio for each IC is expressed as a percentage, and these percentages are then aggregated to achieve a score for the whole CRD [PCD].

As Hawkins (2004:50–51) notes, the on-line EIC metric is in the spirit of another of his principles, Maximize On-line Processing (MaOP), which is formulated first and foremost from the perspective of the hearer, unlike MiD. While it is beyond the scope of the present discussion to outline this principle
calculations for the four sentences are as follows, again including ratios only for those projections for which they do not remain constant, namely IP and VP3:

\[(109) \ [CP \ wo \ [IP\ [DP_1 \ ich]]\ [VP_2 \ [VP_3 \ [Adv_1 \ heute]]\ [Adv_2 \ wieder]]\ [PP \ mit \ dem \ Bulldog]]
\]

\[
\begin{array}{ccc}
1/1 & 2/2 & 3/5 \\
\end{array}
\]

\[
[DP_2 \ noch \ ein \ Stück] \ abgemacht \ [kriegen] \ [soll]]
\]

\[
\begin{array}{cc}
4/8 & 5/9 \\
\end{array}
\]

\[IP = 48\%
\]

\[VP_3 = 73\%
\]

\[AGR = 121
\]

\[(110) \ [CP \ wo \ [IP\ [DP_1 \ ich]]\ [VP_2 \ [VP_3 \ [Adv_1 \ heute]]\ [Adv_2 \ wieder]]\ [DP_2 \ noch \ ein \ Stück]]
\]

\[
\begin{array}{ccc}
1/1 & 2/2 & 3/5 \\
\end{array}
\]

\[
[abgemacht] \ [kriegen] \ [soll]]\ [VP_3 \ [PP \ mit \ dem \ Bulldog]]]
\]

\[
\begin{array}{cc}
2/8 & 3/9 \\
\end{array}
\]

\[IP = 53\%
\]

\[VP_3 = 77\%
\]

\[AGR = 130
\]

\[(111) \ [CP \ wo \ [IP\ [DP_1 \ ich]]\ [VP_2 \ [VP_3 \ [Adv_1 \ heute]]\ [Adv_2 \ wieder]]] \ [soll]]
\]

\[
\begin{array}{ccc}
1/1 & 2/4 & \\
\end{array}
\]

\[
[VP_2\ [VP_3\ [PP \ mit \ dem \ Bulldog]]\ [DP_2 \ noch \ ein \ Stück] \ abgemacht]
\]

\[
\begin{array}{ccc}
3/6 & 4/9 & 5/10 \\
\end{array}
\]

\[IP = 69\%
\]

\[VP_3 = 58\%
\]

\[AGR = 127
\]

here (see Hawkins 2004:49–61 for details), I am concerned with the potential processing benefits for the speaker.
(112) \[ \text{[CP wo [ip[DP1 ich][VP2 [VP3 [Adv1 heute][Adv2 wieder]]] soll]} \]
\[ \begin{array}{ll}
1/1 & \text{IP} \\
1/1 & 2/2 \\
VP_3 & 3/9 \\
\end{array} \]
\[ \text{[VP2[VP3[DP2 noch ein Stück] abgemacht] kriegen][VP3[PP mit dem Bulldog]]} \]
\( \text{IP = 61%} \)
\( \text{VP_3 = 73%} \)
\( \text{AGR = 134} \)

The aggregate scores reveal that if the position of ICs within each PCD is taken into account, then both extraposition and VPR alone result in a score higher than that of the basic sentence, and, in combination, the highest score.\(^81\)

Finally, I calculate aggregate scores for a subordinate clause that displays only VPR.\(^82\)

(113) dass sie da müssen\(_1\)
that they there must

\textit{einen ordentlichen Korb voll Essen} kochen\(_2\)
a a decent basket full food cook

‘that they have to cook a decent basketful of food’ ZWG92

I posit that 113 is related to the basic sentence 114 with conflated IP.

(114) \[ \text{[CP dass [ip[DP1 sie][VP[Adv da][DP2 einen ordentlichen Korb voll Essen] kochen] müssen]]} \]

114 consists of three projections, CP, IP, and VP, with the following ICs:

\(^81\) Note that the on-line EIC calculations corroborate the results of the static EIC calculations for the main clauses as well.

\(^82\) Full underlying and surface representations for 113 in the main text are as follows:

a. \[ \text{[CP dass [ArgP[DP1 sie][TP[VP1[VP2[V2[Adv da][DP2 einen ordentlichen Korb voll Essen] kochen]]]müssen][T] Agr]]} \]

b. \[ \text{[CP dass [ArgP[ArgP[DP1 sie][TP[TP[VP1[VP1[Adv da]_a t_b 1 t_b 1 t_b]müssen_1][VP2[V2 t_a DP2 einen ordentlichen Korb voll essen] kochen]]]_b]]} \]
1. CP = C (dass), IP
2. IP = I (münnen), DP₁, VP
3. VP = V (kochen), Adv, DP₂

I assume that C constructs CP, DP₁ IP, accusative-marked DP₂ VP, and D₂ (einen) DP₂.

In the following, I present the aggregate scores for the basic sentence 115 and attested sentence 116, based on ratios for only the relevant phrases, IP and VP.

\[(115) \text{CP dass [IP[DP₁ sie]][VP[Adv[da]]][DP₂ einen ordentlichen Korb voll Essen]}\]
\[\text{----------------------------------------- IP}\]
\[\text{----------------------------------------- VP}\]
\[\text{kochen] müssen]}\]
\[\text{------------------- IP: 3/9 = 33\%}\]
\[\text{--------- VP: 3/7 = 43\%}\]
\[\text{AGR = 76}\]

\[(116) \text{CP dass [IP[DP₁ sie]][VP[Adv[da]]] müssen][VP[DP₂ einen ordentlichen Korb voll Essen] kochen]}\]
\[\text{----------------------------------------- IP}\]
\[\text{----------------------------------------- VP}\]
\[\text{------------------- IP: 3/4 = 75\%}\]
\[\text{--------- VP: 3/8 = 38\%}\]
\[\text{AGR = 113}\]

A comparison of the scores for 115 and 116 reveals an improved score for the attested VPR construction over that for the competing basic order.

In conclusion, the example calculations presented above suggest a potential processing-related motivation for VPR, namely more efficient language production through domain minimization and/or leftward skewing of ICs within a domain. In this respect, VPR, an operation involving rightward movement, is related to the more general syntactic operation of extraposition, which also involves rightward movement and can result in more efficient structures. Extraposition was shown in section 3.6 to be productive in WCG verbal complexes. In addition, a sample of 1,082
subordinate and main clauses without a verbal complex from 80 Zwirner Corpus transcripts display extrapolation in 366 (33.8%) of these. To be sure, I am not arguing here that the occurrence of VPR constructions in production—let alone the grammaticalization of VPR—can be predicted solely on the basis of MiD.\footnote{However, see Kaufmann (2007:200), who suggests that for some groups of Mennonite Low German speakers, VPR is a grammaticalized performance rule, according to Hawkins’ PGCH.} Rather, efficiency in language production is just one of the factors that can account for the occurrence of VPR.

6.2. Discussion.
Section 6 began with an introduction to Hawkins’ (1994, 2004) efficiency principles Minimize Domains and Early Immediate Constituents. The first principle, MiD, reflects the hypothesis that the shortest possible strings of words necessary to process relations of combination and/or dependency are preferred. The second principle, EIC, follows from the first and predicts the preference for the minimal string of words, the Phrasal Combination Domain (PCD), on the basis of which a mother node M and its Immediate Constituents (ICs) can be constructed. The degree of preference is proportional to the difference in minimization between competing orders.

Applying the metrics for determining aggregate IC-to-word ratios, it was demonstrated, on the basis of clauses from the Zwirner Corpus data, that VPR, alone or in combination with extrapolation, can result in smaller PCDs than those found in competing basic orders. Consideration of the relative position of ICs within the PCD also revealed a potential processing advantage for extrapolation and VPR when this did not follow from domain minimization alone. The increased efficiency was construed as a factor contributing to the occurrence of VPR.

The connection between performance and VPR follows from Lötscher’s (1978) proposal that V(P)R arises in order to make sentences easier to parse, taken up in Kaufmann’s (2007) recent work. However, the approach in this paper differs from the others in that it presents the potential processing advantages of VPR solely from the perspective of the speaker. This view fits well with the notion of the KLAMMERNDES VERFAHREN (bracketing), which involves the creation of discontinuous constituents. This phenomenon, discussed in detail in Ronneberger-
Sibold 1994, is a characteristic (and, for observers like Mark Twain, notorious) feature of German syntax that has developed over the course of centuries from Old High German to modern German and is a development that continues today. Nübling (2006:91) outlines three main types of bracketing as follows, with underlined words indicating the left and right brackets:

(117) Main clause bracket
   a. Grammatical or verbal bracket
      sie hat [das Spiel nach nur
      she has the game after only
      44 Minuten und drei Sätzen] verloren
      44 minutes and three sets lost
      ‘she lost the game after only 44 minutes and three sets’

   b. Lexical bracket
      er zieht [den Vorhang] zu / auf
      he pulls the curtain shut open
      ‘he closes / opens the curtain’

   c. Copula bracket
      er ist [seit drei Jahren] Lehrer
      he is since three years teacher
      ‘he has been a teacher for three years’

(118) Nominal bracket
      das [weltberühmte, aber leider vergiffene] Buch
      the world-famous but unfortunately out-of-print book
      ‘the book that is world-famous, but unfortunately out of print’

(119) Subordinate clause bracket
      obwohl [sie das Buch nie] gelesen hatte
      although she the book never read had
      ‘although she had never read the book’
Following Ronneberger-Sibold (1994:115), Nübling (2006:91, 97–98) sees the function of bracketing in the marking of boundaries in syntactic decoding for the benefit of the hearer. Extraposition and VPR, operations which disrupt this bracketing, are from this perspective a disadvantage for the hearer, but as we have seen, can nonetheless provide processing benefits—if not for the hearer, then for the speaker. As Ronneberger-Sibold (1991:228–229) points out, this competition between the hearer and speaker is found in other linguistic domains such as phonology and morphology (see Ronneberger-Sibold 1980).


Let us now see how the grammatical, functional, and performance factors that have been shown to influence word order in the verbal complex may be unified in a single account. To facilitate discussion, I give a graphic representation of the interactions between these factors in figure 4 below.

The interactions may be exemplified by comparing word order in main clause two-verb complexes in two grammars: standard German and West Central German. We begin with all logically possible word orders, namely 3-2 and 2-3. For each grammar, a grammatical rule first determines which of these word orders is allowed. In standard German, the core syntax licenses only the 3-2 order, so no further factors may influence word order. In West Central German, on the other hand, the grammar licenses both the 3-2 and 2-3 orders, which allows the occurrence of these word orders to be influenced by additional factors. The factor group syntagm first comes into play in the form of a filter. Recall from section 3.3 that the perfect of passive [S], for example,

---

84 Hawkins (1994:402) also mentions that the brace may “facilitate gathering up the ICs in processing, distinguishing them from the ICs of higher containing phrases.”

85 See also Nübling 2006:261–262; Hawkins (1994:402) notes that the German VP bracket is regularly disrupted in order to shorten the VP PCD.

86 Szczepaniak (2006:17), for example, discusses the tension inherent in phonological differences between “syllable languages” and “word languages,” the former advantageous for the speaker, the latter for the hearer.

87 An important question that I plan to address in future work is why particular syntagms favor or disfavor particular word orders.
allows only the 3-2 order in WCG, so here again, no further factors may influence word order. In contrast, the modal IPP [H] allows both the 3-2 and 2-3 orders. The occurrence of these word orders may, in turn, be accounted for by an interaction of grammar, function, and performance.\textsuperscript{88}

Possible Variation

\begin{figure}

\begin{center}

\begin{tikzpicture}

\node {WCG}
    child {node {3-2} child {node {syntaxm filter}}}
    child {node {2-3}}
    child {node {perfect of passive [S]}}
    child {node {modal IPP [H]}};

\end{tikzpicture}

\end{center}

\caption{Interaction of grammar, function, and performance in WCG.}

\end{figure}

\textsuperscript{88} Recall from section 5.1 that in evaluating the main clause two-verb complexes on the whole, the factor group syntagm comes into play twice, first as a filter and then again as a grammatical factor that interacts with the factor group verbal prefix type and the functional and performance factors. This is because syntagm serves not only to differentiate between types of verbal complexes that display variation and those that do not; among those that do, the degree of preference for particular word orders varies by syntagm, as we saw in section 3.1–3.4.
What remains outstanding is a principled way to determine exactly how grammar, function, and performance interact. Indeed, Lötscher (1978:11) is correct in the notion that “alle drei Regeltypen wirken bei der Formation eines bestimmten Satzes natürlich mit, und so sind ihre Interaktionen oft nicht leicht zu erfassen” (All three types of rules play a role in the formation of a particular sentence, and for this reason, their interactions are often not easily captured). It seems clear that the grammatical rule must apply first since if a grammar allows no variation, then none of the factors, for example, the presence of a stressed separable prefix, will have an effect on word order.\footnote{This raises the interesting question of how grammaticality is to be defined. Schmid & Vogel (2004) find that in standard German the 3-1-2 order is possible with the syntagm future MOD + INF under certain focus conditions, while 3-2-1 and 1-3-2 are considered the “normal” orders. Under the analysis outlined here, the 3-1-2 order would have to be included in the set of orders allowed by the grammar, though it is not considered grammatical in the handbooks (see, for instance, Duden).} However, it is difficult to pinpoint the extent to which each factor can account for particular word orders and what the possible additive effects are. I leave this issue to further research.\footnote{For some interesting discussions of the link between performance and information structure, see Hawkins 1994:111–117, 237–242; 2004:122–123, Moore & Polinsky 2002:21–22, and Polinsky 2002:196–197.}

Finally, the question arises as to how the movements involved in verbal complex formation are triggered. Sapp (2006:302–305) argues, following Haider & Rosengren’s (2003) approach to scrambling in German, that the movements associated with focus are not syntactically triggered but are optional movements that take place at the syntax-semantics/pragmatics interface because they are “pragmatically felicitous.” Similarly, movement as a result of grammatical- and performance-related factors could be seen to occur at this and additional interfaces. The key here is that verbal complex phenomena ought not to be analyzed solely in terms of core-syntactic operations, for they can more adequately be captured with concomitant reference to interfaces.
8. Conclusion.
The main goal of this paper has been twofold. First, it classified the types of verbal complex phenomena that occur in WCG, a dialect region that so far has been neglected in this regard, and determined the factors that influence word order in these complexes. In the empirical part of the study, it was shown that the subordinate and main clause two- and three-verb complexes attested in the corpus display multiple word orders, VPR, and extraposition. The subsequent statistical analysis revealed that the factor groups syntagm, verbal prefix type, and focus have statistically significant effects on word order. Second, the paper outlined a multi-causal account of verbal complex phenomena, taking as the point of departure Lötscher 1978 and Sapp 2007. By demonstrating how VPR can be motivated in part as a result of more efficient processing for the speaker, this study furthers our understanding of verbal complexes beyond Lötscher and Sapp’s contributions. Given the insights that accounting for verbal complex word order in terms of interactions at interfaces provides, this study can be viewed as a model for investigations of other phenomena that have been proven difficult to explain solely in terms of core-syntactic operations.

APPENDIX

Subordinate clause two-verb complexes

1. dass wir Schläge hatten₁ gekriegt₂ that we blows had gotten ‘that we had been spanked’ perfect with haben ZW9B8

2. wenn der Vater heim ist₁ gekommen₂ when the father home is come ‘when father came home’ perfect with sein ZWP89

3. dass die Glocken that the bells subjunctive with sein ZWQ77
   nach Rom wäre₁ geflogen₂ to Rome had flown ‘that the bells had flown to Rome’
4. wo der Mist mit heraus wird1 befördert2 werden passive 1-2
   where the manure with out is hauled ZWV96
   ‘with which the manure gets hauled away’

5. dass wir Milch können, bringen2 INF + MOD 1-2
   that we milk can bringen ZWI54
   ‘that we can bring milk’

6. dass die Kinder so INF + tun 1-X-2
   that the children like that
   ihm täten1 immer nachlauf2
tim did always run after
   ‘that the children would always run after him like that’

7. wenn die dann blieben1 stehen2 INF + LEX 1-2
   when they then stayed stand ZWQ83
   ‘when they stayed still then’

Subordinate clause three-verb complexes

8. wie der Baum perfect of passive 3-1-2
   how there the tree ZWB39
   aufgestellt3 ist1 worden2
   put.up is been
   ‘how the tree was put up there’

9. dass die Kirmes subjunctive passive 3-2-1
   that the fair ZWP63
   das Heu noch nicht gemacht3 worden2 wäre1
   the hay still not made been was
   ‘that the hay still hadn’t been cut by the fair’

10. dass auch viel werden passive + MOD 3-2-1
    that also much ZW6G9
    draußen gedroschen3 werden2 muss1
    outside threshed be must
    ‘that a lot of threshing has to take place outside, too’
11. wie das allgemein \textit{werden} passive + MOD 1-3-2
how that generally ZWH83
soll$_1$ gemacht$_1$ werden$_2$
should done be
‘how that should be done generally’

12. dass man zufrieden sein$_3$ können$_2$ hat$_1$ modal IPP 3-2-1
that one satisfied be can has ZWG33
‘such that you could be satisfied’

13. was man hat$_1$ kaufen$_3$ wollen$_2$ modal IPP 1-3-2
what one has buy want ZWG09
‘what you wanted to buy’

14. was ich da habe$_1$ müssen$_2$ schuften$_3$ modal IPP 1-2-3
what I there have must work ZWV96
‘how much I had to work there’

15. dass er was anderes subjunctive modal IPP 3-1-2
that he something else ZWR19
lernen$_3$ hätte$_1$ können$_2$
learn had can
‘that he could have become something else’
(that is, learned another trade)

16. ehe ich schlafen$_3$ war$_1$ gegangen$_2$ INF + perfect 3-1-2
before I sleep was gone ZWQ10
‘before I had gone to sleep’

17. wer das hingelegt$_3$ gekriegt$_2$ hat$_1$ PTC + \textit{kriegen} perfect 3-2-1
who that laid gotten has ZWG66
‘whoever had that placed in front of him’

18. wenn es brachliegen$_3$ bleiben$_2$ müsste$_1$ double INF + MOD 3-2-1
if it lie.fallow stay had ZW7G7
‘if it had to remain lying fallow’

19. dass ... Welsch ... double INF + MOD 1-3-2
that ... Welsch ... ZWP61
200 Taler ... kann_1 abheben_3 gehen_2
200 Taler ... can withdrawal go
‘that ... Welsch can go withdraw 200 Taler ...’

20. dass man dann manches INF + lassen + MOD 1-3-2
    that one then some ZWG92
    muss_3 sein_3 lassen_2
    must be let
    ‘that you have to let some things be’

Main clause two-verb complexes

21. ich habe sie nicht essen_3 k2önnen_2 modal IPP [H] 3-2
    I have it not eat can ZW803
    ‘I couldn’t eat it’

22. die hätten Häuser dafür subjunctive modal IPP [H] 3-2
    they had houses for that ZWA70
    bauen_3 sollen_2
    build should
    ‘they should have built houses for that’

23. habe ich geholfen_2 bohren_3 INF + perfect [H/S] 2-3
    have I helped drill ZWP89
    ‘I helped drill’

24. musst du doch mal gehen_2 gucken_3 double INF [MOD] 2-3
    must you PART PART go look ZW9B3
    ‘you have to go take a look’

25. ich habe ihn lassen_2 laufen_3 INF + lassen IPP [H] 2-3
    I have him let run ZWX02
    ‘I let him run with it’

26. man konnte sich lassen_2 geben_3 INF + lassen [MOD] 2-3
    one could REFL let give ZWW69
    ‘you could have someone give you (X)’
27. ich habe hauptsächlich Kartoffeln  ACI IPP [H] 3-2  
I have mainly potatoes  ZW9G0
   ausmachen, helfen
   dig up help
   ‘I mainly helped dig up potatoes’

28. da werden wir müssen, aufhören  modal future [W] 2-3
   there will we must stop  ZWI54
   ‘we’ll have to stop there’

Main clause three-verb complexes

29. dann hätten wir  INF + lassen + MOD IPP [H] 4-2-3
   then had we  ZWW58
   die Dränage ruhig liegen, können lassen
   the drainage PART lie can let
   ‘then we could have gone ahead and left the drainage lying’

30. hat da auch  double INF + MOD IPP [H] 3-2-4
   has there also  ZWG03
   helfen, wollen, spinnen
   help want spin
   ‘[she] also wanted to help spin (yarn)’

REFERENCES


*SADS–Syntaktischer Atlas der deutschen Schweiz*. Research project at the University of Zurich, Department of German (project director: Elvira Glaser). Swiss National Science Foundation 1114-57121.99.


Wurmbrend, Susanne. 2000a. Verb clusters: Variation at the right periphery. Talk given at the workshop on Syntactic Microvariation, Meertens Institute, Amsterdam.


Department of Modern and Classical Languages
Western Washington University
516 High St. HU 243A
Bellingham, WA 98225
[shannon.dubenion-smith@wwu.edu]