Clitic-Second Languages and Verb-Second languages in a diachronic perspective

The paper discusses the relation of Germanic word order systems based on the Verb-Second Constraint (=V2 systems) and word order systems with clitic clusters in clausal 2nd position (= systems with Wackernagel’s Law ~ W-systems ~ systems with C-oriented clitics). V2 languages are typologically rare outside Europe, while languages with 2nd position clitics are attested both in Old and Modern Indo-European Languages (cf. Hittite, Avestan, Old Greek, Old Novgorod Russian, Serbo-Croat, Czech, Slovene, Pashto, Osset) and in a number of non-Indo-European languages, cf. Kabyle Berber (Afroasiatic), Lummi (Salish), Makah (Wakashan), Warlpiri and Djaru (Pama-Nyungan), Quiavini Zapotec (Otomangean), Mayo (Uto-Aztecan) etc. It is plausible that all languages with Wackernagel’s law, irrespective of their genetic origin, share a number of constraints on clitic placement (clusterization rules, movement patterns, orientation of clitics), cf. [1], [2], [12], [5].

1. V2 languages are mostly attested in languages from the Germanic group\(^1\). W2 languages are widely attested in genetically unrelated languages from different families [13].

2. There are, however, no *V-3, *V4 languages, and no *W-3, *W5 languages etc. Is there a magic of clausal 2nd position or a general principle of constituent ordering?

3. Approaching Second: V2/V1 languages and W2/W1 languages.

<table>
<thead>
<tr>
<th>Topological constraint</th>
<th>Germanic languages</th>
<th>South and West Slavic languages</th>
</tr>
</thead>
<tbody>
<tr>
<td>V2 ~ W2</td>
<td>German, Dutch, Swedish, Norwegian, Faroese</td>
<td>Serbian/Croatian/Bosnian, Slovene, Bulgarian, Czech, Slovak</td>
</tr>
<tr>
<td>V2/V1 ~ W2 / W1</td>
<td>Icelandic, Yiddish</td>
<td>Macedonian, Upper Sorbian</td>
</tr>
</tbody>
</table>

\(^1\) Research is a part of the project “The typology of free word order languages and models of inversion” funded by the Russian Foundation for Humanities (grant RGNF 06-04-00203a), whose support is gratefully acknowledged. I am grateful to the audience of the SWL 3 conference for the valuable feedback. All responsibility for the statements made here is my own.

\(^2\) A similar, though not identical mechanism (the so called Operator-Second Constraint) is attested in Kru languages (Niger-Kongo family). Verb-Second phenomena are also attested in Rhaeto-Romance languages and Kashmiri, in both cases under the influence of Germanic languages.

- The V2 constraint is a language/type-specific mechanism, it is both observable and verifiable. The EPP is a theoretical postulate which cannot be verified, at least in the framework of Minimalist Grammar(s). The EPP is extended to languages which do not conform to a requirement that the preverbal position (SpecTP) should be overtly filled by one and only one phrase.

5. What is on First?

- **Strict V2 languages**: only one maximal projection of a given type is permitted.
- **Liberal V2 languages**: initial strings consisting of different syntactic elements probably count as one projection when they end up in the preverbal position.

<table>
<thead>
<tr>
<th>Language</th>
<th>Example</th>
<th>Preverbal Position</th>
<th>Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Far.</td>
<td>#Men av allari r\text{\textasciicircum}vuni av l\text{\textasciicircum}ufiski R\text{\textasciicircum}stuviku</td>
<td>var</td>
<td>meginparturin hv\text{\textasciicircum}sa.</td>
</tr>
<tr>
<td>Ice.</td>
<td>R\text{\textasciicircum}\textless amenn NATO \text{\textasciicircum}\textasciitilde heims</td>
<td>s\text{\textasciicircum}u</td>
<td>\textless st\text{\textasciicircum}xu til ax hvetja R\text{\textasciicircum}dana til ax binda enda...</td>
</tr>
<tr>
<td>Middle Swe.</td>
<td>#\text{\textasciicircum}E\text{\textasciicircum}n om th\text{\textasciicircum}Fn tima j \text{\textasciicircum}egyptu lande</td>
<td>haffdo</td>
<td>Di\text{\textasciicircum}ffla mykit wald (MS 183).</td>
</tr>
<tr>
<td>O.Eng.</td>
<td>#y\textasciitilde d\text{\textasciicircum}br (1) h\textasciicircum}2</td>
<td>h\textasciitilde ta\text{\textasciicircum}x</td>
<td>hr\textasciitilde nas.</td>
</tr>
<tr>
<td>O.Eng.</td>
<td>#Fela spella (1) \textasciitilde him (2)</td>
<td>s\textasciicircum}don</td>
<td>z\textasciitilde Beormas.</td>
</tr>
<tr>
<td>Middle Eng.</td>
<td>Grehoundes (1) he (2)</td>
<td>hadde</td>
<td>as swift as fowel in flight. (Chaucer, ca. 1380)</td>
</tr>
</tbody>
</table>

6. Crucial difference between V2 (V2/V1) and W2 (W2/W1) systems.

- All W2 systems permit clitic clusters in clausal 2\textasciicircumnd position. V2 systems do not permit predicate strings in clausal 2\textasciicircumnd position: finite and infinite verbals take different positions.
- The W2 clitics are placed in a rigid order, when they take contact position in a cluster\textsuperscript{3}.
- All W2 languages make use of at least two different categories of clitics — clitic pronouns, clitic particles, clitic auxiliaries etc.

Other differences, e.g. the variation W2 clitics after the first phonological word ~ W2 clitics after the first spelled-out maximal projection are less salient.

- The claim that all W2 languages have all clitics in clausal 2\textasciicircumnd position in 100\% cases is not consistent with data.
- One has to develop a new typology of W2 languages.

7. Systems with verbal clitics (Cf. Romance and Balcanic languages).

- Verbal clitics are localized in the VP or in a projection covering the VP but more narrow than clause.

\textsuperscript{3} A diachronic analysis of clitic clusters in a group of W2 languages is offered in [14].
• Verbal clitics move inside their domain.
• Verbal clitics lack a fixed position defined with respect to the clausal left edge.
• Verbal clitics are sensitive to Barrier rules.

8. {Definition}: A Barrier is a syntactic category, which takes effect on the surface position of a single clitic/clitic cluster.

• From a mathematical/computational viewpoint Barrier rules are Context-Sensitive rules, but it is tentatively possible to reformulate them as Mildly-Context-Sensitive rules and integrate them in a Minimalist Grammars, though this issue is currently not pursued formally.
• Barrier rules may have a threefold effect on the position of barrier-sensitive categories: they can either a) change the orientation of clitics/clitic clusters towards the host category; b) change the distance between the host category/edge of the syntactic structure and the clitics/clitic clusters; b) split clitic clusters by moving one or more clitics outside the cluster.

8.1. Barriers changing the orientation of clitics: an illustration from Romance languages.

<table>
<thead>
<tr>
<th>European Portuguese</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clauses without a Barrier</strong></td>
</tr>
<tr>
<td><strong>Base order, enclisis [V — CL OBJ]</strong></td>
</tr>
<tr>
<td>&lt;I know&gt; [that] Joseph (2) it=gave (3)</td>
</tr>
<tr>
<td>A Maria [1], O Jos [2] ofereceu=o (3) ontem to Mary (1) Joseph (2) gave=it (3) yesterday</td>
</tr>
</tbody>
</table>

9. A typology of <non-chomskyan> Barriers:

• Obligatory vs optional Barriers.
• Grammaticalized vs communicative Barriers.
• Cumulative (two or more Barriers count as a single Barrier) vs undoing Barriers (the second Barrier blocks the effect of the first one).
• Blind vs selective Barriers.


• Blind Barriers just shift clitic clusters $n$ steps to the right from clausal 2$^{nd}$ position.
• Selective Barriers are sensitive to the category of clitics.

11. Negation as a selective Barrier.

11. 1. Macedonian. Negation $ne$ is as a barrier only for pronominal clitics, not for auxiliary clitics.\(^4\)

|Maced. | $^{[\text{BAR}_{1}\text{Ne}]}$ (1)=se$_i$ (2) bespokojte! $\Rightarrow$ $[[\text{Ne}](1)$ bespokojte(2)=se$_i$ (3) |

\(^4\) I am applying subscripts of the type $^{[\text{BAR}_{1}\text{B}]} (x, y)$ order to show that sentence category B is as a Barrier only for the element x, not y.
[\text{\text{BAR}} \text{ Not-Neg}] (1) = \text{CL. Refl} \text{ I (2) worry-Imp.2Sg} \ 'Don’t worry'.

Mac. Ti(1) \text{ ne-Neg}(2) = \text{si (3)} = \text{mu (4)} \text{ spomo\ddi (5)}
You-2Sg. \text{ Not-Neg (2) =CL.-Aux2Sg (3)} = \text{CL.DatSgM} \text{ help.Pst.3Sg.M.}
‘You did not help him’.

11. 2. Bulgarian. Negation ne is a barrier only for the question particle ли, not for the other clitics.

Bulg. Детето(1) [\text{\text{BAR}} \text{ ne}] (2) = \text{cu (3)} = \text{ço (4)} \text{ виждал (5) днес?}
Child-Det. (1) [\text{\text{BAR}} \text{ not}] (2) = \text{CL. Aux2Sg (3)} = \text{CL.AccSgM} \text{ see-Pst.2Sg today}
‘The child, did you see him today?’.

Bulg. *[\text{\text{BAR}} \text{ ne}]=\text{лю (1)}=\text{cu} (2) \text{ ходил там} \Rightarrow [\text{ne}]=\text{cu} (1)=\text{лю}
[\text{\text{BAR}} \text{ Not-Neg}]=\text{CL. Q (1)} = \text{CL.Aux2Sg} (2) \text{ walk-Pst.Sg.M there}
\text{Int. ‘Wasn’t it you who walked there?’}

13. Old Novgorod Russian:

- Communicative Barriers move the reflexive pronoun ся out of the clitic cluster:

[\text{\text{BAR}} \text{ како}] = \text{есъмъ} \text{ жибъ} (Birch Bark letter 354)
[\text{\text{BAR}} \text{ as}] = \text{CL.Aux1Sg j} \text{ arrange-Perf.1Sg.M} = \text{CL.Refl i, so and live-Pres.1Sg}
‘I live exactly as I arranged it’.

- Sporadically, the same mechanism moves clitic particles, e.g. question particle ли out of the clitic cluster.

[\text{\text{BAR}} \text{ а оу королева}]=\text{есъ} \text{ жиба} || \text{слышалъ}=\text{ли} \text{ в томъ чётном кръте?} (Ipatievskaja Chronicle, under 1152. list 166 rev.).
[\text{\text{BAR}} \text{ and from king’s}] = \text{CL.2SgAux j} \text{ man // hear=CL. Q i about that worthy cross}
‘Haven’t you heard about this worthy cross from the king’s man?’.

14. Clausal 2\text{nd} position and derived position of clitics in 2W languages.

- In a vast majority of cases the clitic, which leaves its cluster, ends up in a contact pre- or postposition to a verbal form.
- Bulg. future particle \text{ве} + question particle ли:

Bulg. Книгата (1) [\text{\text{BAR}} \text{ ве}] (2)=\text{cu} (3)=\text{як (4)} || \text{прочел (5)}=\text{лю (6) до утре?}
Lit. ‘Book-the (1) [\text{\text{BAR}} \text{ FUT}] (2)=\text{AUX-2Sg} j (3)=\text{AccFSgk (4)} || \text{read-Prf (5)}=\text{Qc (6) tomorrow?’}

It is arguable that at least some V2 languages have had Wackernagel’s law on an earlier stage. This possibility was discussed already by Berthold Delbrück [6], who referred to Germanic languages. Recently, Stephen Anderson [4] repeated the same hypothesis:

(i) Pure W-systems \rightarrow Systems with Verb-Second constraint.

According to Delbrück and Anderson, Germanic Verb-2\text{nd} constraint emerged as a result of analogous leveling: 2\text{nd} position of clitic auxiliaries was later generalized for all verbal forms, stressed and unstressed. I will argue that Delbrück-Anderson’s hypothesis is falsified by Old Germanic Data and propose an alternative explanation:

(ii) Germanic clitics have attracted finite verbal forms to clausal 2\text{nd} position.
15. Where have all Old Germanic clitics gone?

Delbrück-Anderson’s hypothesis predicts that after the generalization of the V2 constraint the class of W-2 clitics was lost. We will demonstrate that Old Scandinavian languages have preserved a large class of pronominal and adverbial clitics; clitic clusters in clausal 2nd or clausal 3d position are attested up to the end of the 16th century. During this period the V-2 constraint was superimposed on an older system with Wackernagel’s law.

- In Scandinavian main declaratives finite verb could either take 2nd position or move to clause-initial position (=V2/V1 constraint). Consequently, Old Scandinavian clitic clusters could end up in 2nd position only in verb-initial clauses (#Vf — CL), but had to take clausal 3d position in all other cases, where the preverbal position was filled by any other category (#XP — Vf — CL).


- Old Icelandic has a rule of clusterization, where particles nd “now”/ z<“then” take the central slots: pronominal clitics stand to the left of nd “now”/ z<“then”, unstressed adverbials and verbal particles stand to the right of nd “now”/ z<“then”. Clitic clusters can consist of 4-5 elements lacking phrasal stress.

<table>
<thead>
<tr>
<th># Vf — X nd/ z&lt;</th>
<th># Vf — XP — nd/ z&lt;Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vil</td>
<td>ek (1) nd(2)</td>
</tr>
<tr>
<td>Lit. &quot;Want</td>
<td>l(1) now (2), that we go.1.Pl.Opt. to Norway both&quot;</td>
</tr>
<tr>
<td>O.Ice. Ferr=3Sg.Pres.</td>
<td>zat =NOM (1) nd (2) allt (3)</td>
</tr>
<tr>
<td>Lit. &quot;Goes</td>
<td>it (1) now (2) all (3)</td>
</tr>
</tbody>
</table>

17. Middle Norwegian (ca. 1600).

- A V2/V1 language. Clitics either in 2nd position or in 3rd position$^5$.
- Clitic template: Subject pronoun > reflexive pronoun > object pronouns > deictic pronouns nu/da > negation / particle vel > postverbs > prepositions.

Middle No. thi giorde |=de=sig| strax fergid at drage fra Byen (PCl 99)

$^5$ The status of Middle Norwegian 2P/3P clitics is established on the basis of their syntactic behavior, in particular, from their ability to clusterize in a rigid order and attach to different types of hosts. Learned native speakers of Modern Scandinavian languages might object that one lacks direct prosodic criteria of clitichood if one deals with dead languages and that sentences resembling those provided in (17) are accepted in Modern Danish, Swedish and Norwegian, where the corresponding pronominal and adverbial elements are usually not treated as clitics. However, Danish, Swedish and Norwegian lack verb-initial orders, which still were possible in Middle Norwegian. Furthermore, it is likely this similarity is not accidental and that Modern Mainland Scandinavian languages just generalized the positions of subject, light sentential adverbs and light objects on the basis of the former clitic positions, which is argued in [13].
Lit. 'therefore made |=they=themselves| at once ready to go out of the town’

Middle No. oc lade |=sig=ikk| myrde inde i Husit (PCI 10)
Lit. 'and let ||=oneself=not| kill inside the house’

Middle No. ginge |=de=da=ind=paa| det store skib som... (PCI 92)
Lit. 'went |=they=than=inside=on| the large ship that…’

Middle No. Oc meente |=jeg=mig=nu| hos hannem at vFr|e| fri for saadan sag (PCI 27)
Lit. 'and considered ||=I=myself=now| by him to-be free from such matters’.

Middle No. oc befand |=han=si| da megit suag at vFr|e (PCI 17)
Lit. 'and found ||=he=himself=then| very weak to be’.

Topical initial constituents as communicative Barriers:

Middle No. efter beggis deris drd 1(1)|| skulde (2) ||=da (3) | deris S|n tage Riget, efter den anden, som ectef|d vaar (PCI 124)
Lit. 'after the death of them both1(1)|| shoudl (2) ||=then (3)|| their son take the kingdom, one after the other, who is born legitimate’.

18. Main Proposal:

- Proto-Germanic W2 clitics moved out of clausal 2nd position and left it vacant for finite verbals: W2 clitics typically moved to the right, if the preceding category on some reasons could not host them.

- An exact parallel is furnished by Modern Slovak and other present-day Slavic languages with W2 clitics and Barrier rules.

19. Modern Slovak: the V2 constraint is in the making.

- Slovak is a W2 language, where clitic clusters end up in 2nd second position, while the placement of finite verbal forms in most cases is irrelevant. This holds for reflexive and possessive clitics, too, cf. example (a).

- However, if the initial constituent is long (consists of more than one word) and has a syntactic structure of its own, possessive and reflexive clitics move one step to the right: in this case, clausal 2nd position is filled by moved verbal forms, cf. example (b).

<table>
<thead>
<tr>
<th>Slovak: ‘weak clitics’</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) Aby That si for-oneself človek a man [predstavil __ srazu] [could imagine__at once]</td>
</tr>
<tr>
<td>XP</td>
</tr>
<tr>
<td>b) Vodič autobusu zapáli si lighted for himself cigaretu a cigarette</td>
</tr>
</tbody>
</table>
20. Struggling for 2\textsuperscript{nd} position: ‘strong’ vs ‘weak’ W2 clitics. The case of Bulgarian:

- Weak W2 clitics of the Slovak or Serbo-Croat type may move out from clausal 2\textsuperscript{nd} position and leave it free for other sentence categories. Languages from this subclass of W2 languages usually have clitic climbing, i.e. raising of clitics from embedded clauses into a higher clause.
- There is another subclass of W2 languages.

Bulgarian/Central Philippine system: ‘strong clitics’

<table>
<thead>
<tr>
<th>Clitics in 2\textsuperscript{nd} position, the verb takes contact position to them</th>
<th>Well-formed orders</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clause-initial clitics</strong></td>
<td><strong>#XP — CL — Vf — YP ~ #Vf — CL — YP</strong></td>
</tr>
<tr>
<td><strong>Clitics stand to the right from 2\textsuperscript{nd} position, the Barrier is absent</strong></td>
<td><strong>#CL — Vf</strong></td>
</tr>
<tr>
<td><strong>Clitics take 2\textsuperscript{nd} position, the verb doesn’t take contact position to them</strong></td>
<td><strong>#XP — CL — [...] — Vf</strong></td>
</tr>
</tbody>
</table>

Tagalog, Central Philippine:

Tag. kahapon = *ka= ba= nila* nakita?
---
\begin{itemize}
\item yesterday 2SgSbj Question p. 3Pl.DO be.seen
\item “Was it yesterday, they saw you?”
\end{itemize}

Tag. *Nakita = ka= ba= nila* (kahapon)?
---
\begin{itemize}
\item “Did they see you (yesterday)?”
\end{itemize}

- Typological prediction: W2 languages with ‘strong W2 clitics’ in the sense of (20) do not allow clitic climbing, since strong W2 clitics cannot leave the second position in their clause.

21. Conclusions:

- In a group of W2 languages, including some modern Slavic languages, clitics move from clausal 2\textsuperscript{nd} position to the right due to Barrier rules. As a result, clausal 2\textsuperscript{nd} position becomes available for verbal forms. The same mechanism can be reconstructed for Old Germanic languages.
- In modern Slavic languages, except for Bulgarian, the order #XP — Vfin — CL with clitics in clausal third position exists as a variant of the base order #XP — CL. In Old Germanic languages with the V2/V1 constraint finite verbals could not move to the right from clausal 2\textsuperscript{nd} position, therefore the order *#XP — CL — Vfin was ungrammatical in main declaratives.
- Another scenario is attested in modern Bulgarian and in Central Philippine languages. Bulgarian/Central Philippine clitics do not give up clausal 2\textsuperscript{nd} position to verbal forms, but attract them to clitic-adjacent positions.
- Bulgarian bans distant position of clitics and verbal forms, the order #XP — CL — Vfin is obligatory for all clauses with a non-initial verb. Bulgarian serves as a mirror image of Old Germanic V2/V1 languages, which generalized the order #XP — [Vfin — CL].
• In clauses with initial verbs Germanic and Slavic languages with clitics are similar: both groups of languages apply to the order #Vfin — CL.
• Old Scandinavian and Middle Scandinavian go back to a subclass of W2 languages, which permitted only one initial Barrier.

22. Problems:

The absence of the V2 in Gothic and considerable freedom in filling the preverbal position in group of V2/V1 languages (cf. Old English, Faroese, Middle Swedish) probably indicates that the V2 ~ V2/V1 constraint could be grammaticalized in a different way in different parts of the Old Germanic areal. Different groups of Germanic languages reveal a close resemblance to different groups of W2 languages.

REFERENCES