1. Introduction

Breton word order:

The prevailing views of Breton word order are the classical assumption that Breton is a verb-initial language with topicalization (Anderson 1981) and the current Principles & Parameters view that Breton shows a verb-second order in affirmative root clauses and a verb-initial order in negative and subordinate clauses (Schafer 1995). One preverbal position in V2 clauses is usually posited--for example, Spec-CP in Schafer 1995. We agree that Breton is a verb-second language, but we posit three preverbal positions: external, medial and internal, although we won’t argue for that here.

Issue:

While there is a consensus on the central role of the preverbal position in Breton grammar, there has been little systematic study of which constituent is selected for placement in preverbal position. This position has been associated with contrastive focus (Hendrick 1990), but preverbal topics occur there as well (Timm 1991).

Claim:

Based on the results of a natural discourse study of preverbal noun phrases, we claim that referentiality constrains word order in Breton. In a sentence containing two overt NPs, one preverbal and one postverbal, the preverbal NP will be more referential than the postverbal NP in the sense that its referent will be more accessible or concrete in the discourse context. However, this Referentiality Constraint can be overridden by Focus.

Corpus:

To test this constraint on preverbal NPs, we studied five texts written in the Gwenedeg dialect, each containing 60 sentences with two overt NPs in the order NP-V-NP. The first text is Yann Ar Baluc’henn, YAB, Jaffré (1986), a collection of short story memoirs originally published weekly in a local newspaper. The second text is Hor Bara Pamdiek, HBP, Guilloux (1992), a book-length childhood memoir, from
which two excerpts have been chosen for this study. These two texts, YAB and HBP, are written in the accepted Gwenedeg orthography. The fourth text is *Evit Ket ha Netra*, Ar Mason (1986), a historical novel written in the rather formal Breton standard orthography. The last text is *Melanie, ur vu hé é Groay*, Corne (1991), an oral narrative recorded in 1982/83. The informant was one of the last native speakers of the variety of Gwenedeg spoken on the island of Groix. This narrative is phonetically transcribed.

The preverbal NPs under investigation are: Subjects (S), Objects (O), Attributes (At), and Left-dislocated NPs (Ld). They occur in five word order patterns: SVO, SVAt, OVS, AtVS, LdVX. The distribution of word order patterns is given in (1).

(1) Distribution of Word Order Patterns

<table>
<thead>
<tr>
<th>Pattern</th>
<th>N</th>
<th>%</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVO</td>
<td>178</td>
<td>59.3</td>
<td>subject-initial transitives</td>
</tr>
<tr>
<td>SVAt</td>
<td>67</td>
<td>22.3</td>
<td>subject-initial equatives</td>
</tr>
<tr>
<td>OVS</td>
<td>18</td>
<td>6.0</td>
<td>object-initial transitives</td>
</tr>
<tr>
<td>AtVS</td>
<td>17</td>
<td>5.7</td>
<td>attribute-initial equatives</td>
</tr>
<tr>
<td>LdVX</td>
<td>20</td>
<td>6.7</td>
<td>left-dislocation-initial</td>
</tr>
</tbody>
</table>

It can be seen that the vast majority of preverbal NPs are subjects--245 out of 300 tokens, or 83%. It should be noted that the NP-V-NP clauses represent only 8% of the total corpus under investigation--300 tokens out of 3,732 clauses--and rank only 7th in the 11 word order patterns identified.

2. The Referentiality Constraint

2.1 Definiteness

The first phase of the study involved coding the NPs for definiteness. Since definite NP tokens denote entities which are uniquely identifiable in the discourse context, they are inherently more referential than indefinite NPs which typically are not used to refer to identifiable entities. Definiteness is realized in Breton by personal pronouns, demonstrative pronouns and determiners, the definite article *er/en* in the singular and plural, and the absence of determiners with proper names and in singular possessive constructions. Indefiniteness is realized by the indefinite article *ur/un/a/an* in the singular, and the absence of determiners with indefinite plural and predicative singular NPs. We coded universally quantified NPs as definite; and we coded negative polarity items and existentially quantified and interrogative NPs as indefinite. The distribution of definite and indefinite NPs is shown in (2).
When one NP is definite and the other indefinite, the definite NP occurs preverbally 72% of the time, or 119 tokens out of 166, as predicted by the referentiality constraint. However, 47 tokens show a preverbal indefinite NP with a postverbal definite NP, a combination which clearly violates the Referentiality Constraint. (This is the class of tokens to which we will later argue that Focus Override most centrally applies.) Moreover, when both NPs are definite (118 tokens) or when both are indefinite (16 tokens), definiteness alone cannot determine whether or not the referentiality constraint is met. We thus have a total of 181 tokens which are not automatically accounted for by definiteness.

### 2.2 Referential Dependence

The next phase of the study involved coding these 181 tokens for Referential Dependence. Referential Dependence refers to one NP dependent on the other for interpretation of its referent. We take the Referentiality Constraint to predict that the referentially independent NP will occur preverbally in such cases. This prediction is realized in all 37 cases, as shown in (3).

(3) **Distribution of Referentially Independent NPs:**

<table>
<thead>
<tr>
<th>preverbal</th>
<th>postverbal</th>
<th>initial NP</th>
<th>N</th>
<th>%</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite</td>
<td>definite</td>
<td>S-26, O-0, At-0, Ld-6</td>
<td>32</td>
<td>27.1</td>
<td>118</td>
</tr>
<tr>
<td>indefinite</td>
<td>definite</td>
<td>S-5, O-0, At-0, Ld-0</td>
<td>5</td>
<td>10.6</td>
<td>47</td>
</tr>
<tr>
<td>indefinite</td>
<td>indefinite</td>
<td>S-0, O-0, At-0, Ld-0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
</tbody>
</table>

The table in (3) shows that only Subject and Left-Dislocated preverbal NPs bind referentially-dependent postverbal NPs, and that Referential Dependence accounts for 20% of the cases problematical to the Referentiality Constraint. (4) shows examples of postverbal NPs which are referentially dependent on preverbal ones.

(4a)  
Ni hor boe hor butun.  
1P 1P.have.PST 1S tobacco  
‘We had our tobacco.’  
HBP-2
2.3 Cognitive Status

As noted earlier, definiteness alone is insufficient for determining relative degree of referentiality in the 134 cases where both NPs are definite or both indefinite. In search of a more finely-tuned method for testing the Referentiality Constraint in such cases, we elected to adapt the Givenness Hierarchy framework of Gundel, Hedberg & Zacharski (1993) for application to Breton, and also to try expand it to cover quantificational and predicative NPs in addition to the definite and indefinite NPs discussed in the 1993 paper. The resulting tentative hierarchy is shown in (5).

(5) **Givenness Hierarchy assumed for Breton:**

<table>
<thead>
<tr>
<th>In focus</th>
<th>Activated</th>
<th>Familiar</th>
<th>Identifiable</th>
<th>Referential</th>
<th>Identifiable</th>
<th>Predicative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clitic</td>
<td>indep.pers.pron.</td>
<td>medial det.</td>
<td>def. article</td>
<td>indef.article</td>
<td>pred.bare sg.N</td>
<td></td>
</tr>
<tr>
<td>Agr.suf.</td>
<td>demon.pron.</td>
<td>distal det.</td>
<td>bare sg.poss.</td>
<td>exist.quant.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal det</td>
<td></td>
<td></td>
<td>univ. quant.</td>
<td></td>
<td>neg.pol.item</td>
<td></td>
</tr>
</tbody>
</table>

This Givenness Hierarchy encodes a number of constraints on NP use, which can be explained roughly as follows. Clitics have to be at the center of attention in the discourse context. Independent personal pronouns, demonstrative pronouns, and proximal demonstrative determiner phrases have to have been just- previously-mentioned in the discourse. Medial and distal demonstrative determiner phrases don’t need to be activated but do have to be known to the hearer. In the case of definite article phrases, bare singular possessive constructions and universal quantifiers, the identity of the referent need not be known to the hearer but must be readily inferrable in the discourse context. An NP is referential when the speaker has a specific referent in mind which is not necessarily inferrable to the hearer. Breton does have a small set of reduplicated demonstrative pronouns which are used only for that purpose, but they are too rare to include in (5). Indefinite articles, existential quantifiers and negative polarity items are often used specifically in Breton, but they can also be used nonspecifically. We tentatively added a seventh status to the 1993 hierarchy because Breton predicate nominals don’t always evoke individual entities.
We take the Referentiality Constraint to predict that the NP with higher cognitive status will always occur preverbally in NP-V-NP clauses—that is, the cognitive status of the preverbal NP must be further to the left on the hierarchy than the cognitive status of the postverbal NP. The table in (6) shows that this constraint was met in 88 of the 134 definite-definite and indefinite-indefinite tokens, and that Cognitive Status accounts for a total of 49% of the 181 tokens problematical for the Referentiality Constraint defined solely on the basis of definiteness.

(6) Distribution of Preverbal NPs with Higher Cognitive Status

<table>
<thead>
<tr>
<th>preverbal initial NP</th>
<th>N</th>
<th>%</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite definite</td>
<td>75</td>
<td>63.6</td>
<td>118</td>
</tr>
<tr>
<td>indefinite definite</td>
<td>0</td>
<td>0</td>
<td>47</td>
</tr>
<tr>
<td>indefinite indefinite</td>
<td>13</td>
<td>81.3</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>88</td>
<td>48.6</td>
<td>181</td>
</tr>
</tbody>
</table>

(7) shows some examples of clauses in which the preverbal NP has a higher cognitive status than the postverbal NP.

(7) a. Yâ des grêt berzél pwarzek.
   he have.PRS done warfourteen
   ‘He did the War of 1914’
   MVG Act/Fam

b. Ur bannig e rehe plijadur.
   a drop.DIM PRT give.COND pleasure
   ‘Some would give pleasure.’
   YAB Type/Pred

c. Un dra aralem es goarnet chonj anehon.
   a thing other 1S have.PRS kept thought of-3SM
   ‘I remembered another thing.’
   HBP-2 Ref/Pred

The table in (8) shows that Referential Dependence and Cognitive Status together account for 69%, or 125 tokens, of the 181 cases problematical for the Referentiality Constraint.

(8) NPs Accounted for by Referential Dependence and Cognitive Status:

<table>
<thead>
<tr>
<th>preverbal initial NP</th>
<th>N</th>
<th>%</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite definite</td>
<td>107</td>
<td>90.6</td>
<td>118</td>
</tr>
<tr>
<td>indefinite definite</td>
<td>5</td>
<td>10.6</td>
<td>47</td>
</tr>
<tr>
<td>indefinite indefinite</td>
<td>13</td>
<td>81.2</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>69.1</td>
<td>181</td>
</tr>
</tbody>
</table>

However, the remaining 31% of the data, summarized in (9), cannot be accounted for either by Referential Dependence or Cognitive Status.
(9) **NPs Unaccounted for by Referential Dependence and Cognitive Status**

<table>
<thead>
<tr>
<th>preverbal</th>
<th>postverbal</th>
<th>initial NP</th>
<th>N</th>
<th>%</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite</td>
<td>definite</td>
<td>S-10, O-1, At-1, Ld-0</td>
<td>= 11</td>
<td>9.3</td>
<td>118</td>
</tr>
<tr>
<td>definite</td>
<td>indefinite</td>
<td>S-10, O-17, At-15, Ld-0</td>
<td>= 42</td>
<td>89.3</td>
<td>47</td>
</tr>
<tr>
<td>indefinite</td>
<td>indefinite</td>
<td>S-3, O-0, At-1, Ld-0</td>
<td>= 3</td>
<td>18.7</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>= 56</td>
<td></td>
<td>30.9</td>
<td>181</td>
</tr>
</tbody>
</table>

3. **The Focus Override**

Of the 56 problematical tokens remaining, the 42 clauses with indefinite preverbal and definite postverbal NPs clearly violate the Referentiality Constraint. In all of these cases, our claim is that the preverbal NP marks a focus of one sort or another. Listed and loosely defined in (10) are the primary subtypes of preverbal focus we identified in the texts we examined.

(10) **Subtypes of Focus**

- A constituent question word presents an **interrogative focus**: answers a constituent question
- An information focus: excludes alternatives
- A contrastive focus: presents an important unactivated discourse referent
- A presentational focus: presents interesting, surprising information
- An emphatic focus: presents an all-news ‘what happened’ report, has an extrasentential topic, or presents an aphorism
- An all-comment sentence: presents an all-news ‘what happened’ report, has an extrasentential topic, or presents an aphorism

The table in (11) shows the distribution of preverbal focus types in the 42 indefinite-definite NP-V-NP clauses which blatantly violate the Referentiality Constraint.

(11) **indefinite-definite initial NP**

- Interrogative focus: S-6, O-7, At-1, Ld-0 = 14
- Presentational focus: S-0, O-2, At-0, Ld-0 = 2
- Emphatic focus: S-1, O-8, At-14, Ld-0 = 23
- All-comment sentence: S-3, O-0, At-0, Ld-0 = 3

10 of the 42 tokens have preverbal subjects, as in the emphatic focus and all-comment examples in (12a)--“a SPADE blow pierced my heart”, “a SHORT fight send their defeat to an END”.

17 have preverbal objects, as in the presentational focus and emphatic focus examples in (12b)--“a SISTER had sir Touz in Sainte ANNE”, “and REGRET had we three”. And 15 have preverbal attributes, as in the emphatic and interrogative focus examples in (12c)--“a MISERABLE village is this Karnasen”, “who IS that young girl?”. 

6
a. Un taoi klezen en deus treuzet ma c’halon.  
A spade blow pierced my heart.

b. Ur hoér en doé en Eutru Touz é Santez Anna.  
Sir Touz had a sister in Sainte Anne.

c. Ur gér truhek é Karnasen-man.  
This Karnasen is a miserable village.

The table in (13) shows the distribution of preverbal focus types in the 3 indefinite-indefinite and 11 definite-definite tokens that cannot be accounted for by Referential Dependence or Cognitive Status.

<table>
<thead>
<tr>
<th>Focus Type</th>
<th>initial NP</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>interrogative focus</td>
<td>S-1, O-0, At-1, Ld-0 = 2</td>
<td></td>
</tr>
<tr>
<td>all-comment sentence</td>
<td>S-1, O-0, At-0, Ld-0 = 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>= 3</td>
</tr>
<tr>
<td>definite-definite</td>
<td></td>
<td></td>
</tr>
<tr>
<td>emphatic focus</td>
<td>S-3, O-1, At-1, Ld-0 = 5</td>
<td></td>
</tr>
<tr>
<td>all-comment sentence</td>
<td>S-6, O-0, At-0, Ld-0 = 6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>= 11</td>
</tr>
</tbody>
</table>

Again, all 14 of these preverbal NPs mark focus. 11 are subject-initial, as in the all-comment sentences (14a) and (14b)--“SOMEONE lost a COIN of four RÉALS”, “The thoughts of a CHILD are NOT the thoughts of a young MAN”. 1 is object-initial, the emphatic focus in (14c)--”May that WONderful NaviGAtion continue to make the HAPPiness of a man”.  And 2 are attribute-initial, as in the emphatic focus (14d)--”the most PROSPEROUS of them ALL was this one”.

a. Unan bennag en doé kollet ur pêh a pear real.  
Someone lost a coin of four reals.

b. Sonjoù ar c’hrouadurn’ int ket sonjoù un den yaouank.  
The thoughts of a child are not the thoughts of a young man.
c. Eürusat ur paotr a gendalc’h ar verdeadenn burzhudus-se. 
become-happy a boy PRT continue.PRS the navigation wonderful-DEM 
‘That wonderful navigation continued to make the happiness of a man.’ EKN

d. Er brasan anehé rah oé hennen. 
the biggest of-3P all be.PST this-one
‘This one was the most prosperous of them all.’ HBP-1

All the NPs not accounted for by Referential Dependence and Cognitive Status are thus accounted for by Focus Override.

4. Gaps in the Corpus

Some combinations of preverbal and postverbal NP types were not attested in the corpus. These are summarized in (15).

(15) a. definite-indefinite OVS, AtVS
b. indefinite-definite LdVX
c. indefinite-indefinite OVS
d. definite-definite topic O-V-S
e. indefinite-definite focus S-V-topic O

The combination in (15a), a definite object or attribute followed by an indefinite subject, is not attested in the corpus, although this combination obeys the Referentiality Constraint. It seems to us, however, that sentences of this sort would tend to pose problems for interpretation in that the definite object or attribute would likely be misinterpreted as subject, since definites are more likely than indefinites to express topics, and topics are more likely to be expressed by subjects than by objects or attributes. To avoid such problems in practice, speakers might tend to avoid using these types of sentences.

The combination in (15b), an indefinite left-dislocated NP followed by a definite NP, is also not attested in the corpus. This combination would violate the referentiality constraint. Furthermore, since left-dislocated NPs normally express topics, they are normally definite. However, we did find in the corpus one abnormal instance of a left-dislocated INDEFINITE NP. It was given above in (7c) and was used in its discourse context to mark a presentational focus. A strict translation would be: “ANOTHER THING, I’ve kept thought of it”.

(7) c. Un dra aralem es goarnet chonj anehon. 
a thing other 1S have.PRS kept thought of-3SM
‘I remembered another thing.’ HBP-2 Ref/Pred
It is interesting that this case does not in fact violate the Referentiality Constraint -- the postverbal predicate nominal (meaning ‘thought’) has a lower cognitive status than the left-dislocated referential indefinite.

The combination in (15c), an indefinite object followed by an indefinite subject, is also not attested in the corpus. Here as well, problems for interpreting such sentences would be likely to arise. Because there is no systematic morphosyntactic information available in Breton for identifying grammatical relations (such as case or agreement), such sentences are likely to ambiguous and thus avoided in practice. One example in the corpus did show a preverbal indefinite ATTRIBUTE followed by an indefinite subject, but the preverbal indefinite attribute was an interrogative pronoun in this case. We assume that a preverbal interrogative object followed by an indefinite subject is also a possible combination.

Given what we have said so far, it might be thought that the Referentiality Constraint is a contraint on grammatical relations rather than on preverbal NPs. A Referentiality Constraint on grammatical relations, such as the one formulated in Schapansky 1992, would predict that in a sentence with two definite or two indefinite full NPs, one preverbal and one postverbal, the preverbal NP will be interpreted as subject. However, examples like (16b) are frequently attested in texts.

(16) a. Ha kleuet e vou nezé trouz er velin é valein.
   and heard PRT be.FUT then noise the mill in grind.INF
   ‘And the noise of the mill grinding will then be heard.’

   b. En drouz-sé e hra ol er melinieu ér bed abéh.
   the noise-DEM PRT do.PRS all the mills in.the world entire
   ‘All the mills in the world make that noise.’

(16b) shows a PREVERBAL activated topic OBJECT followed by a definite subject: “And THEN will be heard the noise of the MILL GRINDING. THAT NOISE all the mills in the WORLD make.”

Furthermore, a Referentiality Constraint on grammatical relations would predict that a sentence containing one definite and one indefinite NP, the definite one would be interpreted as subject. The attested example in (17), however, shows that this is not always the case.

(17) Adreist en auter, én diabarh ag er chapel, é touéh penneéled diùachellet,
   above the altar, in.the interior of the chapel, among heads angels unwinged,
   ‘Above the altar, inside the chapel, among the heads of unwinged angels,

   staget a-glei hag a-zeheu, é sellet get deulegad bras ha divergont,
   attached of-left and of-right PRT watch.INF with eyes big and arrogant,
   hanging to the left and to the right, watching with big and arrogant eyes
In (17), we have a series of five stage-setting topics, followed by an INDEFINITE preverbal presentational focus SUBJECT followed by a definite activated topic object.

What these last two examples show is that the Referentiality Constraint is not a constraint on grammatical relations but rather a constraint on preverbal placement of NPs: The most referential, hence the most topical, NP will occur preverbally, as in (16b); unless the preverbal NP is a marked focus, as in (17).

5. Conclusion

To conclude, we have shown that referentiality constrains word order in Breton. In a sentence containing one preverbal and one postverbal NP, the preverbal one will be more referential, unless marked for focus. One issue remains to be addressed. The overwhelming majority of preverbal subjects (245 out of 300) poses a potential problem for our claim that it is referentiality rather than subjecthood which constrains preverbal placement. Breton might be thought INSTEAD to be SVO. However, 16% of the preverbal NPs in our data were NOT subjects, and if Breton WERE an SVO language, interpretation difficulties of the type shown in (14) would not arise. Furthermore, the prevalence of preverbal subjects follows from the referentiality constraint. Since subjects are more likely than objects or attributes to be topics, and topics are likely to be referential, the referentiality constraint predicts that subjects are more likely to occur preverbally. Finally, recall that sentences with two overt NPs represent only a small proportion of word order possibilities in Breton.

References

Breton Texts

* This research is supported by a research grant awarded to Nancy Hedberg by the Social Sciences and Humanities Research Council of Canada. It is also supported by a doctoral fellowship awarded to Nathalie Schapansky by the Social Sciences and Humanities Research Council of Canada.

Abbreviations:
COND = conditional mood; DEM = demonstrative particle; DIM = diminutive suffix; FUT = future tense; INF = infinitive suffix; IPF = imperfect aspect; M = masculine gender; NEG = negative particle; P = plural; PRS = present tense; PRT = sentence particle; PST = past tense; S = singular; 1 = first person; 3 = third person.