A SO AND SO IN MIDDLE AND EARLY MODERN BRETON: A QUANTITATIVE APPROACH

A particular and isolated feature of the 3rd sg. ind. pres. form so ‘is’ in (early) MB is constituted by the fact that so is mostly used without the preverbal particle a, be it in main clauses as in (1a), or in relative clauses as in (1b).

(1a) Bell. 2  
me so biban ‘I am small’

(1b) Catech. 20  
an re so en dout ‘those who are in doubt’

In MoB, on the other hand, the use of the preverbal particle a in this context is compulsory, and it certainly is from this perspective that Hemon (1975, 261) states: “It [sc. so] is as a rule preceded by the particle a, but the particle is sometimes omitted”. From a historical perspective,

(*) I thank Ricardo Scherschel, B.A., for helping me collect the data and Sonja Schnabel, M.A., for improving the English text.
however, taking MB as a starting point, a negative formulation seems more appropriate, cf. Le Roux (1957, 430): “On ne trouve pas a : i” régulièrement en m. br. avec la forme so […] Peu à peu, par analogie, a s’est introduit et se trouve déjà parfois dans les derniers textes du moyen breton”.

No attempt has been made so far to determine the rules that govern the use of the particle a in either single texts or in a collection of texts. The reason for this may be that minimal pairs as in ex. (2) make it very difficult to determine any syntactic and/or structural conditioning. Neither can it yet be ascertained that referential properties have a significant impact on the choice of a so and so.

(2a) M 3001  da nep so aet
     ‘for anyone who has gone’ -
M 1832  da nep a so rebell
     ‘to anyone who is a rebell’
(2b) Bell. 18  An Credo a so necesser euit an Feiz
     ‘the Credo is necessary for faith’ -
Bell. 18  An dec gourchemen so necesser euit an Carantez
     ‘the Ten Commandments are necessary for the love’

With regard to this variation, the purpose of the following article will therefore be to give a quantitative account of the development of a so from its earliest attestations up to the very beginning of the EMoB period.

2. As a matter of fact, a and so never co-occur in Pre-Classic MB prior to the 16th century, but the use of a so rapidly increases during the Classical Middle Breton (CMS) period during the 16th century (†). A so first appears as early as 1519 in the Mirouer and becomes more common in the last quarter of the century, cf. the figures in table 1.

<table>
<thead>
<tr>
<th></th>
<th>N (1500)</th>
<th>M (1519)</th>
<th>J (1530)</th>
<th>TPM (1530)</th>
<th>Ba (1557)</th>
<th>Catech. (1576)</th>
<th>Cath. (1576)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a so</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>39</td>
<td>1</td>
</tr>
<tr>
<td>so</td>
<td>181</td>
<td>236</td>
<td>249</td>
<td>52</td>
<td>195</td>
<td>109</td>
<td>12</td>
</tr>
</tbody>
</table>

*Table 1: Token frequency of a so and so in some CMB texts*

The distribution of so and a so in the most important texts of this period suggests that metrical texts tend to avoid the variant a so. It is remarkable, however, that Iehan an Archer Coz, the author of the Mirouer, already made use of a so eight times. This in turn makes it very probable that the authors of the other texts of this period were familiar with this use of a as well, but maybe refrained from using it for stylistic reasons. The first author to discard the restrictions on the use of a so was Gilles de Kerampuill in his prose texts (Catech. and Cath.), both of which were printed in 1576.

In order to trace the increasing use of a so and related constructions from CMB to EMoB, the occurrence of so in six relevant texts of this time span has been analysed. The following texts have been taken into consideration:

(†) Schrijver 2011, 417.
1. Iehan an Archer Coz. *Le Mirouer de la Mort, en breton*. Frances Cuburiyen, 1519 (date of composition; printed 1575). (Abbr.: M)

The texts were chosen on the basis of three criteria:

1. The texts must be contemporary compositions.
2. The texts have to cover the whole period from approximately 1500 up to the end of MB around 1650; ideally, the texts are evenly spread over the whole period of consideration, leaving no major gaps.
3. The frequency of *so* and *a so* is reasonably high and thus allows for conclusions that are based on quantitative methods (1).

Note that M is the only metrical text in the text sample and that both text 2 and text 3 were written by the same author, viz., Evzen Gueguen. Table 2 provides an overview of the token frequency of *so* in the text sample.

<table>
<thead>
<tr>
<th></th>
<th>M (1519)</th>
<th>Catech. (1576)</th>
<th>Cnf. (1612)</th>
<th>Bell. (1625)</th>
<th>Beach (1656)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens of <em>so</em></td>
<td>244</td>
<td>148</td>
<td>370</td>
<td>174</td>
<td>109</td>
</tr>
</tbody>
</table>

*Table 2: Token frequency in the text sample*

![Graph showing frequency of *so* and *a so* in the text sample](image)

*Figure 1: Frequency of *so* and *a so* in the text sample*

Figure 1 below illustrates the frequency of *a so* in the text sample. Based on these figures the following points can be made: In the 137 years covered by the texts, the use of *a so* increases

---

(1) It should be kept in mind that in several early metrical CMB texts *so* is never preceded by the particle *a*, cf. table 1 above.
considerably. Starting at 3.3% in the Mirouer, it reaches its highest percentage (79.8%) in the latest text of the text sample.

At first glance Cnf. seems to be an exception to the rule stated above. With 17.8% (rectangular dot in figure 1) it makes markedly less use of a so than Catech. (26.4%). Although the decrease of a so is to be regarded as a variety phenomenon (p value > 0.05), it is still striking. Considering the text genre as well as the distribution of so without a in this particular text, one particular phrase influencing the statistics noticeably can be detected. In the formulaic description of sinful deeds, the peccability of an action is always (139 times) expressed using the set phrase X so pechet maruel/veniel ‘X is a Mortal/Venial Sin’; a is never used in this context. We are obviously dealing with a set phrase. When disregarding those set phrases in the calculation, a total of 28.6% of a so is reached, which blends in seamlessly with the findings of the other texts.

It is also striking that the values of Cnf. and Bell. differ strongly from each other (p < 0.001). Despite the fact that both works were written by Euzen Gueguen, a so occurs significantly more often in the younger Bell. than in the Cnf. The author obviously both mirrors and contributes to the increased use of a so.

The particle a is used in main so-clauses as well as in relative so-clauses. If the introduction of a in conjunction with so had happened independently of the clause type, we would expect to find the same frequency of so (without a) as of a so in both clause types. This, however, is not the case. The statistical account of the use of a so in main clauses and relative clauses is displayed in figure 2.

![Figure 2: A so in main clauses and in relative clauses](image)

In the oldest text of our sample, viz. M, main and relative clauses hardly differ as far as the use of a so is concerned. Since the number of tokens of a so in M is quite low (n = 8), this fact must not be overestimated. Starting with Catech., however, a so becomes much more common in relative clauses than in main clauses (factor 2.16, p < 0.05). This difference between main and relative clauses remains quite stable for the next 50 years (cf. Cnf.: factor 1.82, p < 0.05; Bell.: factor 1.58, p < 0.05). By the time of Beach the inclination to use a so mostly in relative clauses
disappears, the main so-clause with a even surpasses the relative clause with a so in this text. The difference between the two clause types, however, is not significant (p > 0.05). Figure 2 illustrates that during its phase of expansion (between M and Beach) a so is significantly more often to be found in relative clauses than in main clauses. The use of a so in main clauses thus follows its use in relative clauses, only catching up where the percentage of a so is already high.

3. While only two different types of main clauses containing so (viz. a so and so) exist, relative clauses feature more types. Apart from relative clauses with so and a so, which cannot be formally distinguished from main clauses, there is one additional type in which the pronoun pebeny/pere 'who (sg./pl.)' is combined with either so or a so (cf. Hemon 1975, 290, Pennaad 1964, 43–44, Le Roux 1957, 429) (9). The pronoun enables a straightforward formal identification as a relative clause of an otherwise ambiguous structure (cf. ex. 3).

\[(3a) \quad M2612 \quad \text{poan you an yfferm pere so aternal} \]
\[\quad \text{‘the pains of hell, which are everlasting’} \]

\[(3b) \quad \text{poan you an yfferm so aternal} \]
\[\quad \text{‘the pains of hell, which are everlasting’} \]

The extension of the relative form so with a is a decisive factor with regard to the identification of so-clauses as relative clauses in Catech., Cnf. and Bell (cf. Figure 2). In addition to that, the introduction of the pronoun adds a further means of reliably marking and identifying a relative construction. The following paragraphs will be concerned with the behaviour of the pronoun pebeny/pere with and without a and/or its contribution to the spreading of a in relative so-clauses.

The pronominal construction for relative so-clauses occurs but sporadically in Nonne, Barbe, Trois Prèmes and Mystère, while it adopts a predominant position among the marked so-relative clauses of M, as figure 3 indicates.

![Figure 3: Frequency of pebeny/pere so in marked relative so-clauses](image-url)

(9) There is another type of pronoun involved in the construction of relative clauses, viz. nep (cf. Hemon 1975, 290), but nep (a) so is far from frequent in our text sample. Outside M it occurs only five times in the Cnf. Nep almost exclusively serves as a pronominal complement followed by a restrictive relative clause as in M3000–01 paeron ter ne vese da nep so aet 'there will be no poverty to someone who has gone'. Nep as a pronominal element inside a relative clause is exceedingly rare, one of the very few examples being M3359 Dog nep so roe 'god, who is king'.

HOR YEZH 270
A comparison of peheny/pere so and the competing structure a so from a quantitative point of view results in a relation as shown in figure 4. The frequency of peheny/pere so constantly decreases over the course of time. A so, on the other hand, is subject to a (non-linear) overall increase; the remarkable temporary increase in Catech. is particularly striking. In this text the frequency of peheny/pere so and a so differs only slightly. From Cnf. and Bell. onwards, the texts written by Euzen Gueguen, the use of a so increases again.

![Graph showing frequency of peheny/pere so and a so in marked relative so-clauses.]

**Figure 4:** Frequency of peheny/pere so and a so in marked relative so-clauses

In figure 5 the use of the third option, viz. peheny/pere a so, is added to the two constructions mentioned before.

![Graph showing frequency of peheny/pere so, a so and peheny/pere a so in marked relative so-clauses.]

**Figure 5:** Frequency of peheny/pere so, a so and peheny/pere a so in marked relative so-clauses

The run of the frequency curve of peheny/pere a so shows two different patterns: In M and Catech. the number of appearances of peheny/pere a so is negligible, but the construction
dominates statistically from Cnf onwards with a fraction of about 60%. The increase of the mixed construction combining pronoun and particle correlates directly with the decrement of the pronominal construction (pebeny/pere so). Comparing the development by text the following can be stated:

1. M (n = 19): Marked relative so-clauses using pronominal pebeny/pere so are prevalent. A so remains far behind the pronominal construction and the combined pebeny/pere a so construction occurs only marginally.

2. Catech. (n = 54): The pebeny/pere a so construction remains quite limited in Catech., too. The marking of relative so-clauses using a temporarily reaches a level above 50%. Comparing the curves of M and Catech., we see that the portion of a so has increased to the same extent to which pronominal pebeny/pere so has diminished.

3. Cnf. (n = 97): From the Cnf on, the mixed construction (combining pronoun and particle) pebeny/pere a so dominates. In comparison to Catech., it mostly replaces a so which falls back to a remarkably low level. Thus, one may we might say that a so and pebeny/pere so, which were of equal frequency in Catech., have conspired at this stage to become the dominant pattern in the construction of relative so-clauses. The rate of plain pronominal pebeny/pere so continues to decrease, but it does not essentially differ from pebeny/pere a so in terms of statistics (p > 0.05).

4. Bell. (n = 81): The percentage of pebeny/pere a so reaches its highest level in Bell. As pronominal pebeny/pere so continues to decline, the use of a so increases again, obviously compensating to some degree for the vanishing pronominal pebeny/pere so.

5. Beach (n = 28): In this text, which was written at the beginning of the EmoB period, two constructions were prevalent, namely the combined pebeny/pere a so construction and the plain a so construction. The main point of interest here is that by this time in the history of Breton, the plain pronominal construction has become obsolete. Since the difference in the number of occurrences of a so and pebeny/pere a so is not significant in terms of statistics, pebeny/pere so would seem to occur as an optional lexical enlargement of a so. The possibility that the use of pebeny/pere is triggered by e.g. referential values can not be excluded, though. Only a comprehensive investigation of pebeny/pere in all types of relative clauses will help to clarify the role the pronoun plays.

According to the evaluated MB corpus, relative clauses thus played a fundamental role in the spreading of the preverbal particle a in so-clauses. During the decisive span of time, a so occurs significantly more often in relative clauses than in main clauses. The number of occurrences of a so in main clauses only catches up at the beginning of the EMoB period.

4. As figure 1 shows, the use of a so from CMB to EMoB increases constantly and it is therefore tempting, to make a general assumption about the relation of the number of occurrences of a so in each text and its period of composition. This would, in turn, provide us with a criterion for dating texts which otherwise lack information with regard to their time of composition. Using this “a so-value” would allow us to ascribe a given text to a specific point in time. Even though this method can only provide an approximate date and moreover has to be handled very carefully, its benefit might be worthwhile.

The Novelou ancien ba devot may serve as a test object for this approach. The Novelou survive in an edition dated to 1650 (Tanguy Gueguen 1650). According to its title page, this edition is partly based on an older one that has been amended, corrected and augmented "amantet, corriget,
by the 1650 editor Tanguy Gueguen. Based on this evidence, the Novelou can be ascribed to the same period covered by our text sample. A more precise dating of the text may be ventured by using the a so-value as a dating criterion. The a so-value of the Novelou is 12.2% (n = 115), which immediately makes clear that it is indeed most unlikely that the text was written in the 17th century. If we take the a so-value of the Novelou seriously and compare it with the values of our text sample, a date of composition at some point between M and Catech. seems realistic, cf. figure 6.

![Figure 6: Dating of Novelou according to its a so-value](image)

This method can be extended to other texts whose date of composition is even less certain. In G, for instance, only one out of 106 so-clauses is preceded by a (G 146 (1)), which makes it very likely that this metrical text, transmitted in Le Pelletiers manuscript dictionary in the early 18th century, was composed in the first half of the 16th century. At the end of one of the copies which Le Pelletier had at his disposition, 1580 is given as the year of its origin (Widmer & Jørgensen 2011, 9). This date coincides fairly well with its a so-value and suggests at the same time that the 1580 text of Le Pelletier is probably a copy of an even older manuscript (or print).

To give another example, the a so-value of the Trajediën Santex Jenouvefa a Vrabant (JV, an Dug 2008) is 34.7% (n = 276). This value roughly points to a date of composition in the first half of the 17th century.

It is clear that on its own this method only allows approximative dating, and it must be used with due care, ideally in combination with other methods.

5. The results can be summed up as follows. In MB, the relative clauses played a decisive role in establishing the particle a in so-clauses: a occurs significantly more often in relative clauses than in main clauses. The development in relative clauses is closely related to the tendency to overtly mark relative so-clauses already in M. In the earliest text, the pronominal construction using pebeny/pere clearly prevailed, but constantly loses ground until it becomes completely marginalised by the end of the investigated period. The mixed construction combining both the pronoun pebeny/pere and the particle a, quickly exceeds pebeny/pere so and a so used on their

(1) Most interestingly, so doesn’t fit into this line’s rhyme scheme (naireutz, saur-goest) and the whole stanza, which is quite obscure, seems to be an addition in the presumably younger B copy (Widmer und Jørgensen 2011, 144-45).
own. By the beginning of the EMoB period, peheny/pere has disappeared as a relative pronoun on its own, instead occurring exclusively in combination with a. In fact, peheny/pere seems to have become some kind of lexical enlargement in the particle-only construction a so, which is the only alternative to the mixed construction in the latest text. Whether or not the use of peheny/pere at any stage is random or whether it is determined by some other factors (e.g. referential properties) is not clear. Only a comprehensive investigation of all relative clauses can shed light on this issue.

Bibliography
Gueguen, Tanguy 1650. An novelou ancien, ha devot, an oll amantet, corriget, bac augmentet, à vn nóbr re neuez quen Brezonec, ha Gallec. Qvempercavrentin: George Allienne.