Syntactically marked Switch-Reference
Evidence from German

Katja Barnickel

This paper makes two claims: first I will show that languages have the strategy to mark switch-reference (SR) syntactically. This finding is interesting because traditionally switch-reference marking has been considered to be a typically morphological phenomenon. I provide evidence from a corpus-based study that in German SR is marked by the order of the matrix clause and the subordinate clause (linearization). Secondly I account for the observation that SR marking in German does not seem to be a fully grammaticalized phenomenon, but that nevertheless it is reflected in clear statistical tendencies. I provide a modeling of the statistically preferred option to mark SR in the framework of Stochastic Optimality Theory.

1. Introduction

The term switch-reference (SR) refers to the presence of a verbal marker that indicates if the subject of a clause is coreferent with the subject of an adjacent clause within the same syntactic structure. Coreference is marked by a same-subject-marker (SS) (see (1)), whereas non-coreference is marked by different-subject-marker (DS) (see (2)).

(1) Mojave (Munro 1979:145)
   a. nya-isvar-k iima-k
     when-sing-SS dance-Tns
     ‘When he; sang, he; danced.’
   b. nya-isvar-m iima-k
     when-sing-DS dance-Tns
     ‘When he; sang, he; danced.’

The phenomenon of switch-reference was first described by Jacobsen (1967) and is primarily found in North American and Papuan languages. It also appears in South America, in the Caucasus and in East Asia, but it seems to be restricted to languages with the word order OV.
As it can be seen in (1) and (2), SS/DS-marker always show up at the verb and they differ morphologically.\(^1\)

Although morphological switch-reference marking typically does not arise in Indo-European languages, it has been found that SS and DS contexts are encoded within these languages as well. That the encoding of switch-reference does not have to be obligatorily morphological can be seen in (2). In Latin switch-reference is marked syntactically: the ablative in ablative-absolutive-constructions marks DS contexts.

(2) **Latin (Haiman 1983:117)**

a. *Same subject*

   Aristides\(_i\) [\(CP\ pro\ patria pulsus\] Lacedaemonium fugit.

   ‘Being expelled from his home country, Aristides escaped to Lakedaimon.’

b. *Different subject*

   [\(CP\ Aristide\(_i\) patria pulso\] Persae\(_j\) Graecos agressi sunt.

   ‘When Aristides was expelled from his home country, the Persian attacked the Greek.’

In Standard German switch-reference does not seem to be a grammaticalized phenomenon, what means that a reference change within a subordinate clause does not have to be marked obligatorily.\(^2\) However, the observation that in Latin ablative constructions serve as DS/SS markers, provides evidence that Indo-European languages do have a possibility to mark switch-reference non-morphologically but syntactically.

Based on this observation I will show that German, which apparently does not have verbal switch-reference markers, marks the change of subject by structural variation. This structural variation is the linearization of matrix clause and subordinate clause. In German a subordinate clause can follow its matrix clause (\(CP\)\(_{mtx}\) - \(CP\)\(_{sub}\), see (3)) or it can precede it (\(CP\)\(_{sub}\) - \(CP\)\(_{mtx}\), see (4)).

(3) **\(CP\)\(_{mtx}\) - \(CP\)\(_{sub}\)**

a. [\(CP\)\(_{mtx}\) *Der Mann* \(_i\) sah die Frau\(_j\)] [\(CP\)\(_{sub}\) als *er*\(_j\) sich umdrehte].

   ‘The man\(_i\) sah the woman, when he\(_j\) himself turned’

b. [\(CP\)\(_{mtx}\) *Der Mann* \(_i\) sah die Frau\(_j\)] [\(CP\)\(_{sub}\) als *sie*\(_j\) sich umdrehte].

   ‘The man\(_i\) sah the woman, when she\(_j\) herself turned’

---


\(^2\)It has been assumed that switch-reference does not only occur in subordinate clauses, but in coordinated sentences as well (see Weisser 2012) or that switch-reference affects coordinated structures only (see Keine 2012). Concerning German, I will only deal with subordinated clauses.
The subjects in (3-a) and (4-a) are coreferent. In (3-a) we have anaphoric reference, whereas in (4-a) it is cataphoric reference. In (3-b) and (4-b) in contrast there is no coreference relation between the subjects.

In the following, I will show that German exhibits the tendency to mark switch-reference by the order of linearization of the matrix clause and the subordinate clause. If there is a coreference relation between the subjects of two adjacent clauses, the subordinate clause prefers to follow its matrix clause, whereas in the case of different subjects, the subordinate clause prefers to precede its matrix clause. Hence, I claim that the precedence of the subordinate clause corresponds to a (non-obligatory) DS-marker, whereas a following subordinate clause equals a SS-marker.

This claim is based on a corpus study which is presented in section 2 and discussed in section 3. Section 4 provides previous implementations of similar statistical preferences within the framework of stochastic Optimality Theory. Concluding, I will show how the tendency in German to mark switch-reference by the linearization of the matrix and the subordinate clause can be modeled in stochastic Optimality Theory.

2. Corpus study

2.1. Method

Using the corpus search system COSMAS II, I examined the subject coreference in 300 sentences, taken from the public corpora of the IDS Mannheim. I only considered single embedded sentences, in detail that means that all investigated sentences consist of a matrix CP and a subordinate CP.\(^3\) The abstract structure of all items is given in (5).

\[(5) \quad \text{a. } [[\text{CP}_{\text{sub}} \text{SUBJUNCTION PRONOUN} ...,] \ [\text{CP}_{\text{mtx}} \ldots \text{(non-)} \text{antecedent DP} \ldots]]
\]
\[
\text{b. } [[\text{CP}_{\text{mtx}} \ldots \text{(non-)} \text{antecedent DP} ,] \ [\text{CP}_{\text{sub}} \text{SUBJUNCTION PRONOUN} ...,]]
\]

Subjunctions and pronouns were varied the following way:

\[(6) \quad \text{a. } \text{WÄHREND} + \text{ER} \ ‘\text{while} + \text{he}’
\]
\[
\text{b. } \text{WÄHREND} + \text{SIE} \ ‘\text{while} + \text{she}’
\]
\[
\text{c. } \text{ALS} + \text{ER} \ ‘\text{when} + \text{he}’
\]
\[
\text{d. } \text{ALS} + \text{SIE} \ ‘\text{when} + \text{she}’
\]

\(^3\)To provide a clear data set, multiple embedded sentences which introduce a further subject, as well as sentences including direct speech were not included in the study.
75 sentences per combination were chosen at random from the corpora. (7) provides four examples from the study.

(7) a. \[CP_{mtx} \text{ Dadurch erstickte der Fahrer, während er schlief.} \] thereby suffocated the driver, while he slept

‘Thereby the driver suffocated while he was sleeping.’

b. \[CP_{sub} Während sie rührt, CP_{mtx} \text{ telefoniere ich im Wohnzimmer weiter.} \]

while she mixes telephone I in the living-room onward

‘While she mixes something up, I continue doing a call in the living room.’

c. \[CP_{sub} Als er das Haus verließ, CP_{mtx} \text{ fuhr die Polizei auf.} \]

when he the house left, came the police

‘When he left the house, the police came.’

d. \[CP_{mtx} Weltmeisterin Andrea Henkel, hatte Pech, CP_{sub} als sie sich bei einem Sturz den Daumen brach. \]

worldchampion Andrea Henkel had bad luck, when she herself by a fall the thumb broke.

‘It was bad luck for worldchampion Andrea Henkel, when she broke her thumb as she fell.’

2.2. Results

Two of the 300 taken sentences were excluded from the analysis because of unclear reference relations. The remaining 298 valid sentences were the base for a statistical analysis (chi-square-test) for the factors reference (DS/SS) and the position of the subordinate clause (preceding matrix clause/following matrix clause). Table 1 shows that in German there are more SS-sentences (181 in total) than DS-sentences (117 in total).

Furthermore, the linearization \(CP_{mtx}CP_{sub}\) (163 in total) is more common than \(CP_{sub}CP_{mtx}\) (135 in total). That means there is a general preference for subordinate clauses to follow their matrix clauses instead of preceding them. In spite of these general preferences, a highly significant interaction between sentence linearization and referential identity was found.

The referential identity differed for the two possibilities of linearization: \(\chi^2(1, N = 298) = 27.47, p < 0.001\). According to this result, a random distribution of sentence linearization and referential identity can be excluded. Instead, it leaves the following picture: SS-sentences prefer the order \(CP_{mtx}CP_{sub}\), which means that in the case of coreferent subjects, the subordinate clause prefers to follow the matrix clause highly significantly. In contrast, DS-sentences prefer the linearization \(CP_{sub}CP_{mtx}\). This means that if there is no coreference relation between the subjects, the subordinate clause shows the highly significant tendency to precede its matrix clause. These results are summarized in Table 1 and illustrated in Figure 1.
Moreover, it was found that the two subjunctions während ‘while’ and als ‘when’ behave differently with respect to their general preference of linearizing the subordinate clause ($X^2(1, N = 298) = 13.78, p < 0.001$). Während prefers the linearization CP$_{sub}$ - CP$_{mtx}$, whereas als prefers CP$_{mtx}$ - CP$_{sub}$. This is shown in table 2.

Table 1: Total distribution of referential identity and clause linearization ($\chi^2(1, N = 298) = 27.47, p < 0.001$)

<table>
<thead>
<tr>
<th>Reference</th>
<th>DS</th>
<th>SS</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>75</td>
<td>60</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>42</td>
<td>121</td>
<td>181</td>
</tr>
</tbody>
</table>

| total     | 135 | 163 | 298   |

Figure 1: Total distribution ($\chi^2(1, N = 298) = 27.47, p < 0.001$)
Table 2: Distribution of während and als (position) ($\chi^2(1, N = 298) = 13.78, p < 0.001$)

Additionally, both subjunctions differ concerning the factor referential identity (SS vs. DS): ($\chi^2(1, N = 298) = 7.96, p < 0.01$). Als co-occurs more often in SS-sentences than in DS-sentences (Table 3).

Table 3: Distribution of während and als (reference) ($\chi^2(1, N = 298) = 7.96, p < 0.01$)

Despite these general preferences, for both subjunctions there is an interaction between the position of the subordinate clause and the referential identity of the subjects. For the während-sentences (Table 4), as well as for als-sentences (Table 5), in the context of two different subjects (DS), the subordinate clause prefers to precede its matrix clause. In SS-sentences it is exactly the other way around: independently of the subjunction, subordinate clauses prefer to follow their matrix clauses.

Table 4: Distribution of während (reference/position) ($\chi^2(1, N = 148) = 8.41, p < 0.01$)
Figure 2: Distribution of während ($\chi^2(1, N = 148) = 8.41, p < 0.01$)

<table>
<thead>
<tr>
<th>“als”</th>
<th>Position of $\text{CP}_{\text{sub}}$</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>preceding $\text{CP}_{\text{mtx}}$</td>
<td>following $\text{CP}_{\text{mtx}}$</td>
</tr>
<tr>
<td>Reference</td>
<td>DS</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>SS</td>
<td>25</td>
</tr>
<tr>
<td>total</td>
<td>52</td>
<td>98</td>
</tr>
</tbody>
</table>

Table 5: Distribution of als (position/reference): ($\chi^2(1, N = 150) = 15.68, p < 0.001$)

Figure 3: Distribution of als ($\chi^2(1, N = 150) = 15.68, p < 0.001$)
No significant interaction was found for the two examined pronouns *er* ‘he’ and *sie* ‘she’. Neither *er* nor *sie* prefers one of the orders CP<sub>mtx</sub> - CP<sub>sub</sub>/CP<sub>sub</sub> - CP<sub>mtx</sub> ($X^2(1, N = 298) = 0.34, p= 0.56$) or one of the referential identity contexts SS/DS ($X^2(1, N = 298) = 0.35, p= 0.55$).

3. **Discussion**

3.1. Behaviour of *während* and *als*

The results of the present corpus study show that in case of identical reference of the subject in a matrix clause and its subordinate clause (SS), the subordinate clause prefers to follow its matrix clause, whereas in the case of non-identical subjects (DS) the subordinate clause prefers to precede its matrix clause. This observation is valid for subordinate clauses introduced by *während* as well as for those introduced by *als*, although in general the two subjunctions prefer different linearizations. Subordinated clauses introduced by *während* generally prefer to precede their matrix clauses. This could be due to the fact that *während* besides its temporal meaning (8) also has an adversative interpretation (9).

(8) [CP<sub>sub</sub> Während sie kocht,] [CP<sub>mtx</sub> läuft das Radio in der Küche.] While she cooks is.on the radio in the kitchen ‘While she is cooking, the radio is on in the kitchen.’ (temporal)

(9) [CP<sub>sub</sub> Während sie früher joggte,] [CP<sub>mtx</sub> gilt ihre Leidenschaft heute dem Radfahren.] while she in.the.past jogged holds her passion nowadays the cycling ‘While in the past she used to do jogging, nowadays her passion is cycling.’ (adversative)

Adversative contexts, like in (9), mostly involve the presence of two different subjects, which are contrasted. Under the assumption that preceding subordinate clauses mark a subject change, it is not surprising that in the context of *während*, which is predestined for expressing contrasts, subordinate clauses prefer to precede their matrix clauses. Despite the general preceding preference of subordinate clauses with *während*, they prefer to follow their matrix clauses if referential identity of the two subjects (SS) is given.

Remarkably, the same pattern holds for subordinate clauses with *als*, just the other way around: only referring to the position, *als*-sentences generally prefer to follow their matrix clauses. In the case of two different subjects (DS), however they give up their preferred pattern and show up the tendency to precede the matrix clause.

The fact that the two different pronouns *er* ‘he’ and *sie* ‘she’ do not show any statistical preference for a certain linearization pattern confirms that the interaction of clause linearization and referential identity is not due to the presence of a certain pronoun (as expected).
3.2. Dispreferred SR-marking

As has been shown so far, in German there is the tendency to linearize the subordinate clause and the matrix clause according to the referential relation of the subjects: SS-sentences prefer the linearization $CP_{mtx} - CP_{sub}$, DS-sentences prefer the reversed order $CP_{sub} - CP_{mtx}$. These findings are summarized in (10) and (11).

(10) Preferred linearization

a. $[CP_{mtx} \text{ Der Mann},_i \text{ sah die Frau},_j \text{ sich umdrehte}]$. \\
   ‘The man$_i$ saw the woman$_j$ when he$_i$ turned around.’ (SS: $CP_{mtx} - CP_{sub}$)

b. $[CP_{sub} \text{ Als sie},_j \text{ sich umdrehte},] [CP_{mtx} \text{ sah der Mann},_i \text{ die Frau},_j]$. \\
   ‘When she$_j$ turned around, the man$_i$ saw the woman$_j$.’ (DS: $CP_{sub} - CP_{mtx}$)

(11) Dispreferred linearization

a. $[CP_{sub} \text{ Als er},_i \text{ sich umdrehte},] [CP_{mtx} \text{ sah der Mann},_i \text{ die Frau},_j]$. \\
   ‘When he$_i$ turned around, the man$_i$ saw the woman$_j$.’ (SS: $CP_{sub} - CP_{mtx}$)

b. $[CP_{mtx} \text{ Der Mann},_i \text{ sah die Frau},_j \text{ als sie},_j \text{ sich umdrehte}.]$. \\
   ‘The man$_i$ saw the woman$_j$, when she$_j$ turned around.’ (DS: $CP_{mtx} - CP_{sub}$)

The general dispreference for cataphoric sentences (11-a) in comparison to anaphoric ones (10-a) has already been described in the literature. De Beaugrande & Dressler (1981) for example mention that cataphora is used less often than anaphora because they are more difficult to process. The higher processing effort, however leads to a more attentive or ‘deeper’ processing (de Beaugrande & Dressler 1981:65).

Hence, processing difficulties could be due to the fact that normally in SS-sentences the anaphoric linearization is the preferred one (10-a). The precedence of the subordinate clause normally indicates a reference change (10-b), which does not occur in cataphoric sentences (11-a). Therefore it could be the case that speakers make use of cataphoric sentences to increase the attention level of their hearers/readers by choosing the preferred linearization for DS-sentences without intending a reference change.

Conversely, in DS-sentences the processing of the dispreferred linearization (subordinate clause following matrix clause, see (11-b) should be more difficult to process because it is not the expected linearization when the subject reference changes. It is an outstanding issue to test these predictions in experiments.

Interestingly, unexpected SR-marking (DS-marker in SS contexts or SS-marker in DS contexts) is a quite common phenomenon in typical SR-marking languages (which mark SR morphologically) as well. Roberts (1988a) provides examples from Amele which are DS-marked despite of the fact that the subjects are coreferent.
The examples in (12) show that SR-markers obviously are not only used to indicate a reference change, but also to mark contrasts concerning other aspects of a discourse. These aspects could be “foregrounded versus backgrounded events, same-place-setting versus different-place setting, same-time-setting versus different-time-setting and same-world-setting versus different-world-setting” (Roberts 1997:190).

According to this, the presence of the DS-marker co in spite of coreferent subjects in (12-a) is licensed because in the sentence two different time-settings are contrasted. In (12-b) the co-marking indicates a change of places. Taken more generally the DS-marker in Amele appears in SS-sentences if “a surprise change” (Roberts 1988b) or “some unexpected turn in the narrated events” (Stirling 1993) takes place. Such discourse related SR-markings, which indicate contrast in the widest sense, are not a special property of Amele, but have been observed in other SR-marking languages as well, for example in Yakunytjatjara (see Goddard 1985) and Pitjantjatjara (see Bowe 1990), in Dani (see Bromely 1981) as well as in Nankina (see Spaulding 1988).4

4. Modeling of German SR-marking tendencies in Stochastic Optimality Theory

Switch-reference is not the only phenomenon which in some languages is grammaticalized and in others is reflected in statistical tendencies. Bresnan et al. (2001) discuss the interaction between the person hierarchy (1st, 2nd > 3rd) and voice which is a grammaticalized interaction in Lummi. In Lummi the person of the subject argument cannot be lower than the person of a non-subject argument, a so-called Silverstein effect (see Hale 1973; Silverstein 1976; Aissen 1999). If this would happen in the active, passivization is obligatory. Hence, in Lummi the sentence The man sees me is ungrammatical and must be realised as I am seen by the man even though the ordinary information structural trigger for passive is not present. In contrast, if in the active the person hierarchy is fulfilled like in I see the man, passivization is not possible.

What Bresnan et al. (2001) found in a corpus study was that in English there is the tendency to mark the person/voice-interaction as well. This means that, in English, passive sentences which fulfill the person hierarchy (I was attacked by a man) are more common than passive sentences which violate it (Mary was attacked by me).

Bresnan & Aissen (2002) remark that classical generative theories of formal grammar cannot account for the fact that a grammaticalized phenomenon in one language shows up as a clear

---

4For a detailed discussion of SR-markings related to weather verbs and impersonal constructions see Roberts (1987) and Roberts (2001).
statistical preference in another language: “On these theories, frequentistic processes (such as the conventionalization of usage preferences) must belong either to grammar-external ‘performance’ along with speech errors and memory limitations, or to external choices among competing dialect grammars. Yet neither of these alternatives is an adequate model of variation and change, as first pointed out by Weinreich et al. (1968)” (Bresnan & Aissen 2002:2).

A framework where statistical preferences can be modeled is stochastic Optimality Theory (Anttila 1997; Boersma & Hayes 2001; Hayes 2001). According to stochastic OT, grammaticalized patterns and their usage by trend are not two different things, but merely different points on a continuous scale: “Stochastic OT grammars allow us to place the person/voice interactions in English and Lummi at points on a continuum of conventionalization that connects frequentistic preferences in usage to categorical grammatical constraints” (Bresnan et al. 2001).5

In contrast to a classical OT approach (Prince & Smolensky 1993), the basic idea of stochastic OT is that constraints are not necessarily categorically ordered with respect to each other. Rather, their application domains may overlap. An overlap of application domains gives rise to optionality. Categorical vs. overlapping application domains of constraints are illustrated in (13) and (14): the boxes signal the possible application domains of two constraints on an abstract continuum from “more strict” to “less strict” (interpreted from left to right).

(13)  **Categorical order of application domains (grammaticalization)**

```
  C1
  ┌───────┐
  │      │
  └───────┘
    C2
```

(14)  **Overlapping order of application domains (optionality)**

```
  C1
  ┌───────┐
  │      │
  └───────┘
    C2
```

A candidate is evaluated at an evaluation time. It is well formed if it is optimal at that point. For an evaluation, an arbitrary point (indicated with 1 and 2) is chosen in the application domain of a constraint. If we have an overlapping configuration these points can be outside of the overlapping domain (15) or within the overlapping domain (16).

---

5For a detailed discussion and analysis of the person hierarchy/voice-interaction see Bresnan et al. (2001).
A constraint $C_1$ is ranked higher than another constraint $C_2$ at a given evaluation time if the point chosen for $C_1$ is above the point chosen for $C_2$. If the domains of $C_1$ and $C_2$ are categorically ordered, like in (13), then the point for $C_1$ is always going to be above the point for $C_2$. In this case we deal with a grammaticalized phenomenon without any optionality. However, if the domains of $C_1$ and $C_2$ overlap, optionality arises. The winning candidate is determined by whether the point chosen for $C_1$ is above the point chosen for $C_2$ or vice versa. The choice of an evaluation point at a given evaluation time is free as such. However, the smaller the common domain of $C_1$ and $C_2$ is, the more likely it is that the point chosen for the higher ranked constraint $C_1$ is above the point chosen for the lower-ranked constraint $C_2$ (like in (15)).

This way the system captures statistical preferences. The winning candidate depends on the setting of the evaluation point at a given evaluation time. The scenario like in (15), where both points are outside the overlapping domain, leads to the statistically preferred typical result $C_1 \gg C_2$. The same pattern results from the configuration where only one of the points is outside the overlapping domain. However, if we have the picture in (16), which is the statistically rare but indeed possible case, the point of $C_2$ is above the point of $C_1$. This leads to the result $C_2 \gg C_1$, which means that the candidate favoured by $C_2$ is going to be preferred over the candidate favoured by $C_1$.

In the following I will model the preference of German to mark switch-reference by the linearization of subordinate and matrix clause within the stochastic OT framework. I assume the following constraints:

17. a. $SS_{CP_{mtx}-CP_{sub}}$
   If the subject of a clause is coreferent with the subject of an adjacent clause within the same syntactic structure, the subordinate clause follows the matrix clause.

b. $DS_{CP_{sub}-CP_{mtx}}$
   If the subject of a clause is different from the subject of an adjacent clause within the same syntactic structure, the subordinate clause precedes the matrix clause.
But, as we have seen, these linearizations are not fully grammaticalized, so there must be further factors that influence the linearization of clauses. One obvious factor is information structure, hence I assume competing pragmatic constraints which prefer the contrary linearizations. In (18) I give some examples of possible competing constraints related to information structure.

(18)  a. **TOPIC/FOCUS FIRST**
   The clause which contains the topic or the focus of the whole sentence has to be linearized first.

   Example 1:
   A: What did the man see?  
   B: The man saw the woman, when she turned around.

   Example 2:
   There was a man on a party who was looking for a woman with green glasses. The man saw the woman, when she turned around.

b. **PRONOUN FIRST**
   Cataphoric sentences are harder to process than anaphoric ones (Kennison et al. 2009). To increase the attention of the listener/reader, linearize the clause with the pronoun first.

   Example:
   You won’t believe it - when he turned around, the man saw the woman.

There might be even more such constraints. I summarize them all under *information structural reasons* (ISR).
   
The fact that switch-reference marking by clause linearization is not fully grammaticalized suggests that in German the switch-reference constraints and the information structural constraints are not ordered categorically to each other. Assuming instead that the application domains of the constraints which favour switch-reference and those which favour information structure overlap, we get four possible cases. (19) illustrates the case, where both evaluation points are outside the overlapping domain, whereas in (20) one point is inside the overlapping domain.

(19)  Case 1: $SS_{CPmtx−CP_{sub}}/DS_{CP_{sub}−CPmtx} \triangleright ISR$
Cases 1-3: \( SS_{CP_{mtx}−CP_{sub}}/DS_{CP_{sub}−CP_{mtx}} \) is higher ranked than ISR

a. \( SS_{CP_{mtx}−CP_{sub}} \) is higher ranked than ISR

<table>
<thead>
<tr>
<th>Candidates</th>
<th>( SS_{CP_{mtx}−CP_{sub}} )</th>
<th>ISR</th>
</tr>
</thead>
<tbody>
<tr>
<td>( C_1: ) [When he turned around,]</td>
<td>[the man saw the woman.]</td>
<td>*!</td>
</tr>
<tr>
<td>( C_2: ) [The man saw the woman,]</td>
<td>[when he turned around.]</td>
<td>*</td>
</tr>
</tbody>
</table>

For the sake of convenience the following tableaux only contain the English translations of the German examples (for the fully glossed German sentences see examples (3) and (4)).
b. \( \text{DS}_{\text{CPsub} \rightarrow \text{CPmtx}} \rightarrow \text{ISR} \)

<table>
<thead>
<tr>
<th>Candidates</th>
<th>( \text{DS}_{\text{CPsub} \rightarrow \text{CPmtx}} )</th>
<th>ISR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C}_1$: [When she turned around.,] [the man saw the woman.]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>$\text{C}_2$: [The man saw the woman.,] [when she turned around.]</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

Given the rare but possible configuration in (16), where both points are within the overlapping domain and the point for ISR is above the point for SS/DS, the ISR-constraint now is higher ranked than the SS/DS-constraints. The dispreferred candidate of the SS/DS-constraint is the winner, as shown in (24).

(24) ISR is ranked higher than \( \text{SS}_{\text{CPmtx} \rightarrow \text{CPsub}} / \text{DS}_{\text{CPsub} \rightarrow \text{CPmtx}} \)

a. ISR \( \rightarrow \text{SS}_{\text{CPmtx} \rightarrow \text{CPsub}} \)

<table>
<thead>
<tr>
<th>Candidates</th>
<th>ISR</th>
<th>( \text{SS}_{\text{CPmtx} \rightarrow \text{CPsub}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C}_1$: [When he turned around.,] [the man saw the woman.]</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>$\text{C}_2$: [The man saw the woman.,] [when he turned around.]</td>
<td></td>
<td>*!</td>
</tr>
</tbody>
</table>

b. ISR \( \rightarrow \text{DS}_{\text{CPsub} \rightarrow \text{CPmtx}} \)

<table>
<thead>
<tr>
<th>Candidates</th>
<th>ISR</th>
<th>( \text{DS}_{\text{CPsub} \rightarrow \text{CPmtx}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{C}_1$: [When she turned around.,] [the man saw the woman.]</td>
<td></td>
<td>*!</td>
</tr>
<tr>
<td>$\text{C}_2$: [The man saw the woman.,] [when she turned around.]</td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

Summarizing this section, I have shown that in German the constraint which favours SR-marking is not categorically ordered with respect to constraints concerning information structure. Rather their application domains overlap. This overlapping configuration gives rise to optionality. Given that the SR constraint is still the higher ranked one in the overlapping configuration, it is statistically more likely that candidates favoured by the SS/DS-constraint are the winning ones. In contrast, it is statistically less likely that the candidate favoured by information structural reasons wins the competition (but possible). This explains the result of my study that SR-marking in German is not a fully grammaticalized phenomenon but that its marking is reflected as a strong tendency in corpora.
5. Concluding remark: Grammaticalization vs. Statistical tendencies

As Bresnan & Aissen (2002) have shown, the same categorical phenomena which are attributed to hard grammatical constraints in some languages continue to show up as statistical preferences in other languages. This raises the question if there are languages where the interaction of linear clause order and switch-reference is grammaticalized. The answer is: yes. Kazenin & Testelets (2004) describe an interesting interaction between the order of a converb clause and its matrix clause in Tsakhur, a Nakh-Dagestanian (or East Caucasian) language spoken by 30,000 people in Daghestan (north Eastern Caucasus, Russia) and in northern Azerbaijan.

In Tsakhur, the converb clause standardly precedes or follows its matrix clause (25). In SS-sentences the converb clause can also be center-embedded within the matrix clause (26).

(25) Standard order (SS)

\[
\begin{array}{c}
\text{[CP}\text{ }\text{mtx} \text{ }\text{Rasul} \text{ }\text{ark’in-na},] \\
\text{[CP}_\text{con} \text{ }\text{ma-n} \text{ }\text{žuwab} \text{ }\text{iwho]} \\
\text{Rasul(1CL) leave.PFC.1CL-ATR.1CL this-ATR.4CL word(4CL) say.PFC} \\
\text{‘Having said this word, Rasul left.’}
\end{array}
\]

(26) Center-embedding possible (SS)

\[
\begin{array}{c}
\text{[CP}_\text{mtx} \text{ }\text{Rasul, } \text{[CP}_\text{con} \text{ }\text{ma-n} \text{ }\text{žuwab} \text{ }\text{iwho}, \text{ }\text{ark’in-na}.] \\
\text{Rasul(1CL) this-ATR.4CL word(4CL) say.PFC leave.PFC.1CL-ATR.1CL} \\
\text{‘Having said this word, Rasul left.’}
\end{array}
\]

However, center-embedding is not possible in DS-sentences (28).

(27) Standard order (DS)

\[
\begin{array}{c}
\text{[CP}_\text{con} \text{ }\text{še-na} \text{ }\text{solulqa} \text{ }\text{ark’in,}] \\
\text{[CP}_\text{mtx} \text{ }\text{ži} \text{ }\text{žiga-j-l} \text{ }\text{ax-u.]} \\
\text{he-ATR.1CL to.the.left leave.PFC.1CL I(1CL) place-OBL-SUP} \\
\text{stay.1CL-PFC} \\
\text{‘He having gone to the left, I stayed.’}
\end{array}
\]

(28) Center-embedding not possible (DS)

\[
\begin{array}{c}
*\text{[CP}_\text{mtx} \text{ }\text{ži, } \text{[CP}_\text{con} \text{ }\text{še-na} \text{ }\text{solulqa} \text{ }\text{ark’in},] \text{ }\text{žiga-j-l} \text{ }\text{ax-u.]} \\
\text{I(1CL) he-ATR.1CL to.the.left leave.PFC.1CL place-OBL-SUP} \\
\text{stay.1CL-PFC} \\
\text{‘He having gone to the left, I stayed.’}
\end{array}
\]

So, in Tsakhur the influence of switch-reference on sentence linearization seems to be grammaticalized. This is exactly what one would expect.
6. Summary

First, I have shown that languages have the strategy to mark switch-reference (SR) syntactically. This finding is in contrast to the traditional assumption that switch-reference marking is an exclusively morphological phenomenon. I provide evidence from a corpus-based study that in German SR is marked by the linearization of the matrix clause and the subordinate clause: if a matrix and a subordinate clause within the same syntactic structure have coreferent subjects, the subordinate clause prefers to follow its matrix clause. If the subjects are different, the subordinate clause prefers to precede the matrix clause.

Secondly, I made the observation that SR marking in German does not seem to be a fully grammaticalized phenomenon, but that nevertheless it is reflected in clear statistical tendencies. I have shown that fully grammaticalized SR-marking, like in Amele, Mojave or Kiowa, and optional (but statistically preferred) SR-marking like in German both can be modeled in the framework of stochastic Optimality Theory without having to treat optionality as a grammar-external performance phenomenon.

Acknowledgements

For helpful advice, comments and discussion I would like to thank Gereon Müller, Philipp Weisser, the audience of ConSOLE XXII and an anonymous reviewer. My particular and sincere thanks go to Andreas Opitz for help and support dealing with the statistical analysis.

Katja Barnickel
University of Leipzig
katja.barnickel@uni-leipzig.de
www.uni-leipzig.de/~barnickel

References


