

# "Rebellious Subjects"? Subject criteria in modern Germanic languages between valency, text reference and information structure

Pierre-Yves Modicom (Université Bordeaux-Montaigne)

## 1. The problem of the subject

The definition of what should be called a subject and which criteria should be used to identify and compare subjects across languages is a significant issue in typological linguistics, at least since Keenan (1976) and Li & Thompson (1976). While problems were first raised for non-Indo-European languages (e.g. Chinese or Tagalog), things are not much clearer in classical languages, including those of the Germanic family. From a theoretical point of view, there is not even a consensus on what subjects are the “subjects of”, either : of the proposition ? of the clause ? of the verbal phrase ? of the verb ?

Proposition-based accounts take a semantic point of view and usually resort to (a variant of) the classical subject-predicate dichotomy, whose conceptual history is in no way self-evident, either. More recently, information-structural or semantico-pragmatic approaches (see Chafe 1976) have insisted on the central role of topicality as a feature of subjects, partly reinvigorating the original insights of Aristotle. Even though his works inspired the classical dichotomy subject-predicate, his own stance was heavily influenced by what would now be called information-structural issues. This has led to much fine-grained work on the opposition between subject-prominent and topic-prominent languages, an opposition which should not be understood as discrete (Li & Thompson 1976). Consequently, a series of subject properties appear to depend on text-structuring mechanisms, anaphora and information-structural features such as topicality (Lazard 1994).

On the other hand, accounts centered on the clause, the VP or the verb are more concerned with syntax and valency. They are also divergent, partly because of other theoretical disputes on the syntax of the clause. This kind of approach has been much sensitive to the progress of descriptive linguistics, as well. The notion of alignment, i.e. the default morphosyntactic construction of prototypical semantic roles within one language, has cast doubts on many assumptions about the universality of the notion of subject (see the studies in Aikhenvald, Dixon & Onishi 2001), but also on the alleged radical asymmetry between subject and complement or subject and object (so-called configurationality, which is now regarded by many as a language-specific property). An additional issue is the fact that many semantic models still resort to the notion of predicate to describe the function of the verb, or of the narrow VP (excluding, among other, so-called external arguments, such as the subject). Thus, while the pair subject-predicate is banned from some of these syntactic or typological studies, it reappears in some others, at a level which is still different from the information-structural one.

Recent research on subjects in Germanic languages such as German and Icelandic has paid much attention to this crisis of the syntactic definition of subjects (Barðdal 2002 and subsequent). Both languages have preserved a rather strong system of case-marking. For this reason, the mismatch between grammatical case and syntactic function exhibited by non-nominative subjects has become a central issue. This was already the case in the previous wave of grammatical and semantic research on subjecthood in Germanic, during the 19th century, when the analysis of so-called impersonal verbs led to the questioning of subject-obligatoriness in German. Some scholars have even proposed to drop the notion of subject as a uniform cross-linguistic category (Croft 2000) and to work on two notions of subjecthood, informational and argumental (Lazard 1994 and 2012). In this perspective, how both dimensions are intricated or separated in each language becomes a central issue. Instead of resorting to the gradual binary opposition between subject-prominent languages (e.g. English or French) and topic-prominent languages (e.g. Chinese), I claim that three poles of configuration are possible within the set of languages with a basically accusative alignment. Since there are no ergative or split-ergative languages in Germanic, I leave aside the specific issues they raise, although the problems of oblique subjects and deponent verbs also lead me to discuss issues pertaining to alignment.

- Type 1: The syntax of a language can be organized predominantly following argumental features and frame constructions of verbs. Default information-structural values are associated to the different argumental positions; marked information-structural strategies such as clefts or contrastive stress are requested to bypass these default values. This integration of informational subjecthood within argumental subjecthood corresponds to what is classically called a subject-prominent language.
- Type 2: Syntax can be organized predominantly following information-structural features such as the topic-comment distinction. Default semantic roles are associated with such or such informational position. Optional argumental marking (e.g. facultative case particles as in (South-)Eastern Asian languages) are used to flag constructions hurting default templates. This corresponds to traditional topic-prominent languages.
- Type 3: Informational and argumental subjecthood can be two thoroughly independent, cumulative modules, operating with different sorts of markers. This third pole corresponds to languages that are both topic-prominent and subject-prominent.

I argue that type 2 is not represented in Germanic. But while Present-Day English and to a lesser extent Danish side clearly with the first group, High German, Low German and Dutch (and arguably ancient Germanic languages) seem to be located in the vicinity of the third pole.

## 2. Specific features of the proposed monograph

As we have just seen, each of the two major dimensions is characterized by major internal contradictions so that even informational subjecthood and argumental subjecthood have to be defined at the language-specific level. This monograph is intended as a contribution to the ongoing debates on subjecthood in Germanic. It is characterized by two specific features.

First, it expands the comparative stance that until recently was essentially limited to contrastive studies focusing on German and Icelandic. Five Germanic languages are now taken into consideration : High German, Low German, Dutch, English and Danish. Insights from other Germanic languages (Icelandic and non-standard High German varieties) also inform the analysis. Dutch, High and Low German are closely related historically and typologically. All are V2 languages with strong OV biases and a characteristic tendency to use ‘bracket structures’, at least in written discourse. But the downfall of the Dutch case and agreement sets this language apart from the other two, especially when it comes to the morphological ‘coding properties’ of subjects. The fourth West Germanic language in this set, (British) English, has lost the distinctive typological properties of the branch and has developed a strict subject-verb cluster. Case and agreement have all but disappeared and word-order rules are very different from what they are in other West Germanic systems. Finally, like other exponents of Mainland Scandinavian languages but unlike Icelandic, Danish has also undergone an almost complete loss of case and agreement, while preserving Scandinavian word-order rules, which to some extent are comparable with continental West Germanic : most prominently, Danish also allows scrambling and it is a V2 language. On the other hand, Mainland Scandinavian languages are now essentially VO languages like English. These contradictory features hint at the difficulty to find a common core of formal properties for subjects across the Germanic family, whereas some functional dimensions, most prominently the information-structural one, are also likely to be impacted by the different word-order rules, especially when it comes to V2 and scrambling.

Second, this work is centered on the study of corpus from the late 16th century (1587-1592). The corpus is made of a High German chapbook from the year 1587 and its contemporary translations in the other four languages. The High German original is arguably the first printed book dealing with the legend of Johann Faust, a self-proclaimed alchemist who had died in 1540. Its considerable commercial success throughout Protestant Europe is manifested by its translations and its direct influence on other works such as Marlowe’s *Doctor Faustus* (1589), the most important

predecessor to Goethe's well-known drama. These translations are thus of high interest in the perspective of cultural history. This historical relevance has a thorough linguistic dimension : these chapbooks were published during a crucial phase of linguistic *Ausbau* in all five languages, less than half a century after Luther's last revision of his Bible (1545) and after the first Danish translation of the Bible (1550), but more than twenty years before the *Statenvertaling* (1618) and the *King James Bible* (1611), although the Elizabethan age already represented a major shift in the development and standardization of written English. Conversely, this period is crucial in the downfall of Low German, which had long been the language of the Hanseatic league, to a low variety in a situation of diglossia with High German. Finally, the corpus provides critical examples of the relationship between core grammatical features and the history of writing codes. This age, which saw the emergence of a proper vernacular grammatical tradition in High German and English, comes shortly before major changes in writing practices throughout learned Europe during the first half of the 17th century. Texts from the 1580s and 1590s are still representative of the periodic writing style influenced by the Latin rhetorical paradigm of *copia* (abundance) ; most crucially for us, the fact that the rhetorical period, not the sentence, is the communicative unit of reference, has major syntactic implications, regarding word order as well as the distinction of coordination and subordination, which was partly reversible at the time. Text movements are often molded onto the protasis-apodosis structure of classical Latin rhetoric, involving the extensive use of correlative structures and left-dislocations. In High German, Low German and Dutch, this leads to the use of information-structurally marked syntactic patterns and word order constructions that disrupt the frame of the clause. Further, all five languages exhibit flexible patterns of anaphor and pronominal ellipsis putting the traditional test of subject ellipsis into jeopardy. Keeping in mind present-day varieties, further systemic differences pertaining to subjecthood are to be noted. First, reflexive pronouns are also much less systematically used than today, endangering a second classical test for the grammatical entrenchment of informational subjecthood. Second, regarding English, one of the major differences between Early Modern English and Present Day English is the fact that Early Modern English had not developed the system of *do-anaphors* in interrogative and negative questions, that plays a decisive role in establishing a radical syntactic asymmetry between subjects and objects in English. All these factors contribute to make the linguistic examination of a parallel corpus from the years 1587-1592 especially interesting.

### 3. Outlines of the study

The monograph is organized into an introduction, five main chapters and a conclusion. The introduction deals with the problem of subjecthood as a cross-linguistic and language-specific category and presents the issues and questions underlying the book. **Chapter 1** is a panorama of subjecthood in present-day Germanic languages, drawing on descriptive grammars and language-specific theoretical works. This chapter illustrates the tension within and between argumental and informational subjecthood in Germanic. The use of original, authentic examples shows that the previously described oppositions, features and configurations, which were located at the level of language systems, are also observable at the level of individual constructions. This holds for the argumental frames of individual verbs, much in the spirit of recent developments in Construction Grammar (in the tradition of the *FrameNet* program), but it is also true at level of discourse structure and/or information structure. Each verbal construction is marked by the interplay of both levels and can re-negotiate the balance described above between informational and argumental prominence. **Chapter 2** is a presentation of the corpus, followed by a description of the most salient linguistic features of the five languages in the late 16<sup>th</sup> century, based on what can really be found in the five versions of the Faustus chapbook taken into consideration (see above).

**Chapter 3** is a description of subject syntax at the level of discourse structure as well as clause-level information structure. The first section is devoted to ellipsis mechanisms and other transphrasitic features such as consecutive inversion (or consecutive V1), a phenomenon that has all but disappeared from present-day varieties of these five languages but was still very vivid in High

German and Low German at the time. Quotative inversion, a superficially similar but functionally very different word-order construction, is also discussed here for the sake of disambiguation. The second section of this chapter deals with information-structurally marked constructions at the beginning of the utterance: dislocation and correlative constructions, as well as topic change constructions. Once more, High and Low German show more flexibility than the other three languages, but only English is characterized by major differences, which are analyzed separately.

**Chapter 4** is devoted to several related issues at the interface between argumental and informational subjecthood: the use of so-called formal subjects, the existence of one or several expletives (or the lack thereof), and the possibility of subjectless clauses that do not fall under the label of discourse-motivated ellipsis as analyzed in chapter 3. First, I show the limits of the notion of formal subjects in all five languages due to the fuzzy borders between so-called formal subjects and ana- and, above all, cataphoric subject pronouns, at least when we leave the realm of meteorological verbs. I then turn to expletives, which I define as positional fillers (mostly) specialized for the first position in the clause. Two contexts of use have to be examined separately: existential and presentative constructions, which are theoretically redefined in consideration of the corpus data. These constructions involve fillers in all languages, with specific restrictions in English. The study of presentative constructions in English reveals the existence of optional subject agreement in the translator's idiolect, a pattern which had never been signaled before in English, but is attested in similar contexts in non-standard varieties of Germanic (e.g. in Yiddish). In Dutch and Danish, an expletive of locative origin is systematically used (*dær* resp. *der*), with interesting signs of further developments to non-expletive uses as "formal subject" used in existential contexts. The other crucial context, including for the use for expletives, is the passive voice. High German, Low German and Dutch are showed to allow proper subjectless passives, which is not the case of English and Danish – whereas Present-Day Danish can build subjectless passive clauses. The situation of Danish is taken under special scrutiny: like all Mainland Scandinavian languages, two passive constructions (one of them a medio-passive) alternate in discourse, a point already studied carefully in chapter 1. Deponent verbs are built using the "morphological passive" (medio-passive). The argumental frames of these deponent verbs are various and marked by semantically-motivated case alternations. This hints at a certain degree of autonomy for case-assignment mechanisms within the set of argumental subject properties.

**The fifth and last chapter** addresses the specific issue of non-nominative subjects in all five languages. The English translation is dealt with separately at the end of the chapter, due to its numerous idiosyncrasies such as the presence of remnant oblique subject constructions for lexical verbs denoting processes other than internal states. The comparative part of the chapter is divided into three sections. First, I review verbs of internal states, classified on an onomasiological base. It appears that High German uses both dative and accusative subjects, whereas Low German resorts to the dative and Dutch and Danish mostly add either a "formal" or a cataphoric neutral subject. Moreover, the subjecthood of these oblique arguments is better established in High and Low German. Second, I cross-examine the various constructions of the cognates of *think*, which exhibit oblique subject constructions in all five languages. All the verbs at stake allow several case-assignment patterns, corresponding to different semantic features and different argumental frames. Third, I concentrate on the uses of the BE with a dative experiencer subject in High and Low German, and their various translations in Dutch, Danish and English. In general, oblique subject constructions are relics of older language stages in all five languages, but it appears that these constructions are much more vivid in High and Low German than in the other three languages. The alternation between several case-assignment patterns is regulated by semantic factors pertaining to *Aktionsart* and to process elaboration. These regularities are also observed for voice alternation in Scandinavian. In interpret this as a feature of semantic alignment within the broadly accusative systems of Germanic languages other than English: the assignment of case for experiencers appear to depend on retroaction mechanisms between the ontology of the process and the ontology of arguments.

## 4. Conclusion

As a conclusion, the study confirms that Keenan's (1976) core coding properties associated with argumental subjecthood are actually construction-dependent in all five languages at the time, although major variation is observed above all in High and Low German, and to a lesser extent in Danish (essentially restricted to the phenomenon of deponent verbs). Informational subjecthood plays a consistent role in the syntax of Germanic languages, but two modes have to be distinguished: High and Low German are characterized by the very strong influence discourse cohesion, especially anaphoric mechanisms, which might have to do with the critical role of periodic style in the organization of utterances. The informational syntax of Early New High and Low German stands under the sign of optionality and variation, to the absolute contrary of Early Modern English, where informational subjecthood represents a constraint intertwined with dependency, and thus an internal feature of argumental subjecthood, even before the rise of *do*-anaphors, which has made this intrication all the more blatant. The internalization of informational subjecthood within argumental subjecthood is much less salient in Early Modern Danish and above all in Early Modern Dutch, which represent intermediate cases between types 1 and type 3 depicted above.

Further, the study confirms that within the set of argumental subjecthood, two groups of properties have to be distinguished. The first associates nominative case-marking and Keenan's "semantic properties of subjects" insofar as they pertain to the ontology of subject arguments. The other group corresponds to argumental frames and the ontology of processes, including some of the selectional restrictions listed by Keenan as subject properties. The notion of agency is relevant at both levels, hence a series of potential mismatches whose fine-grained semantic dimensions can curb case-assignment. The result is the reorganization of Keenan's hierarchy into three levels of entailment: an informational, a frame-structural and an individual-ontological one; the default relationship between the latter two is characteristic of argumental subjecthood, which is branched onto informational subjecthood in a second step. Both connections involve interfaces and the possibility of conflicts. The various forms and degrees of the external nominative subject constraint and agreement rules in Germanic correspond to those mismatches.

## Selected references

- Aikhenvald, Alexandra, R.M.W. Dixon & Masayuki Onishi (eds). 2001. *Non-canonical marking of subjects and objects*. Amsterdam: John Benjamins.
- Barðdal, Johanna. 2002. Oblique subjects in Icelandic and German. *Working papers in Scandinavian Syntax* 70. 61-99.
- Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. Li, Charles (ed.), *Subject and Topic*. New York: Academic Press. 25-56.
- Croft, William. 2000. *Radical Construction Grammar - Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.
- Keenan, Edward. 1976. Towards a universal definition of "Subject". Li, Charles (ed.), *Subject and Topic*. New York: Academic Press. 303-334.
- Lazard, Gilbert. 1994. *L'actance*. Paris : PUF.
- Lazard, Gilbert. 2012. Devoirs de vacance. *Études de linguistique générale, II : La linguistique pure*. Louvain : Peeters. 189-246.
- Li, Charles & Sandra Thompson. 1976. Subject and topic: A new typology of language. Li, Charles (ed.), *Subject and Topic*. New York: Academic Press. 457-490.