Rosanna Sornicola

Interaction of syntactic and pragmatic factors on basic word order in the languages of Europe

1.   Theoretical and methodological foundations[1]

1.1.   A pragmatic study of word order

The study which is the subject of the present work is an attempt to arrive at an integrated view of the interplay of the various levels of analysis – especially the syntactic and pragmatic ones – on the typology of basic constituent order.[2] Adopting such a perspective represents a move away from traditional typological research. A preliminary explanation is required so that, on the one hand, the concepts and terminology of a "pragmatic study" of word order (henceforth, WO) may be clarified and, on the other, that attention may be drawn to the consequences they have for the typological approach followed here.

In a more restricted and technical sense, the study offers a pragmatic perspective because the effects of pragmatic functions on the "dynamics" of rules of order is considered crucial (cf. Section 1.8). But the term "pragmatic" is also used here with a less technical and wider meaning, one which is linked to the first but which goes beyond it in many respects. In this sense the term "pragmatic" would be better understood as equivalent to "functional."

Three concepts may be defined that characterize the wider meaning of the term "pragmatic" used here. They are in turn interrelated: the *multifactoriality*, the *function*, and the *actual conditions* of WO.

The present study has a *multifactorial* viewpoint. Constituents are seen as carriers of grammatical functions (GF), semantic functions (SF), and pragmatic functions (PF). This representation goes back to an integrated conception of grammatical structures, the properties of which are considered with reference to the interplay of the (syntactic), semantic, and pragmatic grammatical levels. Such a conception is all the more indispensable in the study of order structures, in that these may be subject to alteration of a greater or lesser degree,[3] according to variation in the functional values associated with the representation of the constituents. For example, an NP with the GF S may occur in very different structural positions according to whether it has the PF of TOPIC or FOCUS. On the other hand, an NP with the PF of FOCUS may have

a different distribution according to whether it is syntactically encoded as S or as O and/or whether it has the SF of agent (= A) or of patient (= P). The options relating to a particular distribution are, of course, dependent on the typological characteristics of the various languages.

The second key concept is that of the *function* of WO within individual languages and within types of languages. This regards the need to study the interaction of syntactic, semantic, and pragmatic properties and of other factors, rhythmic and prosodic for example, in determining neutral and non-neutral orders. The attempt to analyze the function of WO may therefore be understood as an attempt to ascertain not only what may be defined as the "static" dimension of the WO, in other words, neutral orders, but also its dynamics, that is, the non-neutral or marked orders.[4] As will be seen, the latter have essentially to do with a pragmatic conception. In a certain sense, it may be said that while neutral orders are considered "static," marked orders are "dynamic." Corresponding to this intuition is the fact that, in some approaches, non-neutral or marked WO phenomena have been represented as the effect of syntactic "processes" on basic structures.[5]

In its multifactorial consideration of constituent order and constituent function, the present study is linked to the European functionalist tradition[6] as well as to more recent issues concerning WO "flexibility."[7]

The third concept which defines our approach concerns the central importance given to the *actual conditions* under which WO phenomena arise. In concrete terms, it has involved constant attention to contextual factors, the type of text, variation in the written/spoken register, and, wherever possible, what is traditionally described as "stylistics," in other words, idiosyncratic variation in individual writers/speakers. Attention to actual conditions has meant attention also to the diachrony of WO mechanisms: from time to time, where documentation has been available, it has seemed appropriate to try and take into account the permanence or discontinuity of a WO property. In fact, variability or continuity of WO patterns over time seems to constitute an important typological parameter.

1.2. Word order between competence and performance

Some of the factors just mentioned concern pragmatic competence, others are connected to discourse planning or perceptual strategies more closely related to performance. In fact, a good part of what has been called the "dynamics" of WO phenomena, for example the characteristic pragmatic processes of emphasizing topicalization or focalization by means of WO, seems to concern the vast intermediate area between competence and performance. On the methodological level, it has meant the use of data elicited by means of grammaticality judgements and, more generally, idealization techniques, as well as data gathered from actual texts.

Although attention has been paid to the intermediate area between competence and performance and to structures actually produced, this has not, however, meant that cognitive or perceptual factors have been given pre-eminence in the study of WO. While this seems to be a legitimate area of research (psycholinguistic factors are indeed of paramount importance in empirical issues concerning WO),[8] it was decided that preference should be given to linguistic examination of the strategies which determine WO structures in the various languages. Of course, concepts such as TOPIC and FOCUS have a cognitive content which is not easily modeled. This is perhaps one of the reasons which have made it difficult to arrive at unanimously accepted definitions of PFs (cf. Section 1.8 and thereafter). Nonetheless, a linguistic modeling may and should be sought. With view to a typological comparison in particular, it seemed necessary in fact not to overestimate so-called "natural" or "iconic" factors. Such overestimation seems to be the modern form of an old idea concerning the existence of a relationship between word order and the order of thoughts.

But the choice of a truly linguistic modeling arises from another necessity. The exploration of psycholinguistic factors takes us closer to the mechanisms of performance and these, in turn, characterize "stylistic" (individual) variation rather than the typological properties of the language.[9] It is true, in any case, that the problem of linearization of constituents, of "how to get words into line"[10] remains a crucial one for WO. The question regarding the relationship between linearization in performance, always and inevitably text-oriented and individual, and linearization as a characteristic of a language, or more especially of a linguistic type, appears here in all its complexity. It is an aspect of a broader problem concerning the relationship between performance and competence, rendered more difficult by the inherent variability of WO patterns. This variability, which, as will be seen, also characterizes languages with so-called "rigid" constituent order, often makes definition of a "type" problematic.

The concept of variability may be understood in differing ways. It may be understood as sensitivity to the combination of interacting factors, or as the co-existence of sometimes very different patterns in the language, which may be explained only by the contribution made by historical and sociolinguistic conditions. In considering the structural properties that determine a type, one must not forget to take into account the extent to which it has been formed by "educated" normativization, from above, how far it relates to spontaneous developments in the spoken language (see Sornicola 1981; Miller and Weinert 1998)), and also the extent to which it is the continuation or persistence of old inherited patterns and how much it is due to the external influence of prestigious languages, such as Latin or Greek.[11] This is all the more important in an investigation into the typology of the languages of Europe, which over the span of several millennia have been exposed to very complex and multi-layered cultural influences. But it would be wrong to confine the problem of interference to the languages of Europe that have been exposed to the Latin and/or

360   Rosanna Sornicola

Greek superstratum. It also arises, in a different but no less important way, in the languages of Inner Asia, for which, in absence of written documents prior to the eighth century AD, it is impossible to carry out a reconstruction of their linguistic history. As Denis Sinor (1990) observes:

> The nature of the relationship between the Uralic (Finno-Ugric and Samoyed) and Altaic (Turkic, Mongol and Tunguz) languages which, as far back as the available data allow us to go, have constituted the dominant linguistic group in the forest and tundra zones of Inner Asia, cannot be established with any degree of certainty. While there are those scholars who aver that some or even all of them are genetically related – that is, that they descend from a common, ancestral **Ursprache** – others, including myself, believe that the elements which they unquestionably have in common are due to constant interaction over the centuries if not millennia, and that they result from convergent rather than divergent development. (Sinor 1990: 16)

With respect to the problem of WO, in addition to the above-mentioned historical circumstances, we should also note the observations of Thomason and Kaufman (1988), according to whom:

> The evidence we have collected does not support the often implicit assumption, in the literature on WO change, that WO patterns constitute a fundamental "deep" structural feature relatively impervious to foreign influence. On the contrary, WO seems to be the easiest sort of syntactic feature to borrow or to acquire via language shift. (Thomason and Kaufman 1988: 55)

For all these reasons, examination of WO in a "pragmatic" perspective should not be confined only to perceptual or cognitive (performance) factors, but should also take into account the interweaving of the structural, stylistic, sociolinguistic, and historical factors which have contributed to establishing certain patterns. It is the very interweaving of these factors that allows for the passing from the level of performance, which is stylistic and individual, to the level of the characteristic structures of the language and more especially of the "type." This is to be understood as a general principle superordinate to processes whose crystallization is visible in the structures of the language. For this reason, it seems to us that we need to reassess the view, common to certain traditions of typological study, that the preliminary condition for research is the security that "we are dealing with a typological pattern that is a result of general principles, not of historical accidents (common descent or contact)" (Croft 1990: 22).

1.3.   Microscopy and macroscopy

But to what extent can the typologist take into account the complexity of this interweaving? The comparative dimension in fact makes it more difficult to check the multiplicity of the factors mentioned in every language.

In short, in our attempt to establish a pragmatic typology of WO, we intend to explore patterns of WO in individual languages on the basis of the particular conditions which determine them and, by this means, to arrive at a typological comparison, inevitably problematic and provisional. As will be seen, it is possible to establish "geometric" rules relative to WO. But such rules are the linguist's precarious constructs. They need to be continuously reexamined with respect to historical, sociolinguistic, and stylistic variability.

In confronting these questions, it was necessary to make a distinction between two different but complementary points of view. The viewpoint of classical typology is the "macroperspective." Here the differences between languages are established "from above," as in an aerial photograph. The pragmatic view, especially in the meaning proposed here, is the "microperspective," which requires assessment of the particular conditions in which WO patterns are determined, as in a ground survey. In this sense it is a "fine grained" investigation, as under the microscope. In the study of WO typology, as in other areas of contemporary scientific research, it is evident that rules appearing at the macroscopic level may disappear like an optical illusion when looked at from the microscopic level, where other rules of a quite different nature may appear. What may seem a clear and well-defined structure from the macroscopic viewpoint breaks up into many different phenomena, for which the same description is uncertain. The uncertainty of "historical" descriptions[12] takes over geometric clarity. On the other hand, an eagle's perspective may reveal connections and interrelations invisible from the ground.

However, the two points of view are not irreconcilable. On the contrary, the results from microscopic examination may be crucial to the vast macroscopic constructions, and, vice versa, the latter may offer a perspective, albeit provisional and precarious, on the difficult task of orienting oneself in the multiplicity of pragmatic phenomena. In the present work we have tried to put together arguments from both perspectives.

1.4.   The languages of the corpus

The sample assembled for the study consists of fourteen languages: Basque, Turkish, Georgian, Hungarian, Finnish, Estonian, Russian, French, Italian, Spanish, English, German, Irish, and Welsh. In choosing the languages we have tried to represent the various types of basic order found in the languages of Europe, also taking account of

362   Rosanna Sornicola

different genetic affiliations and areal distributions. Space limitations of the present work have prevented including in the corpus other languages worthy to be examined as, for example, Modern Greek.

The criterion of the representation of basic order, however, has constituted only a very general point of departure for the study of the pragmatic properties of WO. Within each basic order, numerous subrealizations may in fact be identified. Furthermore, basic order, although fundamental, constitutes only one of the parameters to be taken into account when studying the pragmatic properties of WO. It seems to act in a restrictive way, defining the options which are not structurally possible, rather than in a positive way, that is allowing the specification of pragmatically significant processes (topicalization, focalization, etc.).

The sample shows a bias in favor of certain genetic groupings and one areal block: Indo-European languages have the greatest weight in terms of the number of languages examined (eight out of fourteen). The part of Europe to the west of a line drawn from Finland to Hungary and as far as Turkey is well represented, the area to the east much less so. This may in part be explained by the chosen historical and cultural conception of the European space. While, geographically speaking, the extreme eastern boundary of Europe has been recognized, at least since the eighteenth century, as part of the Urals, it is nonetheless true that historically and culturally, the Mediterranean region to the south, and the region beyond the Volga plains and of the Dnepr to the north have for millennia continued to attract Euro-Asiatic populations. As far as the latter and their languages are concerned, it is extremely difficult to reconstruct their history to any great depth due to the scarcity of direct documentary sources. But, as Sinor (1990) observes, it is possible to make out historical processes over a long period that contrast the two areas in terms of sedentary vs. nomadic worlds and their differing economic and cultural standards.[13] As often happens in the presence of such differences, definition of the societies which may be placed in the lower economic level is made in negative terms with respect to the hegemonic societies and by those hegemonic societies themselves.[14]

Above all, it hardly needs mentioning that not only is "European" not a linguistic concept, it is not even a geographical one. The geographical boundaries of Europe have themselves undergone changes over the course of time, in line with different cultural conceptions.[15] One of the most controversial questions has been the demarcation of the borders between "Europe" and "Asia".[16] In Europe, as a historically defined cultural space, the Indo-European languages have prevailed over genetically different languages which have been confined to the eastern margins (or, in the case of Basque, to a residual zone in the far west).

However, the choice of languages in the sample reflects, in part, the limits imposed by the methodology of gathering data (cf. Section 1.5).

Within certain families, such as Romance or Germanic, more languages were considered in order to represent typological variation within a particular family.

There is good reason to suspect that some of the languages of the corpus are not independent (or that they converge due to historical as well as areal considerations). But, as mentioned in Section 1.2, in an area with a high rate of contact and exchange, as in the European space, it would be unthinkable to construct a sample of languages that did not present such a bias (cf. method for reducing the bias in Section 1.3). On the other hand, processes of convergence are highly visible not only within the Indo-European languages, but also between the Ugro-Finnic and Indo-European languages (see Comrie 1981a: 124).

1.5.   The method of gathering data

The approach taken in the research described in Section 1.1 has required the use of several methods for gathering data. The pragmatic description of WO structures and, in particular, the dynamic analysis of the pragmatic processes associated with them, requires in the first place direct examination of written texts or spoken language produced by native speakers, or of judgements made by native speakers on the acceptability and the context of usage of appropriately preconstructed structures presented to them, thus allowing for the collection of different types of data. Both techniques are more characteristic of "in-depth" research on individual languages than of crosslinguistic studies. As far as the former is concerned, simple elicitation of a text would not have sufficed for the specific aim of the study: certain marked WO structures occur rarely, individual variability may be high, etc. On the other hand, the type of questionnaire which needs to be set up for a study of speakers' judgements would require a direct and thorough knowledge of all the languages in the sample.

Other types of questionnaires, of an "explorative" nature (questionnaires, that is, in which the linguist manipulates "structures" which are not lexically realized), present problems for a pragmatic examination: the actual use of a structure and its association with a particular context or register lie below the threshold of consciousness, often even amongst native speakers who are linguists.

For these reasons, preference was given to the use of reliable descriptions (grammars and other secondary sources), to which, in some cases, additional primary sources were added (texts produced by native speakers and/or native speakers' judgements).[17] At times, this has in some ways influenced the choice of individual languages in the sample: once the typological requirements were established and a balanced representation of the various genetic groupings guaranteed, languages were chosen for which there were comprehensive WO descriptions which accommodated the complexities of what has here been defined as "pragmatic."

Wherever possible, the properties of languages belonging to different branches of the same genetic family were considered in order to establish the diachrony of a given phenomenon.

364   Rosanna Sornicola

Of course, the priority assigned to secondary sources has brought with it problems of another sort, such as the bias of the individual monographs or treatises towards one or another linguistic theory, or at least towards a particular descriptive approach.[18] Recognition of this has in fact highlighted the considerable differences in grammatical traditions on WO, which are often associated with particular language families. These differences are interesting in themselves, and they show that WO is an area of linguistic phenomena where descriptive objectivity must take into account not only the complexity of the phenomena themselves, but also the metalanguage which has developed over time to describe it. For this reason, although an attempt has been made to transcodify individual descriptions according to a single homogeneous model, some consideration has been given to the specific properties of the theoretical models describing the phenomena under examination (cf. Sections 1.6 and 1.7).

1.6.   Topology and syntax: Towards a geometry of the sentence

1.6.1.   Notions of order

The analytical procedure followed for each language in the sample has in the first place provided a description of the dominant order and thus a description of its possible variants and factors associated with it. Such a method seemed appropriate because examination of pragmatically determined variations can be carried out more effectively on the basis of neutral structures, without contextual conditions; in other words, an examination of the statics suitably precedes an examination of the dynamics of the WO structures.[19]

The basic order which, as is well known, does not necessarily coincide with the dominant order,[20] has been considered as an initial macrocondition, on the basis of which finer/more subtle articulations may be studied, in other words microarticulations, such as dominant (neutral) orders, non-neutral orders, marked orders. The macrocondition constitutes a more abstract representation than the microconditions. This conceptual distinction seems to be important, even though from an empirical point of view the macrocondition of the basic order may in many cases coincide with one of the microconditions, that is, with that of the dominant order. Traditional typological parameters of correlation between the basic order of major constituents and other serial structures have been taken into account, especially in problematic cases where the scarcity of accessible data or differences of opinion in the existing literature or even the very complexity of the data make it difficult to determine the microarticulations. On the other hand, the importance of such a correlation must not be overestimated, since there are valid reasons to doubt the existence of an implicational relationship between basic constituent order at the sentence level and constituent order at a lower level.[21]

Where the available data allow it, rules of a statistical nature (frequency) have also been taken into account, such as that relating to the prevalent order.[22]

As the work progressed, the distinction between dominant or neutral orders and non-neutral orders has acquired particular importance. The former are characterized by the fact that they can always occur, while non-neutral orders may occur only under particular textual or contextual conditions. In one section of the literature on word order, non-neutral orders are also called "unmarked,"[23] but here preference has been given to use of the term "marked" as a synonym of "emphatic" and, thus, to a distinction between non-neutral and marked orders. In our terms, the concept of non-neutrality includes that of markedness, but not vice versa. This seemed appropriate since in some languages, non-neutral contextually conditioned orders are not necessarily emphatic (cf. Russian and the discussion in Section 2.4.3.2), while in other languages, in particular the so-called "configurational" languages, most variations on the dominant order are associated with processes of emphatic focalization (cf. the Romance languages, English, and in part, German). As has been observed by Abraham (1986) from a generative perspective, in configurational languages "the underlying WO as defined by grammatical criteria is subject to linear variation as required by textual organization. Movement according to textual parameters results in marked linearization"; on the other hand, in non-configurational languages "any variety of WO is base-generated. Surface order will be obtained only by way of the needs of the discourse organization" (Abraham 1986: 15–16).

1.6.2. Models of order as a result of movement

In generative models, the idea that languages allow basic orders and derived orders has been formalized in terms of movement rules which capture the correspondence between configurations with constituents in situ and configurations with constituents that are not in situ. Although highly prevalent, models which generalize movement rules raise serious problems of a theoretical nature.[24]

Representation by means of constituent movement has nonetheless a heuristic validity, perhaps for descriptive rather than explanatory purposes, and for some languages (configurational ones) more than others. In fact, the markedness of an order generally seems to be the result of a process which moves constituents from their "canonical" position in neutral orders. It is not, that is, an absolute but rather a differential parameter. It constitutes a sort of "segmental" correlate to what happens with contrastive stress at the suprasegmental level. As is well known, there is no single physical correlate to this type of stress pattern which functions in an absolute manner within the phonic sequence; in other words, what is perceived as contrastive stress is a series of differential parameters. In this sense, both contrastive stress and marked WO conform to a more general principle, according to which everything in the language is differential or relational. As will be seen, it is possible to arrive at a typology of differentiality for WO (cf. Section 6).

366   Rosanna Sornicola

Of particular importance has been the treatment of non-neutral "marked" orders, characteristic of structures with contrastive FOCUS (the so-called "second instance sentences" of the Prague School). The patterns associated with these structures have very particular properties in the various languages examined (and, it may be assumed, in all languages) and need to be dealt with separately. The discussion which has arisen in recent years in much of the pragmatic typology literature, concerning the nature of languages which have FOCUS in P_1 or P_n (i.e., initial or final) position seems very confused, and does not distinguish the concepts of marked FOCUS and unmarked FOCUS. In many of the languages of Europe, FOCUS in P_1 is only marked FOCUS: both SVO with weak WO flexibility and VSO languages show such behavior (cf. Sections 2.4.4 and 2.4.6).

Account must equally be taken of Daneš' observations that there are semantic types of sentences which may occur with only one pattern of order, even in languages with so-called free constituent order (consider, for example, a pattern such as that in Czech *Lev je selma* 'Lion is a beast').[25] Daneš makes a distinction between "strong rules," which relate to grammaticalized and fixed orders, "weak rules," which relate to "usual" orders, and "free rules," which relate to so-called "variable orders."[26]

For this reason, it was considered appropriate to make a distinction in the discussion between the typology of structures with two arguments and that of structures with one argument (cf. Sections 2 and 3).

As will be seen, it is possible to establish interesting correlations between basic two-argument patterns and one-argument patterns, and between basic two-argument patterns and certain distributional topological properties of the constituents which carry PF (cf. Section 3).

1.6.3.   The concept of position

Any study of WO presupposes a model of the sentence space organized according to "position." It is possible to have a purely linear or sequential representation of positions (serial topological models), or a representation which takes positions to be projections of structural configurations, as in the generative models (complex topological models; but cf. Sections 1.7.2.1 and 1.7.2.2 for a critical examination of this).

The problem that immediately arises is how positions should be counted. In fact, the concept of "position" is closely related to that of "constituent." In other words, every computation of position refers to a particular level of representation by constituents. It is also clear that simple recourse to a linearly ordered sequence of positions $P_1 \ldots P_n$ is not sufficiently informative/distinctive for describing WO properties in natural languages. If concepts such as "sentence initial position" and "sentence final position" could, at first glance, disregard structural configurational properties

(initial position = P_1, final position = P_n), as far as other crucial positions are concerned, such as the immediately preverbal position or the immediately postverbal position, reference must continuously be made to categorial notions.

The objective of studying the pragmatic typology of basic constituents obviously goes hand in hand with consideration of the constituents NP (S), V, NP (O) which, as is well known, are not all constituents on the same hierarchical level. The constituent NP (S) is on the same level as the VP projection which contains V as its head and the NP (O) branch governed by V. If the criterion of counting positions on the basis of the major constituents were adopted, NP (S) would be unproblematic as a position, but it would be more difficult to establish a correspondence between V and NP (O). If the same criterion applied to NP (= S) were followed, the entire VP would be associated with a single position. In this work, the convention of counting each of the phrases NP (S), NP (O) as a position, and of counting V as a position of its own, has been adopted, even though this creates a dishomogeneity in the criteria followed. However, such an assumption is indispensable for a typological study of WO, in which the special relational properties of the head constituent of VP play a crucial role. In SVO languages, V serves as a demarcation of the two GFs S and O. In SOV languages, V acts as the right-hand boundary of the nuclear sentence (all that follows is extrasentential and backgrounded: cf. Section 2.2.1).

How important and how precarious the correspondence between position and constituent is may be seen in some of the problems concerning the representation of adverbial constituents in position P_1 in the Germanic languages. But other difficulties must also be taken into consideration. The very notions of initial and final position cannot be straightforwardly defined without reference to theoretical models. For example, it is disputable whether in structures with left (or right) dislocation, or with hanging topics, very widespread as "natural" phenomena among the world's languages, the dislocated or hanging element may be considered as being in P_1 (P_n), or otherwise. As is well known, some researchers have claimed it necessary to restrict the computation of positions $P_1 \ldots P_n$ to constituents of the sentence nucleus only, reserving a different positional denotation ($X_1 \ldots X_n$) for extranuclear elements.[27] Now, a typological perspective may make the setting up of a single theoretical grid for the computation of positions more complicated. Extranuclear positions at the extreme left or right periphery have important textual properties concerning anchoring to the preceding and succeeding discourse, as seen in configurational languages such as the Romance and Germanic languages. In Russian, the equivalent of these pragmatic properties are the positions P_1, P_n which are not necessarily extrasentential (cf. Section 2.4.3.2).

A further difficulty is posed by differences between the written and spoken language. The intrinsic structural segmentation of the spoken language complicates the establishment of correspondences between the set of positions and the set of categories/structural functions.

368   Rosanna Sornicola
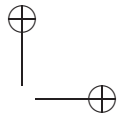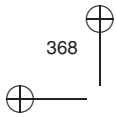
Finally, special mention should be made of some structures that cannot be analyzed in a straightforward way, where the relationship between constituent and position is problematic. This is the case, for example, with cleft sentences in the Celtic languages (or in French), where establishing which constituent occurs in P_1 is at the least controversial (cf. Sections 2.4.4.1.2 and 2.4.6.2).

1.6.4. Representation by fields

The model of representation by "fields" determined by the verb is much used in works of functional typology. Thus a "preverbal field" is distinguished from a "postverbal field."[28] This model is influenced by the Germanic languages, where the verb has effectively a strong property of topological demarcation of the sentence. However, it is questionable whether the fields model is absolute. It is not, in fact, a representation which can lay claims to universality. For example, what may be considered the postverbal field in SOV languages has structural and pragmatic properties which differ from those of what may be considered the postverbal field in an SVO language. Structurally, if X is the constituent occupying the postverbal field, in SOV languages this will yield a representation:

S O V X

while in SVO languages it will yield a representation:

S V X

As far as pragmatic properties are concerned, in SOV languages the postverbal field always has a backgrounding function, while in SVO languages it may have a focalizing function (in neutral structures)[29] or a thematizing/backgrounding function (in marked structures).

The conclusion, therefore, may be that it is difficult to arrive at a consistent topology of the sentence of typological importance. Yet on an empirical basis it is worthwhile making attempts in this direction. Some spatial properties can be captured and can be described in a fairly analytical way, even if the task may not be considered exhausted, but merely started.

It is as well to bear in mind, in any case, that topology is only one of the aspects of a typology of WO.

1.7. The descriptive traditions

1.7.1. Diversity in traditions of word order study

A review of the reference literature on WO in individual languages (or groups of languages genetically interrelated) shows quite clearly to what extent assessment of the
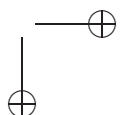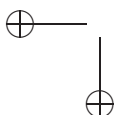
actual phenomena, a prerequisite of any typological comparison, is mediated by the grammatical representations and models with which they are described. In this respect, there are different descriptive and explanatory traditions; some coincide with the study of specific linguistic families: the functionalistic models of Praguean inspiration dominate in the literature on Slavonic languages, generative models prevail in the literature of the last twenty years on the Germanic languages. At times, as in the case of the Romance or Celtic languages, the co-presence of descriptive traditions is noted. It is difficult to get away from the conclusion that these metalinguistic differences reflect more complex cultural differences.

With respect to the comparison a no lesser difficulty is constituted by: (a) the different empirical data which are examined; (b) the different structure of the descriptions. For example, in the functionalist literature great importance has been given to the correlation between structural factors and prosodic factors, of which there are detailed and very interesting studies (cf. Cruttenden 1986, and this volume). On the other hand, in generative literature interest is focused on structural factors, while prosodic correlates such as stress (FOCUS) are looked upon as dependent variables with respect to them.

Some ideas already present in the literature in the nineteenth century seem to have had and continue to have considerable importance, and have been taken up now and then in recent studies. Among them, those which may be worth of a mention are "WO freedom," the influence of "case," and the distinction between grammatical WO and pragmatic WO. The first two concepts are already found in strict relationship with each other in Weil (1879) and in other later works;[30] the third is found in Meillet (1912: 147–148) and also in Mathesius (1939, 1941–1942), and more recently in Thompson (1978). In contemporary literature these ideas seem to have been given a somewhat more rigorous treatment compared to earlier formulations. A general problem in comparative typology, and one that resurfaces in a conspicuous way, is the relative nature of the parameters used, with respect to the language adopted (more or less implicitly) as the element of comparison. In fact, neither the concept of freedom (or rigidity) of WO, nor that of the pervasive influence of case, nor the distinction between languages with grammatical WO and languages with pragmatic WO are absolute principles. It is all the more important to bear this in mind in works on pragmatic phenomena which, as has been said, require a viewpoint which takes multiplicity and flexibility into account. Similar caution applies to other parameters of pragmatic typology which have been widely used in recent years, such as the distinction between topic-prominent languages and subject-prominent languages.

The problem is complicated by the fact that both the generative and "neo-empiricist" typological[31] literatures adopt English as the preferred language of reference for comparison.

370 Rosanna Sornicola

1.7.2.   The d-configurationality model

Partly in the light of this, the recent development of models such as d-configuration-ality will be considered, where the problems which arise in generative theories (and even more so in the EST phase) and in some functional theories seem to become more acute. The basic problem is that such theories do not distinguish between the two different dimensions of syntactic and topological representation. The notions of S, O, and categorial notions (N, V, NP, VP) belong to specifically syntactic representations, while concepts such as "position" and "field" belong to topological representations. In generative models, functional notions such as S, O are defined on the basis of positional notions. A similar albeit implicit assumption can also be made out in Dik's functional models, which axiomatize GFs and associate them with linearization rules: "S and O are typically tied to specific positions, such that the S position precedes the O position. If S/O are not relevant in a language, constituent ordering will mainly be determined by pragmatic factors" (Dik 1989: 365). Elsewhere it is observed that languages such as Hungarian seem not to have specified GFs.[32] Now, position is only one parameter amongst others for codification of GFs,[33] but here it becomes a basic parameter: if a language does not display it, it follows that it does not have specified GFs. This conceptualization is biased by the restriction of the analysis to languages, such as English, in which GFs are codified with position as the element of comparison.

1.7.2.1.   Heltoft's criticism of generative models

The observations of Heltoft, who analyzes generative models on the basis of Diderichsen's traditional Danish model, seem to be no less important with respect to the confusion between syntactic and topological representations. According to Heltoft,

> the development of trace theory within EST appears – seen from the point of view of Diderichsen's paradigm – to be an attempt to solve the problems involved in a topological description, but without introducing any particular topological level of analysis. No particularly topological notions are introduced into the theory (like fields or positions, for example), but the notion of "empty position" is replaced by two different complex notions: 1) empty categories, i.e. nodes where no lexical material is generated in the base, and 2) traces, i.e. nodes from which base-generated lexical material is moved transformationally. This leads to a peculiar kind of ambiguity with regard to the interpretation of the syntactic trees, which, consequently, must function simultaneously as syntactic and topological models. The branching in the trees (and the corresponding PS-rules) become ambiguous, and their significance either as syntactic categories or positions is determined by the origin of
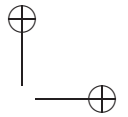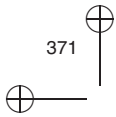
> the lexical material: base-generated lexical material is dominated by a category, whereas moved lexical material occupies a position. This, in turn, implies several things, partly of a general theoretical nature . . . The theoretical problem lies in the fact that it is impossible to interpret the tree diagrams as constituent structures, i.e. we are no longer dealing with a division of the sentence into its component parts. (Heltoft 1986: 61)

Heltoft's observations make reference to crucial questions, such as the relationship between positional representation and representation by constituent. We will see why forthwith.

In topological models, *actual* representations and *virtual* representations may be distinguished. The former are distinguished by the presence of *actual* positions, the second by the presence of *virtual* positions.

A string of *actual* positions must always be associated with a configuration by constituent, in such a way that every P_i corresponds to one and only one constituent X_i (this implies that there are as many positional representations as there are constituent levels). The total topological representation of a language is the set of the sequences $P_1 \ P_2 \ldots P_n$ which are associated in such a way to structural configurations. For example, in the Celtic languages V-initial and S-initial structures give rise, respectively, to the following correspondences:

P_1	P_2	$\ldots$	P_n
V	S		O

P_1	P_2	$\ldots$	P_n
S	V		O

A syntactic representation (categorial or functional), in its constituent or functional relational reality, is by its very nature always associated with a model of actual positions. In fact, every constituent of a given syntactic structure is always found in a particular actual position. In such a syntactic representation, therefore, there is never an empty position (as in structuralist models and the first generative models). The syntactic representations are more real than positional representations by fields.

But topological representations may, as has been said, contain virtual positions. The *Felder* (fields) model drawn up for the Germanic languages, such as those of Diderichsen or others (Section 2.4.5.3) are models of *virtual* position. They express an abstract or virtual maximal schema,[34] independent of any particular structure.

In these virtual representations, particular positions may be empty. In this case, an empty position expresses the fact that the structure does not saturate all of the cells of the virtual schema.

Therefore, as Heltoft observes, in the EST/GB generative models there are positions which are not categories (as for example TOP). The EST models thus mix topological and syntactic representations. Correspondences between categories (GF)

372   Rosanna Sornicola
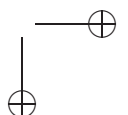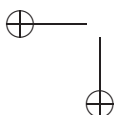
and positions are established aprioristically. A second point of confusion is the fact that both actual and virtual positions are used.

A not very different conceptual mix also arises in a functional model such as Dik's, where positional symbols co-occur with categorial/functional symbols in representations of the type:

$$P_1 \quad V \quad S \quad O$$

where P_1 merely expresses the potential occurrence of a constituent in such a position.

These mixes may be considered the consequence of not having decided which representation is the independent variable and which the dependent. In a topological model, the independent variable is always position and the dependent variables are categorial or functional structures. In the functionalist models, the independent variable is always the abstract relation, in itself nonlinear, between constituents which carry a given function, while positional properties are the dependent variable. In other words, they are a spatial (linear) realization of logico-semantic relations.

In this work representations have been adopted in which correspondences are established between topological (positional) models and categorial/functional configurations. In fact, the positions of a string were considered first and, hence, every position has been associated with a category/function pair. The WO patterns of a language are therefore the set of the correspondences thus defined.

1.7.2.2.   Configurationality and d-configurationality

The concept of d-configurational languages has developed in contrast to that of configurational languages. With regard to the latter, traditionally defined as languages in which the constituents with the GF of S and O have preestablished positions, d-configurational languages are defined by the fact that the PFs of TOPIC and FOCUS may occupy fixed positions in them.

The specificity of the conditions for d-configurationality has often been pointed out in the literature. The fact that WO is conditioned by discourse factors is not in fact sufficient for this to be determined. As Vilkuna (1989: 18) observes, "the additional requirement is that an essential part of the sentence structure be considered by positions for constituents with specific discourse roles. Provided that such functions have an independent status in the syntactic description of a language, one can claim that the language has syntactic functions with a discourse basis." More recently, Kiss (1995) has defined d-configurationality thus:

> The (discourse-)semantic function "topic," serving to foreground a specific individual that something will be predicated about (not necessarily identical with the grammatical subject), is expressed through a particular structural relation (in other words, it is associated with a particular structural position) ... The (discourse-)semantic function "focus,"

expressing identification, is realized through a particular structural re-
lation (that is, by movement into a particular structural position). (Kiss
1995: 6)

Kiss observes that although these two properties often co-occur, they are indepen-
dent. This in a certain sense is obvious, since it concerns the positional realization
of quite different semantic (pragmatic) functions. Furthermore, at the level of re-
alization, there are also clear differences between the range of positions which are
characteristic of the TOPIC and those which are characteristic of the FOCUS.

The d-configurational model may be criticized in various respects. Some of the dif-
ficulties are of a general nature. Although in Kiss' (1995) more recent formulation,
PFs are defined independently on the semantico-pragmatic level and, hence, posi-
tional properties are described as realizational properties, the model strongly favors
the latter in a way that gives scant (or no) account of the complexity of interrelation-
ships on the topological, syntactic, and pragmatic levels. The independent definition
of PFs, on the semantico-pragmatic level, is not developed to its full potential. In
fact, PFs are assimilated *tout court* to constituents which occupy a given position in
the sentence, in conformity with the generative models critically examined in Section
1.7.2.1.

In fact, it may be observed that if we assign the correct definition to PFs such
as FOCUS in pragmatic terms, it will be seen that such a function cannot coincide
entirely with a nominal or verbal constituent, etc., but with the whole sentence struc-
ture.

These difficulties are also linked to more specific problems. For example, in much
research dealing with characteristic FOCUS positions, no distinction is made be-
tween nonmarked and marked orders.

As far as determining the positions of PFs is concerned, account must be taken of
the fact that in the vast majority of languages, at least for the FOCUS function, there
is no unique characteristic position, but a range of characteristic positions, according
to variation in the syntactic and pragmatic parameters. Assessment of the position
is further complicated by the fact that there is not even a biunique correspondence
between position and the constituent which carries a given PF, in languages in which
the range of positional options is very limited, as Turkish or Hungarian. In fact, in
these languages as in others, the categorial nature of the constituent which carries
the FOCUS function is a basic factor in determining the position: when FOCUS is
the verb itself, the position of the FOCUS is different from that which the func-
tion would have if it were carried by a nominal constituent (cf. Sections 2.2.1 and
2.3.2).

A more interesting version of d-configurationality has recently been elaborated
by Abraham (1995), who critically examines the relationship between theme-rheme
structure (T-R) and constituent structure.[35] He observes that while other languages

374   Rosanna Sornicola

such as Hungarian may be considered [−configurational] and [+organized according to T-R], languages such as German are both [+configurational] and [+organized according to T-R],[36] which constitutes an anomaly in the traditional formulation of d-configurationality. Regarding this, it may be observed that the same situation may be found in languages other than German, such as the Romance languages. Of the three logical possibilities:

(a) the T-R structure determines the constituent structure;

(b) the T-R structure may be derived from the constituent structure;

(c) the T-R organization and the sentence grammar are autonomous;

Abraham claims that possibility (b) is methodologically the more interesting one.[37] The claim that the T-R structure may be derived from the constituent structure in such languages is to an extent possible (cf. Section 2.4.5).

1.8.   Form and function in the study of word order

Finally, it is necessary to specify some theoretical assumptions of the work, such as the inventory of terms and concepts adopted, relating to the functions of the various levels (syntactic, semantic, pragmatic).

It was felt appropriate to keep the dimension of constituency distinct from that of the functions of the various levels, as indeed the latter between themselves, and thus to study the interrelationships between these properties. This corresponds to the methodological principle of identifying variables and then studying their interrelationships. The topological (positional) structure has been assumed to be the independent variable, while the other variables (constituent properties, functions of the various levels) have been considered dependent variables.

Definition of the structural domain within which WO phenomena are studied has necessitated a choice. Obviously a pragmatic study requires that the context of analysis be a whole portion of text. We shall call this the "maximal domain," distinct from the "proper domain," constituted by the sentence structure. In fact, although it is necessary to refer to a preceding or subsequent context in order for some WO phenomena to be understood in their textual dynamics (on the semantic/pragmatic level), it is always within a sentence structure that the topological properties of WO must be studied. The maximal domain is therefore connected to semantics and pragmatics, while the proper domain is connected to topological and syntactic structure (for an examination of the topological and syntactic properties of the proper domain, as determined by the preceding and subsequent maximal domain, cf. Section 5).

It must be made clear that it is the simple declarative sentence that is examined as the proper domain here.[38]

For these reasons, examination of topological and syntactic structure is considered methodologically more important than examination of context. It reflects an orienta-

tion where, at least from a descriptive point of view, first the structure and then the function must be determined, even though the two directions of analysis, from form to function and from function to form, must continually overlap. In short, the chosen point of view is that the structures must be established first and then the dynamics of the WO. This has various methodological consequences: (a) neutral structures must be analyzed first, without necessarily referring to context, and then the marked structures, where reference to the context is more or less indispensable; (b) examination of the position-constituent-GF/SF correlations[39] precedes examination of the correlation between these and PFs.

1.8.1.   Grammatical functions, semantic functions, pragmatic functions

On the basis of what has been observed, it is clear that each basic constituent has been considered as having three possible functions:

$$X = (GF, SF, PF)$$

GFs range over the set S, O, indirect O (IO); SFs over the set agent, patient, beneficiary, etc.: PFs over the set TOPIC, FOCUS, TAIL.

As far as GFs are concerned, some problems are caused by the cases crucial to their determination, as in the ergative-absolutive languages in the sample (in Basque and Georgian, in fact, a problem arises in determining S with respect to the constituents with ergative or absolutive case) or in the Balto-Finnic languages (in Finnish and Estonian a problem arises in determining the functions S and O with respect to the partitive case). A further question is constituted by the treatment of constituents marked for morphological dative case but corresponding to a direct O (as in German: cf. Section 2.4.5.5.4, examples (70e) and (70f)).

The biggest problem in defining functions is, however, posed by the PFs TOPIC and FOCUS. As is well known, they are terms with strong polysemy due to the varying uses made of them in the different linguistic traditions.[40] In line with the principle just stated, that the functions of each level must first be defined in terms of the notions of that level (which, of course, does not exclude them being correlated with properties defined on other levels), TOPIC and FOCUS have been defined respectively as the function of "aboutness" and the function of "highlighting." The PFs are therefore represented as primitive concepts of a semantico-pragmatic nature, whose structural manifestation is distributed over more than one level, namely, at the prosodic and the syntactic levels. What seems to be interesting for a theory of PFs is an examination of the interrelationships between the correlates on different levels. An example of an outcome of this type of examination may be given by the generalization which emerges in Section 2.4, according to which a relationship exists in configurational languages between the movement of a constituent into a position in

376   Rosanna Sornicola

which it is not in situ and the fact that it attracts nuclear stress with contrastive function. Examination of such interrelationships seems to be a prerequisite for a study of the pragmatic typology of WO.

The notions of "aboutness" and "highlighting" involve a sort of "solidarity" with other concepts, which may be represented thus:

Aboutness – GIVENNESS

Highlighting – foregrounding – NEWNESS

Note that as far as the first set of concepts is concerned, they do not form a necessary relationship: the function of "aboutness" does not automatically assign the feature [+GIVEN] (think of TOPICS which introduce a new discourse referent). For this reason it is distinct from the textual properties traditionally used to define the concept of "theme" in textual linguistics.[41] It is, furthermore, neutral with respect to the backgrounding/foregrounding dichotomy.

A further characterization of the concept of "aboutness" concerns its relationship with referentiality and argument status. The function of "aboutness" is in fact more typically conveyed by referential elements than by predicative elements. Furthermore, it may not be associated with expletive elements (for example "dummy" pronouns). But the concept of "aboutness" is also related to that of argument status. As is well known, this last property is, in turn, strictly interrelated with referentiality, in the sense that if an element is an argument, it must be referential.[42] Now, the function of "aboutness" is typical of elements with values high up in the argument hierarchy.

The relationship between the function of "highlighting" and that of "foregrounding" is biconditional: the one implies the other and vice versa; neither of the two functions, however, implies the feature [+NEW], nor does the feature [+NEW] imply either of the other two functions. Highlighting is then a different pragmatic property from rhematization. The latter in fact is a typically textual property, connected to the progression of information in the sentence and the text, while this is not necessarily the case for highlighting.

The definition of FOCUS in terms of highlighting could be more subtly articulated in terms of the subfunctions that FOCUS covers at the semantic-pragmatic level. Dik et al. (1981: 60 ff.) identify at least six subtypes:

Completive FOCUS
Selective FOCUS
Expanding FOCUS
Restricting FOCUS
Replacing FOCUS
Parallel FOCUS

Of course, even here, as in all semantic classifications, the number of subtypes could be expanded. However, such finer subclassifications are not very easy to use in a systematic way in large-scale typological surveys. We shall not resort to them here.

A separate problem is that of the distinction between FOCUS in neutral sentences and FOCUS in marked sentences. In the first case, a characterization of FOCUS is in itself problematic (cf. Dik et al. 1981: 51). But even the relationship between the function FOCUS and that of "contrast" is not indisputable (cf. Dik et al. 1981: 57 ff.). In this work, the function of contrast is considered a special subtype of FOCUS.

A recurring question in much of the literature on different languages (and particularly crucial in the study of WO in the Celtic languages: cf. Sections 2.4.6.2 and 2.4.6.4) is the confusion between the concept of TOPIC and that of FOCUS at the semantic-pragmatic level. In fact, it has been noted in many places that the function of establishing center of attention may be associated not only with "highlighting," and thus with FOCUS, but also with "aboutness," and thus TOPIC. However, this does not seem to be an irreconcilable opposition. In a certain sense, it could be said that both functions belong to a single, more general pragmatic function, which is that of "centering attention." However, compared to FOCUS, TOPIC takes the lowest place in such a general function, while FOCUS takes the highest. This definition, which has its basis in the linearity of the discourse, finds its raison d'être in the sentence. In this sense, the neutrality of the TOPIC with respect to the backgrounding/ foregrounding dichotomy may be reinterpreted as the ability of the TOPIC to assume one or the other function according to the textual features specified in it: topics which introduce [+NEW] referents have something of a foregrounding value, although inferior to that of the constituent in FOCUS (for this hypothesis, cf. Sornicola 1993a). This representation allows, on the one hand, the functions TOPIC and FOCUS to be assimilated in terms of the increasing (or decreasing) hierarchy of "centering of attention" and on the other hand, the function TAIL (afterthought) to be distinguished from the first two functions. Thus, the following schema may be arrived at:

Table 1. Distribution of features of "centering attention" among PFs

	Backgrounding	Foregrounding
TOPIC	+	+
FOCUS	−	+
TAIL	+	−

1.8.1.1.   The syntactic and prosodic realization of pragmatic functions

At the level of syntactic structure TOPIC and FOCUS always coincide with a particular constituent, the length and structural complexity of which are variable. The

378   Rosanna Sornicola

methodological principles outlined with respect to the relationship between the maximal domain and the proper domain in the analysis of WO (cf. Section 1.8) are also valid here. At the maximal domain level (text), TOPIC and FOCUS may each coincide with a whole sentence (cf. Sornicola 1981, and this chapter Section 4), while at the minimal domain level (sentence), TOPIC and FOCUS each coincide with a constituent or with a sequence of constituents.[43] Of course, identification of the constituent structure which realizes the PFs, and in particular FOCUS, depends in part on the presuppositions associated with the structure.

Although definition of the concepts of TOPIC and FOCUS are independent of their projection onto a syntactic structure, from a descriptive point of view it seems essential to take the syntactic structures as the point of departure from which to arrive at the pragmatic functions (cf. Sornicola 1993b). However, even from a descriptive point of view, it is important for the analysis to be multifactorial and bidirectional; in other words, it should proceed from GFs to PFs and vice versa.

As has been seen in the previous paragraph, there are basic semantico-pragmatic differences between the function TOPIC and the function FOCUS. Similarly, they show considerable asymmetries on the level of realization.

Despite what has been said about the relationship between TOPIC, aboutness, and referentiality/argument status, at the minimal domain the function TOPIC always coincides with constituents with the feature $[-V]$.[44] From a topological point of view, it may be observed that the tendency to encode the function TOPIC in one of the sentence initial positions is a widespread phenomenon amongst the world's languages.

The function FOCUS has a wider range of realizational options. It may in fact be realized by means of suprasegmental (accentual) and/or positional devices, but may equally be realized by means of processes which isolate the constituent, as in French for example (cf. Section 2.4.4.1.2) or, again, by inserting into the structure lexical material which has a characteristic function of calling to attention: particles as in the case of Lapp (cf. Fernandez 1986, 1994) or pieces of more complex configurations, as in the case of the Celtic languages (cf. Section 2.4.6.2) or English (cf. Miller, this volume).

Categorially as well, FOCUS may coincide with a wider range of constituents than TOPIC; in fact it also allows predicative constituents as its categorial domain, and in this sense it may be said that it ranges over any type of constituent.[45]

Finally, FOCUS is realized in a wider and different set of characteristic positions than that of TOPIC; this set varies according to the type of language and to whether it concerns an unmarked or marked FOCUS.

In the generative literature, FOCUS has often been made to coincide with prosodic correlates of the phrase structure.[46] It is in fact defined as the constituent that receives the main stress. From this point of view, differentiation between unmarked and marked FOCUS may be made on the basis of "default stress" and "contrastive stress,"

which, however, involve interesting empirical problems in the prosodic analysis.[47] On the other hand, in much of the functionalist literature, stress is considered to be a prosodic correlate of pragmatic FOCUS.[48] In this work, the latter perspective is adopted.

2. Word order in sentences with two arguments: The relationship between position, syntactic function, and pragmatic function

2.1. TOPIC position

The range of TOPIC positions in the languages of the corpus is much more restricted than that of FOCUS positions. In both unmarked and marked orders, SOV languages have the constituent carrying the PF TOPIC in one of the positions preceding P_k (i.e., the immediately preverbal position [see Figures 1–2]). In the SVO languages in the unmarked sentences the constituent carrying the PF TOPIC is in one of the positions preceding V, while in the marked sentences it may occupy either any preverbal or any postverbal position (see Section 2.4). The lack of one-to-one mapping between position and TOPIC function is also typical of VSO languages, where the TOPIC position varies according to the unmarked or marked character of the structure of order (see Section 2.4.6.6).

2.2. Languages with O in preverbal position in the basic order and FOCUS in preverbal position

In a vast linguistic area, encompassing Altaic languages of the Turkic (Turkish, Tatar) and Mongolian branch, Uralic languages such as the Ugric (Ob-Ugric and Hungarian) and Volgaic (Mari) languages, as well as Basque, the constituent which bears the function FOCUS is obligatorily placed in the position immediately preceding the verb.[49] Specification of the nature of the constituent in this position is necessary because, as will be seen in Section 4.1, different configurations arise when it is V that carries the function FOCUS. This has interesting repercussions on the theoretical side because it challenges attempts to express the function FOCUS in unequivocal terms of position (configuration) (cf. also the discussion on Hungarian).

This characteristic of argument FOCUS, that it occurs in immediately preverbal position, co-occurs with a collection of other syntactic and morphological properties: (1) all the languages mentioned have basic SOV order (except Hungarian, the assessment of which is problematic: cf. Section 2.3.2); (2) all are agglutinating; (3) in addition, all occurrences of WH-words in interrogative sentences are obligato-

380   Rosanna Sornicola

rily placed in the immediately preverbal position.[50] The areal distribution might lead to the belief that these characteristics are part of the Uralo-Altaic family inheritance, reinforced perhaps by processes of areal convergence (as in the Mari-Turkic relationship), but a broader study, extended to other Euro-Asiatic areas, shows that a number of the features in question are present in most of the modern Indo-Aryan languages as well (Hindi, Nepalese, Bengali, etc.[51]): with the exception of Kashmiri,[52] they all have basic SOV order, and have FOCUS and WH-words positioned immediately before V,[53] although defining their morphology as inflectional or agglutinating is more problematic.[54]

Moreover, the situation with respect to the oldest phase of the Indo-European languages, the basic pattern of which was SOV,[55] shows that word order as such cannot be correlated with the WH-word position parameter. In fact, in these languages WH-words in interrogative sentences occurred in a position preceding the verb, often in P_1, but were in no way specifically associated with the immediately preverbal position.[56]

In fact, none of these three apparently connected parameters – SOV order, agglutination, and WH-words in immediately preverbal position – seems to form an implicational relationship with the obligatory placing of FOCUS in immediately preverbal position. For convenience, we shall indicate the placing of FOCUS in immediately preverbal position as A, SOV order as B, agglutination as C, WH-words in immediately preverbal position as D. It can be said that *A does not entail B* (cf. the situation in Hungarian described in Section 2.3.2), and, conversely, *B does not entail A* (cf. Japanese, an SOV language which, however, does not exhibit parameter A). Also worth mentioning is the fact that some SOV languages, such as Somali and Quechua, do not have property A, but have a property that may be called A', defined as:

> If X is the constituent with the feature [+FOCUS], X may occur in one of the preverbal positions.[57]

A′ is obviously a quite different property from A: the only focal position defined by A is a proper subset of the range of possible focal positions defined by A′.

As far as C is concerned, this parameter seems to be only indirectly related to A. It is, however, well known that agglutination exhibits a strong correlation with parameter B.[58]

More promising instead is the relationship between A and D. Making crosslinguistic generalizations, Horvath proposed the existence of a close relationship between the position of FOCUS and the position of WH-words in the following terms:

> The syntactic position(s) in which non-echo interrogative WH-phrases can appear in a language L will be identical to or be a proper subset of the positions in which FOCUS constituents can appear in a language L. (Horvath 1986: 122)[59]

In languages in which X with the feature [+FOCUS] occurs in immediately preverbal position, WH-words are expected to occur in the same position, and therefore it is expected that *A entails D*. In fact, there does not seem to be a single case in which A but not D is found.[60]

In conclusion, it seems, as far as is known, that the immediately preverbal position for X [+FOCUS] is indeed a property which has developed in SOV languages; it is, however, genetically and areally too localized to be considered a characteristic of substantial typological significance (the only clear evidence of it comes from some Ural-Altaic languages and Basque). Although the relationship between association of the FOCUS constituent with immediately preverbal position and SOV type is not implicational, it nonetheless seems to be far from accidental: in Section 6, the hypothesis is advanced that it arises from a syntactic property, that is, the occurrence of O immediately preceding V.

2.2.1.   Basque and Turkish

Basque and Turkish are two languages in which obligatory placement of the argument constituent in FOCUS in preverbal position has been well studied.[61]

Consider the following examples from Basque:

(1)   Aita-k   untzia   aurdik-i   d-u.
        father-ERG   vase(ABS)   throw-PST.PART   3SG.OBJ-have(3SG.SBJ)
        'The father has thrown the vase.'

(2)   Untzia aitak aurdiki du.
        'The vase, it is the father who has thrown it.'[62]

Sentence (1) shows "normal" word order in Basque. The pragmatic function attributable to the O constituent is that of unmarked FOCUS: it is clearly analogous to what occurs in SVO languages where in sentences with a "neutral" prosodic contour, the pragmatic function of unmarked FOCUS may be assigned to the constituent with the syntactic function O. On the one hand, as is clear from example (2), *untzia* in P_1 is the TOPIC, while whichever constituent other than O occurs in the immediately preverbal position constitutes a marked (contrastive or emphatic) FOCUS. If, on the other hand, an O occupying the immediately preverbal position (i.e., its canonical position in Basque) is also associated with a prosodic contour which carries heavy stress, it may be considered a marked FOCUS.[63] Note that O occurring in its "normal" (i.e., immediately preverbal) position may carry what Lafitte (1962: 17) defines "affective stress." It is not clear whether in this case it can be considered a marked FOCUS (e.g., a constituent with a contrastive value) or merely a constituent with a prosodic intensification of affective (emotional) nature.

Similar considerations apply to the examples from the Turkic languages. In Turkish, a structure such as (3a) has a neutral constituent order:

382   Rosanna Sornicola

(3)   a.   Murat para-yı   bu adam-a   ver-di.
          Murat money-ACC this man-DAT give-PST (3SG)
          'Murat gave the money to this man.'

Structures (3b)–(3f) are also possible, each of which has a marked order, in the sense of being both nonbasic and pragmatically marked:

(3)   b.   Murat bu adama parayı verdi.
          'As for Murat, it is the money that he gave to this man.'
      c.   Parayı Murat bu adama verdi.
          'As for the money, it is to this man that Murat gave it.'
      d.   Parayı bu adama Murat verdi.
          'As for the money, it is Murat who gave it to this man.'
      e.   Bu adama Murat parayı verdi.
          'As to this man, it is the money that Murat gave to him.'
      f.   Bu adama parayı Murat verdi.
          'As to this man, it is Murat who gave the money to him.'[64]

As can be seen, the FOCUS constituent must be placed in the immediately preverbal position. It should be noted that, as in Basque, unmarked or neutral FOCUS (that is "the element that conveys the new information in a sentence with an unmarked order") is that "immediately to the left of the verb."[65] The situation with respect to WO in Turkish is further complicated by the effects of semantic and pragmatic factors such as features of definiteness, animacy, and referentiality (cf. Section 5.1). The structure with FOCUS (whether neutral or unmarked) in immediately preverbal position is also found in Tatar:

(4)   a.   Kolxozčə-lar   uzgan jel bu traktor-nə   satəp al-gan-nar
          collective.farmer-PL last   year this tractor-ACC buy-PST.PART-PL
          i-de.
          be-PST
          'The collective farmers bought this tractor last year.'
      b.   Bu traktornə uzgan jel kolkozčəlar satəp algannar ide.
          'As for this tractor, the ones who bought it last year were the collective farmers.'[66]

In those languages which follow this pattern, the following linear ordering of the pragmatic functions TOPIC and FOCUS is clearly discernable:

$$P_1 \ldots P_k \text{ V, where } P_1 = \text{TOPIC}; P_k = \text{FOCUS}$$

This characteristic is related to another one which is also found in the languages examined: the demarcation of two distinct fields with respect to V, the "preverbal field" and the "postverbal field," the first reserved for codification of the functions TOPIC and FOCUS, the second for linear deployment of informationally GIVEN or of

"backgrounded" material, in accordance with conditions partly linked to afterthought processes (cf. Schroeder 1995).

As far as the pre-field is concerned, on the basis of what was said in Section 1, the sequence $P_1 \ldots P_k$ may be represented as a scale of highlighting increasing in value from P_1 to P_k. This is a reformulation of the interpretation laid down by classic functional sentence perspective theory, according to which the "normal" distribution of information in a sentence would go from the minimum in P_1 to the maximum in P_n, where $V = P_j$ and the constituent in P_n follows V.[67] It confirms that it is not the position of V with respect to the argument constituents S and O that is crucial for information flow, but only that of S and O.[68]

It may therefore be claimed that when FOCUS is not distributed over the entire sentence, but instead coincides with a particular argument of V, its position is determined by the position (or range of positions) that O may assume in a language (cf. Section 6). However, when it is V that carries FOCUS, the languages in question show additional important properties. A different analysis is offered by Rebuschi (1984: 77) who observes that in the Navarro-Labordino variety of Basque, if a sentence has the order O + participle + Aux, the focal element may only be the participle or the O + participle phrase.[69] In Navarro-Labordino, disambiguation of this double possibility may be achieved by means of the Aux – V participle inversion structure which unambiguously identifies the constituent preceding Aux as FOCUS. In any case, these properties force a reappraisal of the d-configurationality model, which establishes a biunique relation between position and PF.

Bearing in mind the limitations just mentioned, it seems that the following generalizations with respect to the languages under examination may be put forward:

– Biuniqueness of the relationship between position and FOCUS (P_k in the configuration P_k V is the FOCUS); this property holds for constituents with referential value;
– Isomorphism of the relationship position/FOCUS between structures with marked FOCUS and those with unmarked FOCUS;
– Lack of influence of grammatical functions with respect to the allocation of FOCUS.

On the basis of what has been said, the representation in Figures 1–2 may be arrived at; this shows a specific correspondence between position, pragmatic function, and grammatical function.

384   Rosanna Sornicola

P_1 … P_k V

T F unmarked order

O

Figure 1. Alignment of Positions, PFs and GFs in unmarked orders of SOV languages

P_1 … P_k V

T F marked order

$\begin{Bmatrix} S \\ O \end{Bmatrix}$

Figure 2. Alignment of Positions, PFs and GFs in marked orders of SOV languages

As can be seen, the structure shows a certain rigidity and uniformity; the only possible movement, which produces a marked order, is constituted by S in P_k.[70] One may speculate as to whether this property, which is quite different from that found in languages with basic SVO word order (cf. Section 2.4), does not also have something to do with the agglutinating character of the languages in question (cf. Section 6.2.4.3).

In any case, in dealing with this phenomenon it seems appropriate not to separate the examination of PFs from that of GFs and other typological parameters.

2.3.   Problematic cases concerning the relationship between position, grammatical function, and pragmatic function

2.3.1.   Georgian

Although none too clear for the time being, the example offered by Georgian is interesting. There is a lack of systematic studies on Georgian word order in a pragmatic framework. The available data seem to indicate a complex and in some ways hybrid situation. In fact, Georgian oscillates between SOV and SVO orders, a characteristic which, although more typical of southern Caucasian languages, is not unknown in other branches of the Caucasian family where the preferred word order is SOV.[71] Determining what the dominant neutral order is in Georgian does in fact appear to be problematic. Variation in the types of text has a bearing on the issue. Two surveys conducted by Vogt on a novel and on popular tales produced quite different results. In the first case, SOV order occurs in 75 % and SVO in only 17 % of the sentences in

the corpus. In the second case, SVO order has a frequency of 36 % and SOV 32 %; while the remainder is divided between various orders, including V-initial orders.[72] Determining the subject function among constituents marked for nominative, dative, and ergative case constitutes an additional problem. This differentiation has an important bearing on WO, such as in constructions with so-called "inverted verbs" (cf. Harris 1981: Ch. 8). Vogt, however, who has given due consideration to the question, holds that "quel que soit l'ordre des termes dans une proposition donnée, leurs fonctions respectives par rapport à la forme verbale ne font jamais de doute" (Vogt 1974: 54).

In sentences with four constituents (S, V, O, IO) in the corpus of popular tales, Vogt draws up the following frequency distribution for the dominant types, all S-initial, S O V IO, S IO O V, S V O IO, and S V IO O:[73]

(5)   S O V IO (55 examples; 25 %)

maq'urebl-eb-ma   q'ižina   da   t'aš-i
spectator-PL-ERG shout.(NOM) and applause-NOM

dascxes   gamaržvebul-s.[74]
3PL.cover.3(SG).AOR winner-DAT
'The spectators noisily applauded the winner.'

(6)   S IO O V (41 examples; 19 %)

Ivane-palavan-ma   k'ameč-eb-s   c'q'al-i
Ivan-champion-ERG   buffalo-PL-DAT water-NOM

daalevina.
3SG.make.3(PL).drink.AOR
'Ivan the Champion watered the buffalos.'

(7)   S V O IO (22 examples; 10 %)

Dev-ma   daart'q'a   xmal-i   c'iskvil-s.
dev-ERG[75] 3SG.strike.3(SG).AOR sword-NOM milestone-DAT
'The dev struck the milestone with the sword.'

(8)   S V IO O (19 examples; 9 %)

Upros-ma   ʒma-m   moaxsena   xelmc'ipe-s
elder-ERG brother-ERG 3SG.remind.3(SG).AOR king-DAT

taviant-i   tavgadasaval-i.
their.own-NOM adventure-NOM
'The elder brother reminded the king of their adventure.'

None of the examples (5)–(8) seems to be a marked structure. The occurrence of O in immediately preverbal position is nonetheless related to features of referentiality and definiteness, as is also attested by many examples with less frequent orders. In fact, in many of the examples with O in preverbal position, O has [−referential] and/or

386   Rosanna Sornicola

[−definite] features. On the other hand, functional sentence perspective sometimes plays a role,[76] as in examples (9)–(10) where O in initial position is a GIVEN element:

(9)   Es   amat-i   nalap'arak'ev-i metvalq'ure-m xelmc'ipe-s
         this(NOM) their-NOM   spoken-NOM   supervisor-ERG king-DAT
         moaxsena.
         3SG.remind.3(SG).AOR
         'The supervisor reminded the king of what they had said.'

(10)   Kal-eb-i   p'at'ron-eb-s   mihgvara
         woman-PL-NOM relative-PL-DAT 3SG.take.IO.3(PL).AOR
         q'mac'vil-ma.
         young.man-ERG
         'The young man took the girls to their relatives.'

Position P_1 may be used for the "foregrounding" of a constituent:

(11)   Q'mac'vil-s   mamobil-dedobil-ma male   col-i
         young.man-DAT   adoptive.parents-ERG immediately wife-NOM
         šertes.
         3PL.give.IO.3(SG).AOR
         'The adoptive parents quickly gave a wife to the young man.'

In (12) the O constituent seems to have a true emphatic FOCUS value:

(12)   Šandl-eb-i-c   me   ševucvale   am
         candlestick-PL-NOM-also I(ERG) 1SG.change.for.3(SG).AOR this(OBL)
         kal-eb-s.
         woman-PL-DAT
         'I have changed these girls' candlesticks as well.'

(13)   Amistana   ambav-i mikna   me   im
         such(NOM) news-NOM 3SG.do.IO.1SG.AOR I(DAT) this.OBL
         dac'q'eul-ma dev-ma.
         cursed-ERG   dev-ERG
         'That cursed dev did such a thing to me.'

This evidence would seem to indicate that the parameter concerning occurrence of WH-words in immediately preverbal position is indeed present in Georgian,[77] yet it does not conform to the rules of FOCUS placement which hold for Turkish and Basque. The situation seems comparable to that described by Comrie (1984) for Armenian.

2.3.2.   Hungarian

Hungarian requires special mention. The synchronic and diachronic study of WO in this language, which presents numerous descriptive and interpretative problems, seems to give rise to some interesting speculations which may lead to a better understanding of the driving force behind the formation of the type P_1 = TOPIC, P_k = FOCUS, with P_k immediately preceding V.

Kálmán et al. (1986), Kenesei (1986), and Horvath (1986) have, in different ways, made a useful distinction between neutral and marked sentences, the first with rigid constituent order and a "level" prosodic contour, the second with free constituent order and "eradicating" prosodic contour.[78] It clears up a misunderstanding recurrent in the literature of various approaches, which is that Hungarian is a language with free constituent order. It shows once again the relativity of the concept "freedom of WO" and the problems it raises. The discussed alternation between SVO and SOV patterns does indeed correspond to features of definiteness and referentiality of O, a situation having some points in common with Turkish (cf. Sections 2.2.1 and 5.1). As for the so-called "d-configurationality" defended by Kiss, Komlósy (1986: 219) observes that such a hypothesis assumes "certain invariant properties of Hungarian sentence structure which, in fact, typify only marked sentences."

Equally clarifying is the distinction drawn between unmarked and marked FOCUS. The first, if associated with certain features of definiteness and referentiality, may occur in sentence-final position, in line with what happens in SVO languages,[79] while the second may occur only in immediately preverbal position.

Consider the following sentence with SVO order:

(14)   'János 'sétál-tat-ja   a 'kutyá-já-t   a
         John   walk-CAUS-3SG.SBJ:3DEF.OBJ ART dog-3SG.POSS-ACC ART
         'park-ban.
         park-INESS
         'John is walking his dog in the park.'

and the related structures with "scrambling" in (15)–(18), which exhibit a considerable freedom of constituent order, but only if associated with an "eradicating" prosody:[80]

(15)   'Sétáltatja 'János a 'kutyáját a 'parkban.
         (Major stress on *sétáltatja*; possible secondary stress on *János*);
         FOCUS on V
         'John WALKS his dog in the park.'

(16)   A 'kutyáját 'János 'sétáltatja a 'parkban.
         (Major stress on *János*; secondary stress on *kutyáját*, which may only be an extrasentential constituent);

388   Rosanna Sornicola

FOCUS on S, minor focus on the extrasentential constituent 'As for his dog [lit. his dog], it is JOHN who walks (him) in the park.'

(17)   A ˈparkban ˈJános ˈsétáltatja a ˈkutyáját.
         (Major stress on *János*; secondary stress on *parkban*, which may only be an extrasentential constituent);
         FOCUS on S, minor focus on the extrasentential constituent
         'As for the park [lit. in the park], it is JOHN who walks his dog (there).'

(18)   A ˈparkban a ˈkutyáját ˈsétáltatja ˈJános.
         (Major stress on *kutyáját*, secondary stress on *parkban*, which may only be an extrasentential constituent);
         FOCUS on S, minor focus on the extrasentential constituent
         'As for the park [lit. in the park], it is HIS DOG that John walks (there).'

As can be seen, the immediately preverbal position in the sentences quoted is the FOCUS position, while the first is the TOPIC position.[81]

The appropriateness of the distinction between unmarked and marked FOCUS is further confirmed by a distinction made by Kálmán et al.: they differentiate between a position of proper FOCUS, typical of marked sentences, and a HOCUS position, filled by "the first stressed element in front of the finite ending in level-prosody sentences" (Kálmán et al.: 1986: 133). From a structural point of view, it is the immediately preverbal position in both cases. However, the HOCUS position has no emphatic or contrastive interpretation. It is often, although not exclusively, occupied by a "verb carrier," or a V modifier or complement, as in examples (19a)–(19b):[82]

(19) a.   Fá-t   vág.
             wood-ACC   cut(3SG)
             'He is cutting wood.'
      b.   ˈFá-t   vág-ott.
             wood-ACC   cut-PST(3SG)
             'He cut wood.'

This situation seems to call for a conceptual distinction between a "preverbal" position as the place where a constituent happens to be in unmarked, neutral sentences, and the P_k "preverbal" position, which is the place where FOCUS constituents occur in marked sentences.

In opposition to the suggestion, defended by Kiss (1987), that there is a biunique correlation between position and PFs (the so-called "discourse configurationality": cf. Section 1.7.2), Komlósy maintains that "focus interpretation is not limited to focus position," since V may also be in FOCUS,[83] and that, on the other hand, "focus position does not necessarily entail focus interpretation," as in the case of unmarked sentences with constituents in FOCUS.[84]

Results emerging from a diachronic study of WO in Hungarian seem particularly interesting. In particular, according to Behrens (1989: 142), Old Hungarian differs from Modern Hungarian in that SOV order did not require a contrastive interpretation of O. Behrens (1989: 149) makes the crucial observation that in the period when the preverbal position of O was not so marked, as in the sixteenth and seventeenth centuries, a FOCUS interpretation of the constituents in front of the verb was not obligatory. This fact seems to confirm that the immediately preverbal position arises as a consequence of basic SOV order, and that it is the position of O which determines the position of the constituent with the FOCUS function.

2.4. The interplay of position, grammatical functions, and pragmatic functions in languages with O in postverbal position

2.4.1. General considerations about the position of the FOCUS constituent in SVO languages

In these languages, unmarked FOCUS occurs in postverbal position, thus coinciding with O. Marked FOCUS tends to occupy a position in the preverbal field. This description, however, only takes into account the relationship between position and PF. Here, as in the previous case, the influence exerted by a further element, syntactic function, should also be taken into consideration (cf. Fig. 3).

P_1 … P_k V … P_n

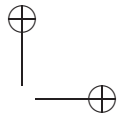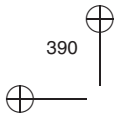T   F   unmarked order

S   O

Figure 3. Alignment of Positions, PFs and GFs in unmarked orders of SVO languages

Those orders shown in the representations in Figures 4, 5, 6, and 7 are all marked.

P_1 … P_k V … P_n

F   T

O   S

Figure 4. Alignment of Positions, PFs and GFs in OVS orders of SVO languages

390   Rosanna Sornicola

P_1	…	P_k	V	…	P_n
F		T			
O		S			

Figure 5. Alignment of Positions, PFs and GFs in OSV orders of SVO languages

P_1	…	P_k	V	…	P_n
T		F			
S		O			

Figure 6. Alignment of Positions, PFs and GFs in SOV orders of SVO languages

P_1	…	P_k	V	…	P_n
T					F
O		CLT			S

Figure 7. Alignment of Positions, PFs and GFs in O CLT V S orders of SVO languages

Compared with the properties exhibited in Figures 1–2, Figures 3–6 show:

– Unimportance of compartmentalization of the sentence space into pre-field and post-field with respect to codification of TOPIC and FOCUS: the constituents which bear such functions may "move over the entire field" of the sentence;[85]

– Lack of isomorphism of the position/FOCUS relationship between structures with marked FOCUS and those with unmarked FOCUS (where P_n = FOCUS); the relation P_n = F in particular, which applies to unmarked structures, does not uniquely identify unmarked structures (cf. Figure 7);

– Influence of grammatical function on the assignment of FOCUS. To arrive at a complete specification in these languages, values must be assigned to each of the three coordinates P (= position), GF (= grammatical function), PF (= pragmatic function).

It should also be pointed out that in the so-called "V-second languages" there are further conditions on WO: the relationship expressed by Figure 5 is not possible (consider the well-known phenomenon of S–V inversion when a constituent other than S occupies position P_1: cf. Section 2.4.5). The relationship between GF and PF expressed in Figure 7 (which represents left dislocation structures) applies in many
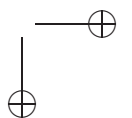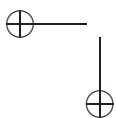
languages which are basically SVO (for example, the Romance languages, Modern Greek), but not in all of them. For example, as will be seen in Section 2.4.3.2 in some Slavonic languages such as Russian, the closest pragmatic equivalent of the pattern [O clitic pronoun V S] of Figure 7 is an OVS syntactic configuration, which is not necessarily associated with a FOCUS V TOPIC pragmatic configuration as in Figure 4. This may be considered a consequence of the greater freedom of constituent movement in these languages compared to languages which show the relationship expressed by Figure 7.

In any case, a generalization seems possible. The criterion for movement of a constituent from its canonical position to a noncanonical one is far more characteristic of languages with basic SVO word order than of those with SOV order. Note that this characteristic has nothing to do with syntactic functions (in Turkish or Basque O may occur in the post-field, but in such a case it may not take on either TOPIC or FOCUS functions), but instead relates specifically to pragmatic functions. The differential character of marked FOCUS in WO variations may be seen to a far greater extent in SVO languages than in SOV languages.

2.4.2.   SVO languages

Languages with basic SVO word order exhibit properties very different from those described in Section 2.2. Amongst the languages of Europe, the majority of modern Indo-European languages have this basic word order, even though they differ with respect to the degree of "freedom" to which they depart from this arrangement, as seen in *actual* orders which may be observed in corpora. It is well known that this difference is strongly interrelated with the presence vs. absence of a morphological case system. Balto-Finnic languages, such as Estonian, Finnish, the Slavonic languages, Lithuanian, Modern Greek, which have maintained a case system, have a range of possible permutations of WO unknown in the Romance languages and Germanic languages, such as English and Swedish, which have lost morphological case.

The existence of a relationship between (a greater or lesser) "freedom" of WO and the presence of a morphological case system was recognized at least as far back at Weil 1844 (= Weil 1879); on the basis of the co-presence of these two characteristics, Weil made a distinction between word order in ancient Indo-European languages and that in modern Indo-European languages. In the latter, the loss of case systems is accompanied by a greater rigidity of WO.[86]

It cannot be denied that the development of some modern Indo-European languages, such as the Romance languages from Latin, English or Swedish from older phases of Germanic, offers solid evidence of this.[87] However, to take the principle just mentioned as absolute would be a mistake for various reasons.

In the first place, there are languages which, despite conserving a case system, have strong tendencies to a nonfree order: German and Icelandic, V-2 languages,

392   Rosanna Sornicola

and Modern Greek exemplify this situation in different ways.[88] The prevalence of nonfree word order may be more conspicuous in the spoken than in the written language: in Modern Greek, for example, the literary language exhibits a range of word order patterns that are not found in the spoken language (cf. De Simone Brouwer 1921: 217). A second problem which argues against labelling the aforementioned principle as absolute is that detailed philological studies of the development of the Romance languages from Latin and of the development of English have shown that the relationship between crystallization of a particular word order and loss of case markers is far more complex than has often been assumed. In Latin, reordering from OV to VO occurs before the period of the first literary texts (cf. Adams 1976: 72). The development of SVO order as the predominant (basic) one in the Romance languages did not therefore result from the loss of the case system; it was instead an option which was available right from the time of the oldest documentation and which surfaces here and there in subsequent literary documentation (cf. Linde 1923: 169 ff.), acquires ground, and presumably aids the process whereby the case system is lost. On the other hand, SVO order was already one of the dominant options in Old English prose,[89] in a period in which erosion of case morphology was only just beginning and was more typical of some areas such as the northern regions. It is true that stabilization of the SVO pattern is clearly attested in the fourteenth century, by which time simplification of case morphology is apparent, but the whole trend does favor a by no means unilinear correlation between the fixing of word order and loss of case.

Of interest, in any case, is the fact that in all the SVO languages that make up our corpus of languages of Europe, the SVO pattern, although appearing in written registers of various sociolinguistic levels, is the preferred trend in the spoken language (cf. Section 6.1.1). It presumably has something to do with the well-known psycholinguistic properties of linearization associated with the SVO type, which constitute tendencies so general as to render the definition of the type in question hardly relevant to the study of genetic affinities. On the other hand, as has been recognized by Hawkins (1983: 16), the very characterization of an SVO type is highly problematic.[90]

The examination of the available data shows that:

1. In all the SVO languages of the corpus, deviations from this order have a more or less "marked" character and depend on textual factors and/or factors determined by the pragmatic operations of topicalization and focalization;

2. WH-words occur in sentence-initial position;

3. SVO languages in the corpus vary according to:
   (a) the freedom with which GFs may occupy nonbasic positions;
   (b) the freedom with which the segmental linear axis is exploited by pragmatic strategies (topicalization, focalization, functional sentence perspective [GIVEN/NEW]);

(c) differing use made of the segmental vs. suprasegmental axis with respect to pragmatic strategies.

Points 3a–3c are aspects so closely related as to be interdependent. Such differences have already been recognized by Mathesius and by the Prague School scholars as being those which distinguish languages with grammatical WO and languages with WO sensitive to the influence of textual factors (cf. Mathesius 1941–1942; for a more recent proposal, see Thompson 1978). However, even this distinction should not be assumed to be a consistent typological parameter (cf. Section 1.7.1).

We will deal with issues 3a and 3b in this section, while for 3c, see Sections 2.4.3.2 and 2.4.4.2.

The SVO languages of the corpus will be divided into languages with highly flexible WO, such as Slavonic and Balto-Finnic (see Section 2.4.3) and languages with weak WO flexibility, such as Romance languages and English (see Section 2.4.4). The V-2 type, represented in our corpus by German, requires a separate discussion (see Section 2.4.5).

2.4.3. SVO languages with highly flexible word order

2.4.3.1. Estonian and Finnish

Estonian and Finnish have interesting properties in common. In both languages, the syntactic restrictions on WO are very slight: that is, there are numerous possible permutations of the constituents departing from the basic order.[91] Nonetheless, of a certain interest for diachronic typology is the fact that in the language used in the translation of the Agricola New Testament, which is one of the oldest written texts in Finnish (circa mid-sixteenth century), potential traces of an order with the constituent in FOCUS immediately in front of the verb turn up.[92]

Tauli (1983: 24–25) observes that in Estonian, both the initial and the final positions are endowed with a certain degree of stress. In particular "the end-position is in general more stressed than the front position, and is particularly appropriate for expressing the unknown, unexpected and contrasting" (Tauli 1983: 25). On the basis of the examples reported by him, it may be concluded that movement of a constituent along the linear axis, towards the initial or the final position, produces a strong focalization effect (marked focalization),[93] as is clear from example (20), where O is in preverbal position (Tauli points out that in this case, O is always stressed):

(20)   ˈTiibu   ei   näi-nud.
         wing:PRTV.PL   NEG see:PST.PART
         'The WINGS I didn't see.'

and from example (21), where S is in postverbal position:
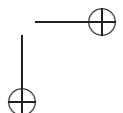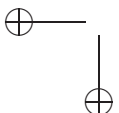
394   Rosanna Sornicola

(21)   Neid   ühenda-s   ühine   huvi.
        they:PRTV unite-PST(3SG) common(NOM.SG) interest(NOM.SG)
        'A common interest united them.'[94]

Similar properties are found in Finnish, where unmarked FOCUS occurs in sentence-final position. Marked FOCUS may occur conversely at the beginning of the sentence. From a normal, unemphatic, sentence such as:

(22)   Juho   lyö   Heikkiä.
        John(NOM) hit(3SG) Henry:PRTV
        'John hits Henry.'

the following six alternatives are possible:

(23)   a.   ꞌJuho lyö ꞌHeikkiä.
            'John hits Henry.'
       b.   ꞌJuho Heikkiä lyö.
            'JOHN and no one else hits Henry.'
       c.   ꞌHeikkiä Juho lyö.
            'It's HENRY that John is hitting.'
       d.   ꞌHeikkiä lyö ꞌJuho.
            'It's JOHN that is hitting HENRY (and someone else the others).'
       e.   ꞌLyö Juho Heikkiä.
            'John IS hitting Henry.'
       f.   ꞌLyö Heikkiä ꞌJuho.
            'Henry IS being hit and it's JOHN that's doing it.'[95]

Sentence (23a) is judged by Hakulinen as "normal," although it differs from the structure in (22) with respect to prosodic contour: it has, in fact, stress on the initial constituent and stress on the final constituent (Hakulinen does not make it clear what the difference between the pragmatic interpretations of the two structures is). It seems clear, in any case, that in structures (23b)–(23f), all of which are marked,[96] the initial position P_1, and the final position P_n, with which a primary and a secondary stress may be associated,[97] are the places assigned to the allocation of FOCUS: if there is only one stress, as in examples (23b), (23c), and (23e), it may only occur in position P_1, which is thus recognized as the FOCUS position; if there are two stresses, this results in a structure with two foci which may be represented as:

$$P_1 \quad \ldots \quad P_n$$
FOCUS1   FOCUS2

Determining a FOCUS hierarchy is more problematic. The interpretation that Hakulinen (1961: 316) assigns to the sentence *ꞌHeikkiä lyö ꞌJuho*, that is, "the fact of hitting is taken as known and what is to be expressed is that as far as Henry is concerned it is John that is doing the hitting," suggests that the functional sentence

perspective determines a FOCUS hierarchy to the advantage of the constituent that occurs in final position. This conclusion goes against the WO model recently drawn up for Finnish by Vilkuna (1989) in terms of "d-configurationality."[98] Vilkuna defines the following d-configuration for Finnish:[99]

> K T V-field

K (= contrast) is an optional constituent; it is obligatorily specified only in WH-constituents or relative clauses and these must precede any other constituent in the sentence. T (= topic) is also optional, in the case of imperative or elliptical sentences (cf. Vilkuna 1989: 38). The preceding schema, in any case, has different realizations. Consider:

(24) a. SVO order
           T   V-field
           Mikko   pesi   astiat.
           Mikko(NOM)   wash:PST:3SG dish:ACC.PL
           'Mikko washed the dishes.'
      b. OSV order
           K   T   V-field
           Astiat Mikko pesi   .
           'The dishes, Mikko washed.'
      c. SOV order
           K   T   V-field
           Mikko astiat pesi.
           'It was Mikko who washed the dishes.'
      d. OVS order
           T   V-field
           Astiat pesi   Mikko.
           'The dishes were washed by Mikko.'

According to Vilkuna, only the configuration in sentence (24a) is neutral, while all the others are marked, in the sense that they are "all-NEW sentences."[100] Vilkuna adds, furthermore:

> A sentence has a syntactically marked order, if:
>
> (a) it has a non-interrogative, non-relative K, or
> (b) the verb follows some of its complements in the V-field, or
> (c) a default topic is present but is not T. (Vilkuna 1989: 43)[101]

The sentence with OVS order (24d) is precisely an example of a structure with "default T." However, it is not clear what the author intends by "syntactically" marked order and what relationship there is between this notion and that of contrast. It is not clear why Vilkuna claims that the only position in which the contrast function may occur is P_1.[102] In this case, one would in fact have:

396   Rosanna Sornicola

    P_1
    K (FOCUS)

It is in fact the OVS example reported earlier that shows that S in final position has a focalizing function (treating the distinction between K and FOCUS with due caution). Nonetheless, the same conclusions may also be drawn from example (23f).

The model proposed by Vilkuna, however, makes comparison with Hakulinen's (1961) data difficult.[103] The absence of a distinction between marked and unmarked sentences has perhaps contributed to the lack of clarity with respect to the complexity of the relationship between GFs, PF, and WO in Finnish. Some of these difficulties seem to arise from descriptive, procedural, and theoretical problems. The definitions of K and T are inconsistent: in some places these concepts are linked to specific syntactic categories (cf. Vilkuna 1989: 38), elsewhere they express pragmatic notions in terms of "dialogue games" (cf. all of Chapter 3). A second problem is constituted by overuse of the linear model of information organization (GIVEN/NEW). This is even odder considering that the author criticizes the classic functional sentence perspective model and its subsequent modifications by, among others, Chafe or Tomlin. The *argumental* conception of TOPIC/THEME chosen by Vilkuna[104] is superimposed to the informational one, although not consistently. Particularly unsatisfying is the definition of K, defined in places as TOPIC (cf. Vilkuna 1989: 91 ff.) and elsewhere as FOCTOP (cf. Vilkuna 1989: 102 ff.) as in the "dialogue games" model.

A third difficulty, the fact that the notion of FOCUS is presented as a concept (term) subordinate to that of K (contrast), seems to be even more problematic. While in many studies in the functional grammar framework contrast is represented as a set of interrelated but varying values which constitute a subset of the values of the pragmatic function FOCUS,[105] here FOCUS (defined simply as "NEW" information) is a subset of the properties of K. The question is anything but terminological: it refers to an unsatisfying modeling of pragmatic functions.

On a more general note, the model presented by Vilkuna seems questionable with respect to the integration of marked and unmarked options into a single structure. Use of the concept of a V-field, devised with respect to Germanic languages and presumably unmotivated for Finnish,[106] is not convincing either.

Finally, Vilkuna's model does not capture the obvious similarities between Finnish and other Balto-Finnic languages such as Estonian.

In conclusion, one may well ask to what extent the K T V-field model is typical of Finnish and not also of languages with a certain freedom of constituent order. If it is true that unmarked sentences tend to conform to the functional sentence perspective, the situation in Finnish as put forward by Vilkuna illustrates a phenomenon long known to be present in the Slavonic, Romance, etc., languages and variously called "ordo artificialis" (cf. Firbas 1964) or "secondary topicalization" (cf. Fillmore 1968: 57). In this sense, Karttunen and Kay's generalization, taken from Vilkuna, would fit

into the more general property whereby in conditions of emphasis the first position is employed for FOCUSING (a property recognizable for Finnish in the generative literature as well: cf. Holmberg 1989: 113). Similarly, the K T V-field configuration could also apply to many languages with a certain amount of freedom of constituent order.

2.4.3.2.   The Slavonic languages

These languages are of great interest to the study of the pragmatic typology of word order because the basic SVO order is interrelated with numerous possibilities of constituent movement along the linear axis.[107] Amongst the languages of Europe, they seem to be those with greater flexibility of WO, in both unmarked and marked sentences. The multiplicity of factors (semantic, prosodic, textual) which influence WO in these languages has already been emphasized in the Praguean literature. Jakobson (1963) observed that in the Slavonic languages "recessive" alternatives to the dominant order are numerous. All six orders SVO, SOV, VSO, VOS, OSV, and OVS may occur, although with different frequencies. Jakobson stressed that the conditions governing occurrence of the six variations are not grammatical in nature. Only one of the variations is stylistically neutral, while the others represent differing shifts of emphasis. Two studies carried out on Russian, using spectro-acoustic techniques of analysis, have shown the range of WO possibilities (Fougeron 1989) and the interaction of factors which bear on them (Yokoyama 1986).[108] They have also revealed some problems in Jakobson's and other authors' conclusions with respect to the concepts of "emphasis" and "expressiveness."

As far as the FOCUS position is concerned, a distinction must again be made between the unmarked and the marked position. In recent functionalist literature, the sentence-final position has been described as that of unmarked FOCUS in Polish (cf. Siewierska 1988: 123 ff.), in Czech (cf. Kim 1988), and in Bulgarian (cf. Stanchev 1987). The picture with respect to marked FOCUS appears to be more complex. It has been claimed that in Serbo-Croatian the immediately preverbal position may be used for the constituent in FOCUS.[109] It should be noted, however, that this is not the same property as that found in Turkish and Basque, since the immediately preverbal position is a subset of the positions accessible to the constituent in FOCUS. Besides, WH-pronouns are located in sentence-initial position.[110] As far as the other Slavonic languages are concerned, Czech and Russian are most frequently described. These descriptions, however, are not immediately comparable, being the product of different theoretical frameworks and different analytical methodologies. It would therefore be useful here to deal first of all with the phenomena in a way which closely follows the various pieces of research that have been carried out and only afterwards try to draw some more general conclusions.

In Russian, marked focalization of a constituent may be realized by either one or by both of the following possibilities:

398   Rosanna Sornicola

(a) movement of stress along the suprasegmental axis;

(b) movement of the constituent along the segmental axis.

The first possibility needs to be examined in detail, since it is crucial for the pragmatic organization of the sentence in Russian. From different perspectives, both Yokoyama (1986) and Fougeron (1989) have critically discussed the view that the neutral position for sentence stress may be either configurationally determined or may be obligatorily placed in the final position. In this model, any movement of stress away from such a position would have the consequence of "assigning prominence" to the constituent on which the stress falls.[111] Fougeron agrees with Nikolaeva, according to whom accentual prominence is not necessarily linked to expressiveness or emphasis. Such a phenomenon is distinguished from the occurrence of stress in final position because it takes the information beyond the immediate communicative context of the sentence (Fougeron 1989: 232; for examples, see here below). For Fougeron, then, initial nuclear stress is an indication of communicative reorganization (Fougeron 1989: 207).

It must be noted that while Fougeron makes a distinction between sentences with a theme-rheme division, entirely rhematic sentences and sentences with a marked rheme, this subdivision poses problems for an approach to comparative typology as the one followed in the present work: what are called "marked sentences" here seem to recur in each of the categories proposed by Fougeron.[112]

Yokoyama (1986: 180) also takes a stand against the idea that all the rhemes which are not in final position must necessarily be accompanied by a certain expressiveness of the utterance. The author distinguishes two types of intonation, the first characterized by the absence of sentential stress and by the fact that "the underlying tones are implemented on the strings, regardless of the location of various kinds of knowledge, and simply from left to right" (Yokoyama 1986: 190–191); the second by the presence of sentential stress. The latter is defined as "that stress which marks the knowledge item that would occur in utterance-final position, were the same sentence to be uttered with intonation Type I instead" (Yokoyama 1986: 191). The first type of intonation appears in neutral, context-free variants, while this is not the case for the second type.

As far as constituent movement on the segmental axis is concerned, the theoretical question of the relationship between movement of stress and constituent movement is dealt with in different ways by both Yokoyama (1986: 191) and Fougeron (1989: 205). The relationship that Yokoyama draws between intonation type and WO is important. The two options identified by Yokoyama in Russian could also be extended to other languages which allow focalization through constituent movement forwards on the linear axis, such as the Romance languages or, in a different way, Hungarian (cf. the notion of "eradicating prosody" discussed for Hungarian in Section 2.3.2). In fact, "when an item is moved to utterance-initial position, the result is obviously a monosyntagmatic intonational structure with sentential stress at its left

edge" (Yokoyama 1986: 194). Besides, "in all instances of the absence of sentential stress, WO is available for the non-propositional transfer of information" (Yokoyama 1986: 198). On the other hand, "WO . . . must be meaningful not only in Type I utterances as a whole [i.e., utterances without sentential stress], but also in the choice of the landing site of the preposed item in Type II utterances [i.e., utterances with sentential stress] and within all those structures of segmental material that bear no sentential stress before and after the word that bears it as well" (Yokoyama 1986: 199).

Nonetheless, the very WO structures with which Yokoyama exemplifies the two types of intonation make it clear that the WO associated with utterances with Type I intonation are neutral, while those with WO associated with a melodic contour of Type II are marked or "expressive":[113] consider the difference between the reply (26a) (with Type I intonation and WO following functional sentence perspective) and the replies (26b) and (26c) (with Type II intonation and WO in which the rheme is moved either to initial position or to a position before the verb):

(25) Kto napisal Evgenija Onegina?
who:NOM write:PST:M.SG Evgenij:ACC Onegin:ACC
'Who wrote Evgenij Onegin?'

(26) a. Evgenija Onegina napisal Puškin.
Evgenij:ACC Onegin:ACC write:PST.M.SG Puškin(NOM)
'Puškin wrote Evgenij Onegin.'

    b. Puškin Evgenija Onegina napisal.

    c. Evgenija Onegina Puškin napisal.

In both (26b) and (26c), sentential stress falls on the constituent which has been moved forward. The translation in both cases could be 'It was Puškin who wrote E. O.'.

The situation described by Fougeron, although primarily concerned with prosodic analysis, nonetheless provides very useful data for an examination of WO variation. Of particular interest are data from a corpus of spontaneous speech containing around 1000 utterances together with the results from comprehension tests carried out with a group of native speakers. Therefore, although the author's interpretation of pragmatic phenomena seems sometimes disputable (cf. here n. 343), the following discussion will be based on Fougeron's data and analysis.

Fougeron points out that, according to specific pragmatic conditions, the orders

SVO, SOV, OSV, VSO, VOS, OVS

all with nuclear stress on S, are "normal" in the sense of being "nonexpressive."[114] The orders VSO and VOS appear exclusively in narrative contexts.[115] Of particular interest to us are structures with O preposed to V, as in (27) and (28). They display certain properties peculiar to Russian:

(a) an O constituent may be moved from its canonical position with a high degree
     of freedom;
(b) nuclear stress may not be placed on it.

Consider (27):

(27)   ['Why are you making coffee today?']

          Mámočka   kofe   prosit.
          mother:DIM:NOM.SG   coffee:ACC.SG ask:PRS:3SG

          'Mother is asking for it.'[116]

Here *kofe* is a GIVEN element, while *mámočka* is NEW. This latter constituent, which
carries the nuclear stress, is obviously an implicit corrective contrastive FOCUS ('it
is mother who asks for coffee') with respect to an underlying conversational impli-
cature 'I or someone else wants the coffee'.[117]

   In a structure such as (28):

(28)   ['It's stuffy in here. Why don't you open the windows?']

          Otca   šum   bespokoit.
          father:ACC.SG   noise(NOM.SG) disturb:PRS:3SG

          'It's because of father, the noise disturbs him.'[118]

where O is in P_1, while the nuclear accent is carried by S, the O constituent is a
TOPIC [+NEW].[119] According to Fougeron, when it is preposed to the rhematic se-
quence, it always has a specific informational role which is greater than that of the
SOV structure cited earlier (cf. example (27)).[120] On the other hand, S expresses the
element of choice: "le sujet, porteur de l'accent nucléaire, est opposé à tous les autres
sujets eventuellement possibles qui sont rejetés en bloc" (Fougeron 1989: 297). This
interpretation confirms that *šum* is, according to the terminology used here, the FO-
CUS. The same pragmatic properties may be seen in OVS structures such as:

(29)   ['Come on, I've poured your coffee. It's not for me.']

          Kofe prosit mámočka.

          'It's mother who wants it [in the morning she cannot do without her cof-
          fee].'

(30)   ['So, you have allocated the roles? Yes.']

          Juru   čitaet   Serëža.
          Jura:ACC   read:PRS:3SG Serëža:NOM

          'It is Serëža who will play Jura [Volodja's voice is a little too soft].'[121]

   In sentences in which the element which occurs in first position (theme) is demar-
cated by a break in sonority (or a pause) on the right hand boundary, Fougeron speaks
of a "mise en relief" of the theme, which allows the introduction of a contrast *on the
level of the theme*. The author observes that the sentences in question represent the

Interaction of syntactic and pragmatic factors   401

first element of an opposition, generally introduced by an adversative conjunction. Here we see an example of S / / OV:

(31)   ['Shall we have this breakfast then? What do you want? Coffee? Tea?']
          Mámočka / / kófe prosit.
          ['Mother wants coffee and me, I'd prefer tea if possible.'][122]

Example (31) shows that Fougeron's analysis may be incomplete from a pragmatic point of view: it contains not only a contrast of theme (*mámočka* vs. *moi*), but the O constituent in preverbal position, on which the nuclear accent falls, constitutes a contrastive FOCUS ('coffee' vs. 'tea').

As far as marked FOCUS is concerned, it would seem then that in Russian movement of a constituent to a position preceding or following the in situ position can, but not necessarily must, be accompanied by the falling of the nuclear accent. The following generalizations with respect to marked FOCUS may therefore be drawn:

(a) If the nuclear accent falls on a constituent moved from its in situ position, such a constituent is the FOCUS.[123] If we call that which specifies the constituent moved from its canonical position the segmental parameter, and that which specifies on which constituent the nuclear accent falls the suprasegmental parameter, it may be said that in such a case, the values of the segmental parameter and of the suprasegmental parameter are isomorphic, that is, they determine the same constituent.

(b) If the nuclear accent falls on an in situ constituent and there is at least one other constituent in the sentence which is moved from its in situ position (in other words, if the values of the segmental and suprasegmental parameters are not isomorphic), the FOCUS is the element on which the nuclear accent falls. In such a case, the prosodic parameter of stress movement takes precedence over the segmental parameter of constituent movement along a scale of accessibility to FOCUSHOOD.

Characteristic (b) clearly distinguishes Russian from the Romance languages, where an OSˈV structure would not be possible (cf. Italian **la carne LA MAMMA mangia* [lit. 'the meat MOTHER eats']). The closest structural types would be either an OʹSV pattern (*LA CARNE la mamma mangia*) or a pattern with left dislocation of O (O S Pro V) (*la carne LA MAMMA la mangia* [lit. 'the meat MOTHER eats it'] 'it is MOTHER who eats the meat').[124]

The fact that the constituent on which sentential stress falls is the strongest candidate for FOCUS may also be seen from structures X / / Y Zʹ, such as:

(32)   ['Your letter gave me so much pleasure.']
          Pis'm-o / /   napisal   BRAT.
          letter-ACC.SG write:PST:M brother:NOM(SG)
          'The letter, it's my brother who wrote it [me, I just added the card].'

402   Rosanna Sornicola

(33)   ['We could not let such an opportunity to buy a flat go by. But we didn't
         have enough money for the down payment so we decided to ask our father
         for it.']
         Brat // napisal PIS'MO.
         'My brother therefore wrote a letter [where he told him how we came
         upon this opportunity].'

or from structures X Y // Z′, such as:

(34)   ['Did the birthday go well?
         Very well! It's true we had trouble getting the presents, but we finally
         found something:']
         Katja   podarila //   SUMKU.
         Katja:NOM give:PST.F.SG bag:ACC.SG
         'Katja gave a bag [I gave a shawl and Andrej managed to find a pressure
         cooker].'

(35)   ['I liked your present very much. You had a good idea. I've wanted a bag
         like that for ages.']
         Sumku   kupila //   MAMA.
         bag:ACC.SG buy:PST.F.SG mother:NOM.SG
         'It's mother who bought the bag [me, only the shawl].'

(36)   ['How shall we sort out the shopping?']
         Sumku mama // KUPILA.
         'The bag, mother bought it [You no longer need to concern yourself with
         it].'[125]

In (32)–(33) the element carrying nuclear stress occurs in structure final position,
while in (34)–(35) it occurs in the extrasentential tail position. Bearing in mind these
structural differences, it may nonetheless be said that in all four sentences Z′ is the
FOCUS.[126]

   It would therefore seem that in Russian, the rule which determines marked FOCUS
is:

          Associate FOCUS with the nuclear stress.

In other words, Russian seems to display autonomy[127] and predominance of the
prosodic over the syntactic configuration in determining the PF FOCUS.[128]

   These conclusions coincide only partially with those of a recent study by Holloway
King, according to whom

     under neutral intonation, i.e. non-emotive sentences, the WO is strictly
     determined by the discourse function organization of the clause: topics

precede the verb and discourse-neutral material, which precede the FO-
CUS. In emotive sentences, i.e. sentences with emphatic or sentence
stress, the WO appears to be less constrained because focused con-
stituents need not appear sentence finally. However, these non-final foci,
which are contrastive in meaning, are always marked with sentence
stress and generally occur immediately before the verb, following the
preverbal topics. (Holloway King 1995: 93)

The latter point, in fact, would make Russian a language with marked FOCUS placed
in the immediately preverbal position, while here it has been assumed that Russian
fits into the typology described in Section 2.4.1.[129]

2.4.4.   SVO languages with weak word order flexibility

2.4.4.1.   The Romance languages

The Romance languages discussed below (Italian, Spanish, French) have a basic
SVO word order pattern, and differ significantly with respect to their pragmatic
properties from the Slavonic and Balto-Finnic languages examined here. In compar-
ison with the latter languages, the Romance languages have a greater configurational
rigidity, which can be seen in at least two crucial properties:

(a) with the exception of a structural type which is quite rare and restricted to stylis-
tically rather elaborate written registers (cf. Section 5.2), a constituent may un-
dergo topicalization by movement into P_1 from a subsequent position only if
resumed by an anaphoric pronoun;

(b) with respect to the SVO configuration, in general any movement of a basic con-
stituent which is not linked to the topicalization process described in (a) gives
rise to a contrastive structure;[130] there is a prosodic correlate to this characteristic
in Italian and Spanish in that the moved constituent must carry the nuclear stress.
In order to make such movement possible in French, the moved constituent must
be inserted into the identifying construction with *c'est*.[131]

Although the SVO pattern is in itself a weak distinguishing characteristic (cf. Sec-
tion 2.4.2), the Romance languages may be said to share basic structural and prag-
matic affinities with respect to word order. In this matter, however, there are inter-
esting differences between French on the one hand and Italian and Spanish on the
other: (1) the relative order of the basic constituents is even more rigidly configura-
tional in written standard French compared to the other two languages examined; (2)
spoken French has more conspicuous phenomena with respect to clefting than spo-
ken Spanish or Italian. The first difference, which distinguishes the educated written
language, seems to be mainly determined not so much by internal typological factors
as by the external historical and sociolinguistic factors to which French syntax has
been exposed (cf. Sornicola 1995a).

404 Rosanna Sornicola

In the languages examined, sentences containing verbs with two arguments show a limited range of WO variation induced by pragmatic factors, both in terms of the types of structures allowed and of the frequency with which they occur.

2.4.4.1.1. Italian and Spanish. In the Italian sentences (37)–(42), only (37) is an unmarked structure, while the orders in (38)–(40) are possible only if the nuclear stress falls on O, and O, which is clearly a marked FOCUS, has a contrastive interpretation; (41) is acceptable as a neutral sentence only in a poetic style, while in everyday speech it is acceptable only if S carries the nuclear stress and if this constituent is a contrastive FOCUS; similarly, (42) is unacceptable as a neutral sentence, but does allow an interpretation with marked FOCUS:[132]

(37) Mario ha aiutato Lucia. SVO
       Mario have:PRS.3SG help:PST.PART Lucia
       'Mario helped Lucia.'

(38) Mario LUCIA ha aiutato. SOV
       'Mario helped LUCIA [and not someone else who has been mentioned in the preceding discourse].'

(39) LUCIA ha aiutato Mario. OVS
       'It is LUCIA that Mario helped.'

(40) LUCIA Mario ha aiutato. OSV
       'It is LUCIA and nobody else that Mario helped.'

(41) Ha aiutato MARIO Lucia. VSO
       'It is MARIO that helped Lucia [and not someone else who has been mentioned in the previous discourse].'
       (Corrective interpretation)

(42) Ha aiutato Lucia MARIO. VOS
       'It is MARIO [and nobody else] who helped Lucia.'

Some of the sentences reported may possibly be marginal in the spoken language. In fact, some of the speakers to whom the sentences were submitted for judgement, commented that (40) is less acceptable than (38) and (39), and that (42) is less acceptable than (41). They claimed that they would prefer *Lucia l'ha aiutata MARIO* to (42), but that they would also prefer *L'ha aiutata MARIO, Lucia* to (41).

Nevertheless, the constituent on which the nuclear stress falls is, as can be seen, that which does not occur in its canonical position (cf. (38), (40), (42)). Note that if the constituents that are not in situ are both S and O, as in (39), it is O in P_1 that must carry the nuclear stress. More problematic is the case of (41), where O is not adjacent to V on the right-hand side. Comparison of (40) and (41) could lead one to assume that the nuclear stress coincides with the constituent which occupies the

field which is not its own rather than with the constituent not in situ. The problem seems to be linked to the basic constituents, and to the potential ambiguity which arises when they occupy a field that is not the canonical one: cf. in fact *Sandro aveva portato orchidee gialle a Maria* 'Sandro had brought yellow orchids for Maria' and *A Maria Sandro aveva portato orchidee gialle* 'For Maria Sandro had brought yellow orchids' where in the second structure, IO is in P_1 but does not necessarily carry the nuclear stress.

From a structural point of view, in sentences (39), (40), and (41) the O constituent occupies the extreme edge of the sentence, to the left or to the right, while in (42) it is S which occupies the extreme right hand edge. Even if the intonational pattern described does not show pauses or changes in intonation contour,[133] it is possible that these types are the result of a process moving a constituent in a nonargument position.

From a semantic point of view, sentences (38)–(42) have an interpretation in terms of exclusive contrast: in (38)–(40) the contrast is realized by O, in (41)–(42) by S.[134]

Furthermore, in all the cases mentioned, the sentences must be associated with a particular context in order to work. (38) works only with the presupposition 'Mario helped someone other than Lucia' and therefore has a corrective value. If the corrective interpretation is obligatory for (38), it is optional for (39) and (40). (41) and (42) have the same value of exclusive contrast, this time relative to S. However, (41) also has a corrective value, which is not necessarily associated with (42). This last sentence may instead have a characteristic identifying function, textually typical of progressions featuring the addition of information,[135] and may be paraphrased with a pseudo-cleft structure such as *chi ha aiutato Lucia è MARIO* 'who has helped Lucia is MARIO' (cf. the possible context: ['ora ti dirò la verità'] *Ha aiutato Lucia MARIO* ['now I'm going to tell you the truth'] lit. 'has helped Lucia MARIO').

The situation with respect to WO in the Romance languages is further complicated by the property relating to topicalization processes, summarized in (a) in Section 2.4.4.1. It is exemplified by the structures in (43)–(44), which are typical examples of left dislocation:

(43)   Lucia / l'   ha   aiutata (/)   Mario.
         Lucia   she:ACC have:PRS:3SG help:PST.PART Mario
         'As to Lucia, the one who helped her was Mario.'

(44)   Mario / Lucia / l'ha aiutata.[136]
         'Mario helped Lucia.'

Note that these sentences do not generally allow the nuclear stress to fall on referential constituents.[137] In particular, it may be said that in left dislocations a contrastive FOCUS on the dislocated constituent (i.e., on O) is never allowed, unless it carries a FOCUS of exclusive contrast (see Section 2.4.4.3).

406   Rosanna Sornicola

An obligatory resumptive pronoun in topicalization processes is only one aspect of the greater configurational rigidity of the Romance languages compared to the Slavonic and Balto-Finnic languages. A basic constituent may not, in fact, be moved to the extreme edge of the sentence without leaving behind it a pronominal copy which serves as a case marker. This property is shared by Modern Greek (cf. De Simone Brouwer 1921: 217–221; Tsimpli 1995), a language in which WO varies considerably between the written and spoken registers. In fact, Italian and Spanish and Modern Greek also share another property, and that is that deviations of arrangement of basic constituents from SVO orders result in structures with FOCUS of exclusive contrast on the moved constituent (cf. *állin aɣapó* 'it is indeed another that I love [and not you or the one you thought]', cf. De Simone Brouwer 1921: 220).

The Spanish translation of sentences (38)–(42) (noted here as (38′)–(42′)) and the judgements elicited from several native speakers show that if structural isomorphism is maintained, only (38′), (41′), and (42') are grammatical but not (39′) and (40′):

(38′)   Mario a LUCIA ha   ayudado.
         Mario to Lucia   have:PRS:3SG help:PST.PART
         'Mario helped LUCIA [not somebody else].'
         (Contrastive FOCUS on O)

(39′) a. ??A LUCIA ha ayudado Mario.

(40′) a. *A LUCIA Mario ha ayudado.

(41′)   Ha ayudado MARIO a Lucia.
         'It is MARIO [and not somebody else] who helped Lucia.'
         (Contrastive FOCUS on S)

(42′)   Ha ayudado a Lucia MARIO.
         'The one who helped Lucia is Mario.'

The three grammatical sentences have the same semantic and prosodic interpretation as the corresponding Italian sentences. Interestingly, the grammatical translations of (39) and (40) cannot be structurally isomorphic with Italian, because in Spanish the O constituent may appear in P_1 only if it is resumed by an anaphoric pronoun, in other words only in a left-dislocation structure such as:

(39′) b.   A Lucia la ha ayudado MARIO.
            'The one who helped Lucia is Mario.'

(40′) b.   A Lucia Mario la HA AYUDADO.
            'Mario DID help Lucia.'

In (39′b)–(40′b) the O constituent in P_1 is a TOPIC. However, in Spanish it is impossible for structures with left dislocation of O to have FOCUS of exclusive contrast on this constituent.

The structural property that we shall call "strongly SVO" is also reflected at the surface level. Surveys carried out on different types of text in Italian have shown that the SVO order is by far the predominant one. The distributional frequency of the SVO pattern ranges between a lower threshold of 59 % and an upper threshold of 98 % according to the type of text:

Table 2. Percentage of the SVO pattern in different types of Italian texts

Type of text	% of SVO pattern
Scientific prose	98
Newspaper reports	87
Literary prose	78
Spontaneous speech	77
Thriller novels	66
Comic strips	59

Note: These percentages are taken from Sornicola (1994: 53) and from Como (1995: 56, 97, 133). They refer to main clauses.

Particularly interesting is the fact that all the values in the range are above 50 %, and that the average frequency (77.5 %) comes close to the frequencies for spontaneous speech and literary prose.

The percentages obtained for SVO order are surprisingly similar to those calculated by Hakulinen et al. (1980: 145) for Finnish: S precedes finite V in 61 % of all sentences and in 79 % of sentences containing S in standard written prose. This last figure shows a difference of only 1 % from that obtained from the corpus of literary Italian by Como (1995). This result shows the unsatisfactory character of concepts such as "language with free word order" vs. "language with non-free word order" of constituents, and confirms a uniform trend amongst SVO languages towards strong dominance of the pattern with S preceding V.

SOV order turns out to be almost always associated with a pronominal O, while OVS order, always contrastive, has in any case a frequency below 10 %; V-initial orders are either not represented at all (cf. Como 1995) or have a frequency of between 1–4 % (cf. Sornicola 1994: 31, who records VOS order in newspaper articles at 1.25 %; Como 1995: 56, for comic strips).

2.4.4.1.2. French. Structures isomorphic to (38)–(42) are not possible in French, not even as marked orders: a cleft structure with *Lucia* as a postcopular constituent in FOCUS, restricted to an "initial" position,[138] is the only construction which may correspond to (38)–(40):

408   Rosanna Sornicola

(43′)   C'   est LUCIA que   Mario a   aidée.
                 that is   Lucia   REL:OBJ Mario have:PRS:3SG help:PST.PART
                 'It is LUCIA that Mario helped.'

Cleft constructions with *Mario* as a postcopular constituent in FOCUS, equally restricted to an "initial" position, correspond to (41)–(42):

(44′)   C'   est MARIO qui   a   aidé   Lucia.
                 that is   Mario   REL:SBJ have:PRS:3SG help:PST.PART Lucia
                 'It is MARIO who helped Lucia.'

In other words, movement possibilities are more restricted in focalization processes in the written language: O may be moved into the preverbal field only in the construction with *c'est*; on the other hand, S may never move from the preverbal field to the postverbal field.[139] Note that in spoken French, an S/OV order with O "encapsulated" in the construction with *c'est* would be possible:

(44″) a.   Mario / c' est LUCIA qu'   il   a
                 Mario   that is Lucia   REL:OBJ he:NOM have:PRS:3SG
                 aidée.
                 help:PST.PART
                 'As to Mario, it is LUCIA that he helped.'

with the obvious result of structure segmentation.[140] The O/SV type given in (44'b) would be equally possible:

(44″) b.   Lucia / c' est Mario qui   l'   a
                 Lucia   that is Mario REL:SBJ he.CLT:ACC have:PRS:3SG
                 aidé.
                 help:PST.PART
                 'As to Lucia, it is MARIO who helped her.'

Note, however, that in French (especially in spoken registers) cleft structures of the kind described here seem to be undergoing a process of demarking. In such cases, the postcopular constituent no longer carries the PF FOCUS and the whole sentence functions instead at the broader textual level as an all-in-FOCUS, presentative structure (cf. Section 4.2; for similarity of this situation to that of Welsh, cf. Shisha-Halevy 1995: 154–155).

    As far as topicalization is concerned, movement processes are possible in the spoken language as right or left dislocation:

(45)   Moi / je   le   lui   ai
              I.EMPH I.CLT:NOM he.CLT:ACC she.CLT:DAT have:PRS:1SG
              donné   le livre.
              give:PST.PART the book
              'As to me, I gave it to him, the book.'

Interaction of syntactic and pragmatic factors   409

(46)   Cet   élève / je   l'   aime   bien.
        DEM:M.SG pupil   I.CLT:NOM he.CLT:ACC love:PRS:1SG well
        'This pupil, I love him much.'

As in Italian and Spanish, the dislocated constituents also have noncontrastive FO-
CUS in French types (45) and (46).[141]

   However, isolation (demarcation) of constituents is used more often in French than
in Italian or Spanish as the normal means of highlighting in the spoken language (see
further on, Section 6[142]).

2.4.4.2.   English

Interestingly, modern English conforms to the Italian and Spanish type rather than
to the French or Germanic type, at least with respect to leftward movement possibil-
ities.[143] A structure such as (47), typical of V-second languages, such as are many of
the Germanic languages (cf. Section 2.4.5) and, at least in certain respects, literary
Romance languages of the Middle Ages (cf. Renzi 1984; Salvi 1993, 1997), is con-
sidered by many native speakers of contemporary English as "archaic," "typical of
old poetic texts," and surely ungrammatical in present-day ordinary usage:

(47)   MARY loves John.   OVS

On the other hand, (48) and (49) would be possible as structures with marked FOCUS
on O and left dislocation respectively:

(48)   MARY / I love.

(49)   MARY / I love her.

Judgements elicited indicate that both (48) and (49) are possible in both the spoken
and written language (but (49) would be more typical of "informal writing" styles or
of spoken language simulated or recorded in the written language). Some judgements
also show that in (48) *Mary* is a marked FOCUS, with a contrastive value,[144] while
in (49) it is a TOPIC, without such values. Other judgements point to a text such as
(48a) as the typical context for (48):

(48)   a.   Mary I love / Sue I love / Jane I love / Ann I don't love.

Such a context would make it more plausible to consider *Mary* a contrastive TOPIC
rather than a contrastive FOCUS.

   The situation is, therefore, similar to that in Italian and Spanish as far as the left
side of the sentence is concerned. That the similarity is restricted to the leftmost edge
seems further confirmation of the fact that leftward movement processes for marked
focalization are more "natural" than rightward movement processes (see, e.g., the
impossibility of structures like (42) and (42′) in English).

410   Rosanna Sornicola

On the other hand, English and Romance languages such as Italian and Spanish group together by virtue of having a wider range of nuclear stress movement than French, as may be seen from the English examples (50a)–(50d), and from the corresponding Italian (51a)–(51d) and Spanish (52a)–(52d) examples.

(50)   a.   Sandy knows Peter.
       b.   SANDY knows Peter.
       c.   Sandy KNOWS Peter.
       d.   Sandy knows PETER.

(51)   a.   Sandy   conosce   Peter.
              Sandy   know:PRS:3SG   Peter
       b.   SANDY conosce Peter.
       c.   Sandy CONOSCE Peter.
       d.   Sandy conosce PETER.

(52)   a.   Sandy   conoce   a   Peter.
              Sandy   know:PRS:3SG   to   Peter
       b.   SANDY conoce a Peter.
       c.   Sandy lo CONOCE a Peter.
       d.   Sandy conoce a PETER.

However, the difference between English on the one hand and Italian and Spanish on the other appears to lie in the fact that English has an even wider range of structural possibilities associated with nuclear stress movement than Italian or Spanish, while the latter two languages have a wider range of segmental constituent movement than English.[145]

The difference between French and English may be seen in at least two characteristics. English does not have to use a clefting process in order to move O into preverbal position. On the other hand, cleft structures, normally associated with a process of focalization of a nonverbal constituent,[146] are not as frequent in English as in French.[147] Furthermore, the prosodic and syntactic structures interact in the two languages according to very different typologies. As has been pointed out by Daneš (1967: 227), the range of possible intonational patterns with marked FOCUS differs in English and in French: "in English … the highly fixed WO is compensated for by a great variety of the possible positions of the CI [= intonation contour] in the utterance." Cruttenden (1986: 146–147) has pointed out that the number of words in an intonational group is lower in French than in English; that is, intonational groups in French are shorter than those in English. Cruttenden emphasizes that this does not constitute a radical difference in structural options, but merely different uses of the available options. He justifies in an interesting way the peculiarity of French, observing that "French does not have the mobile nucleus characteristic of English

nor the same potential for pre-nuclear accents, and hence is forced to introduce extra intonation groups for the purpose of highlighting" (Cruttenden 1986: 147).

These remarks highlight a very important typological problem, which is that of the relationship between processes which move the stress nucleus and processes which move nonprosodic constituents. Cruttenden (1986: 150) claims that "it is almost always the case that languages which use nucleus movement also use WO variation, even if only infrequently (like English), whereas the reverse is not true." The relationships between the prosodic and the segmental levels do indeed seem to be more complex. As far as the languages of our corpus are concerned, one can sum up the results which have emerged with the following sketch:

– English: wide freedom of nucleus movement. Relative freedom of nonprosodic constituent movement (to the left edge of the sentence). Possibilities of using only nucleus movement without WO variation. If, however, there is WO variation, nucleus movement onto the moved constituent is necessary;
– Italian and Spanish: possibilities of using only nucleus movement without WO variation. Moderately free nonprosodic constituent movement to both the left and right edge of the sentence. If, however, there is WO variation, nucleus movement onto the moved constituent is necessary;
– French: no nucleus movement. Very reduced possibilities of nonprosodic constituent movement. Use of isolation and structures with *c'est* for focalization processes;
– Russian: wide freedom of nucleus movement. Wide freedom of nonprosodic constituent movement. Absence of isomorphism between the prosodic and segmental axes: possibilities of using only nucleus movement without WO variation; if there is WO variation, nucleus movement on the moved constituent is not necessary.

2.4.4.3.   Some remarks on Italian and Spanish

On the basis of what has been said so far, it would seem that, unlike Russian, where movement of a constituent from its in situ position does not necessarily imply a contrastive interpretation, but rather a change in textual functions (cf. the constituent *otca* in example (28)), in Italian and Spanish such movement may occur only if the moved constituent is a marked FOCUS (cf. the Italian example (40)). Furthermore, unlike in the Slavonic languages, where the moved constituent does not necessarily carry the nuclear stress, in Italian and other Romance languages the moved constituent always coincides with the nuclear stress, except in very rare cases.

The situation in Italian and Spanish may be represented schematically as follows:

Unmarked distribution of FOCUS: this occurs when the FOCUS coincides with P_n, as in Figure 8.
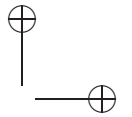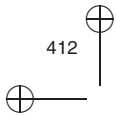
412   Rosanna Sornicola

P$_1$ …   V   …   P$_n$

S   O

T   F

Figure 8. Unmarked distribution of FOCUS

The left dislocation type, *La carne la mangia MARIO* [lit. 'the meat it-ACC he-eats MARIO'], is represented in Figure 9.

P$_1$ …   CLT   V   …   P$_n$

O   S

T   F

Figure 9. Left dislocation type

As can be seen, in SVO structures without a stress nucleus and in left dislocation structures the relationship between position and PF remains unchanged. Similar considerations apply to structures with a double TOPIC (cf. for example, *La carne Mario la mangia volentieri* [lit. 'the meat Mario it-ACC he-eats gladly']). This is shown in Figure 10.

P$_1$ P$_2$ …   CLT   V   …   P$_n$

O   S

T$_1$ T$_2$ F

Figure 10. Structure with a double TOPIC

P$_1$ V

O

F

Figure 11. Marked distribution of FOCUS
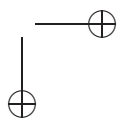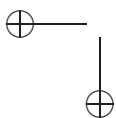
P$_1$ …   P$_i$ V

IO   O

T   F

Figure 12. Marked distribution of FOCUS

Marked distribution of FOCUS (as in the type: *La CARNE mangio* [lit. 'the MEAT I-eat'], with the implication 'but not any other food') is represented in Figure 11. This is, in fact, a structure with FOCUS of exclusive contrast, which implies a second element of comparison. It can occur in Italian, but not in Spanish (see Section 2.4.4.1.1). If this structure is compared with the corresponding Russian ones (cf. Section 2.4.3.2), it can be seen that the latter have no contrastive value.

Marked distribution of FOCUS (as in the type: *A Maria la CARNE ho dato* [lit. 'to Maria the MEAT I-have given'] with the implication 'not any other food') is also represented in Figure 12.

The picture drawn so far of the differences between SVO languages with highly flexible order and SVO languages with weak WO flexibility should not obscure the basic syntactic and pragmatic structural similarities which exist between all SVO languages. A crucial property serves as an example; that is, the awkwardness/low frequency of V in P$_1$. Such a configuration, impossible with Vs with two arguments in modern English and French, is extremely rare in Italian and Spanish and is also very sporadic in the Slavonic languages.[148]

The differences between the languages in question are often attributed to the presence/absence of morphological case. However, such a factor does not in itself seem to guarantee a greater freedom of WO: consider Modern Greek, a language that has morphological case but behaves like the Romance languages, which do not have case, with respect to WO. On the other hand, in Russian as in other Slavonic languages, SVO is by far the most predominant order (cf. Holden and Krupp 1987; Fougeron 1989).

2.4.5.   V-2 languages

2.4.5.1.   The role of scientific traditions in the formation of the V-2 concept

The traditional approach to the study of WO in the Germanic languages has been highly influenced by theoretical models linked to specific historical developments of linguistic ideas. Investigation of WO properties in such languages cannot therefore be separated from a preliminary examination of the traditional views on this (cf. Scaglione 1981: especially 108–180).

414   Rosanna Sornicola

Topological models which identify sentence spaces which are characterized by special position/type of constituent relationships have taken on great importance. Diderichsen's (1936a, 1936b) investigation on Danish and Drach's ([1937] 1963) on German show how during the 1930s in German linguistics, theories of WO developed which had certain characteristic principles in common: (a) the idea that every sentence could be reduced to an abstract schema, consisting of a fixed number of positions, and in particular (b) the idea that the basic element in the sentence was the finite verb and that all the other elements were structurally and spatially defined in relation to it.[149]

Although the second idea rests after all on assumptions of a functional nature,[150] it has converged, especially for German, with WO studies in the neogrammatical traditions, which in the first place turned to examination of structural regularities.[151] Within these traditions, conceptions developed in the Indo-European framework, which favored examination of the relationship between accentual factors and WO, have had particular importance for the origin of the concept "verb-second." The more general notion of "intermediate position" (*Mittelstellung*) of the verb has been substituted by that of "second position" (*Zweitstellung*), identified as the position occupied by unstressed elements, particles, pronouns, and the verb itself.[152]

These ideas have generally set apart discussions of Germanic languages[153] and have contributed to the characterization of a linguistic type defined as "verb-second." Moreover, in the historical origins of this concept, a weakness may be recognized which also seems to apply to some modern theories constructed with respect to the V-2 type: a WO property characteristic of modern phases and of particular registers of certain languages becomes absolute, diachronically (in comparison with properties in phases of Indo-European languages going back several millennia) or synchronically (in describing it as the basic property of WO in main clauses in Germanic languages).

2.4.5.2.   The Germanic languages

The scientific traditions referred to have influenced and perhaps biased the analyses of empirical data, at least those relating to the modern phases of the Germanic languages. In fact, in main declarative clauses, all contemporary Germanic languages, except English, have finite V occurring in what may be reliably defined as the second position in the sentence. The second position may be reliably defined because: (a) in P_1 one and only one constituent may occur (for some apparent exceptions, see further on); (b) whatever the type of constituent which occurs in P_1, the inflected verb always occupies position P_2. In other words, while the argument constituents such as S and O may occur in various positions, V has a fixed position.[154]

Consider the following Swedish (53a)–(53b) and German (54a)–(54b) examples:

(53) a. Erik köpte   verkligen boken.
          Erik buy:PST really   book:DEF.SG
          'Erik really bought the book.'
      b. Erik hade   verkligen köpt   boken.
          Erik have:PST really   buy:PST.PART book:DEF.SG
          'Erik had really bought the book.'[155]

(54) a. Er   liest   das Buch heute.
          he:NOM read:PRS:3SG the book today
          'He is reading the book today.'
      b. Er   hat   das Buch gestern   gelesen.
          he:NOM have:PRS:3SG the book yesterday read:PST.PART
          'He read the book yesterday.'[156]

What the Swedish and German examples have in common, independently of the differences in the VP structure,[157] is the fact that finite V occurs in P_2. Such a property does not only concern cases of SV order such as (53a)–(53b) and (54a)–(54b). In fact, characteristics (a) and (b) just mentioned define a phenomenon which is peculiar to V-2 languages, which is S inversion: since only one constituent may occur in P_1, if P_1 is occupied by an NP = O or by a PrepP or an AdvP, S will occupy the immediately postverbal position P_3, as can be seen in the following examples from German, Dutch, and Swedish, respectively:

(55) Das habe   ich   nicht gesagt.
      that have:PRS:1SG I:NOM NEG   say:PST.PART
      'I haven't said that.'

(56) Gisteren   kocht   Wim dat   boek.
      yesterday buy:PST Wim that book
      'Yesterday Wim bought that book.'[158]

(57) Den boken   köpte   Erik i   London.
      that book:DEF buy:PST Erik in London
      'That book, Erik bought in London.'[159]

The situation may therefore be represented by means of the following correspondence between position and the constituents of a sentential configuration:

$$P_1 \quad P_2 \quad P_n$$
$$XP \quad V_{fin} \quad XP$$

where XP stands for "any phrasal category" (cf. Sigurðsson [1989: 5], who however has a quite different formula). Alternatively, it is possible to arrive at a representation such as

$$XP \quad V_{fin} \quad \ldots$$

416   Rosanna Sornicola

which leaves the postverbal part of the configuration unspecified.[160] It is clear, therefore, why the V position may be taken as the absolute parameter from which the positions of the other constituents may be determined.

2.4.5.3.   Vorfeld, Mittelfeld, Nachfeld

In many topological models the *Mittelfeld* is defined as the particular field of finite V, on the basis of which the *Vorfeld* and *Nachfeld* are defined as, respectively, the field that precedes and the field that follows the verb.[161] This spatial subdivision is only partly comparable with the linear representation by sequence $P_1 \ldots P_n$, which creates a problem when languages traditionally described by topological models are compared with languages which follow different descriptive traditions. In particular, the lack of a suitable means of comparison is more acute with regard to the sentence space which extends from V to P_n.

In this discussion we shall follow Engel's topological model of German sentence, which is built upon the notion of "predicative space" (*prädikativer Rahmen*). This is defined as the space occupied by the verbal complex (which may consist of a finite verb [as in (54a)] or a finite and a nonfinite verb [as in (54b)]), plus the various complements and/or modifiers of the verbal complex. The space within the "predicative space" is termed *Mittelfeld*, while the spaces before and after the whole predicative space are termed *Vorfeld* and *Nachfeld*, respectively.[162] Note that in this definition the latter notion coincides with the space beyond the "framed" part of the sentence.

While the *Mittelfeld* is relatively flexible with respect to the type of constituent which may occur in it (pretty much every type of constituent may be found in it[163]) both the *Vorfeld* and the *Nachfeld* are restricted in this respect. If the set of elements which may occur in P_1 (i.e., in the *Vorfeld*) coincides in part (but not completely) with the set of elements which occur in the *Mittelfeld*,[164] the already noted condition which excludes more than one constituent in P_1 constitutes a powerful constraint.[165]

As far as the *Nachfeld* is concerned, the set of its characteristic elements is very limited. Normally[166] only prepositional "complements" (*Ergänzungen*) and circumstantial elements (*Angaben*) may in fact occur there.[167] As in many other languages, this field may have the double function of being the location for the addition of information, which not uncommonly assumes the form of an afterthought (TAIL) structure[168] especially in spontaneous spoken speech, or the location of focalization. The double interpretation may be seen in the example:

(58)   Ich   habe   mich   geärgert   über   diesen
        I:NOM   have:PRS:1SG   I:ACC   annoy:PST.PART   because.of   this:ACC
        Typ.
        fellow
        'I got angry (/) because of this fellow.'

The afterthought interpretation of the constituent in the *Nachfeld* applies if pause occurs, while the FOCUS interpretation of such constituent is to be associated to the structure without pause. However, unlike the Romance languages, where FOCUS in P_n is unmarked, in German FOCUS in the *Nachfeld* is stronger than FOCUS in the *Mittelfeld* (on the question of unmarked and marked FOCUS, see further on):

(59)   Ich   habe   gekündigt   aus diesem   und keinem
        I:NOM have:PRS:1SG quit:PST.PART for that:SG.DAT and no:SG.DAT
        anderen   Grunde.
        other:SG.DAT reason:SG.DAT
        'I quit for this and no other reason.'

(60)   Ich habe aus diesem und keinem anderen Grunde gekündigt.

When making comparisons with other languages, it is of course necessary to give due consideration to differences in the type of constituents which may occur in this part of the sentence. For example, the Romance languages do not have the categorial constraints shown by German.

The position of V for spatial demarcation of the sentence is even more crucial in languages with *Satzklammer*, such as German and Dutch. In examples such as (54b), or (61):

(61)   Bei uns   hat   es schon   wieder   Spaghetti
        by we:DAT have:PRS:3SG it already again   spaghetti
        gegeben   heute.
        give:PST.PART today
        'We have had spaghetti again today.' (cf. Engel 1991: 316)

the auxiliary and the V participle form a sort of frame which defines the VP area (*verbaler Rahmen*).[169]

2.4.5.4. Some non-V-2 structures

It should be noted, in any case, that if the second position for V is dominant in other Germanic languages, this is not, the situation in older diachronic phases of the languages under examination. Furthermore, even in modern phases, various types of exceptions surface here and there, which may be regarded as traces of previous stages.

V in P_1 in modern German is a characteristic of some dialects and some spoken registers, as had already been documented by Behaghel (1932: 38). The popular character of the type is also linked to its presence in the *Volkslieder* and in some poetic texts which are modeled on it, such as Goethe's *Heideröslein*:

(62) a. Sah   ein Knab'   ein Röslein stehn.
          see:PST(3SG) a boy:SG.NOM a rose:DIM stand:INF
          'A boy saw a little rose.' (cf. Behaghel 1932: 38)

418   Rosanna Sornicola

It is interesting that in the vast majority of cases the type seems to occur with single argument verbs, such as intransitives of movement or saying, the verb 'to be' with locative-existential meaning (cf. Behaghel 1932: 38),[170] although a few examples with transitive verbs are found:

(62) b. Tut   doch   kein   Mensch auch nur   den Mund
          open:PRS:3SG however no:SG.NOM person also only the mouth
          auf.
          PREV
          'But nobody even opens their mouth.' (Behaghel 1932: 38)

On the other hand, in contemporary spoken German V-initial orders are possible, as a result of omission of the subject pronoun:

(63)   Habe   es schon   gelesen.
          have:PRS:1SG it already read:PST.PART
          '(I) have already read it.' (Haider 1986: 56)

It has been noted with respect to sentences such as this, particularly in the generative framework, that they are related to surface representation which do not alter the underlying V-2 configuration of the sentence.[171] However, the existence of this type certainly makes the idea of an absolute V-2 character problematic.[172]

The type with V in P_1 is also found in literary prose in modern Icelandic, and in Swedish where it is typical of oral narration (it is defined as "narrative inversion"). The following are Icelandic examples:

(64) a. Gaf   ég   honum bókina.
          give:PST(1SG) I:NOM he:DAT book:DEF:ACC
          'I gave him the book.' (cf. Pétursson 1978: 138)
      b. Tog   han   sen foto.
          take:PST(3SG) he:NOM then picture
          'Then he took a picture.'[173] (cf. Platzack 1986: 47)

The Icelandic and Swedish type is today typical of literary prose; it does, however, seem to have a correspondence in older language states (for Old Icelandic, cf. Freeman 1937: 102–103). A similar dynamic may be observed in German, where the phenomenon is found in literary prose throughout a period of many centuries, although with variation from writer to writer,[174] and surfaces here and there in dialects and in contemporary spontaneous speech. Again, it may be deduced from this that the V-2 character is anything but absolute in the Germanic languages.

Similar conclusions are arrived at particularly in cases where V occurs in positions other than P_2 (i.e., P_3, P_4, P_5) in Swedish and Icelandic. The following are Swedish examples:

(65)   Erik kanske kan svara   på din   fråga.
        Erik maybe can answer:INF on your:NNT question:SG
        'Maybe Erik can answer your question.'

(66)   Kanske Erik verkligen har   köpt   bok-en.
        maybe Erik really   have:PRS buy:PST.PART book-DEF.NNT.SG
        'Maybe Erik really has bought the book.'

(67)   Nu kanske jag svarade   fel   på den   här
        now maybe I   answer:PST wrong on DEM.NNT.SG here
        fråg-an.
        question-DEF.NNT.SG
        'Now, maybe I answered wrong on this question.'

(68)   Då   kanske han inte bryr   sig om   att stanna.
        then maybe he NEG bother:PRS REFL about to stay:INF
        'Maybe he does not bother himself to stay, then.'
        (cf. Platzack 1986: 29; for Swedish Sigurðsson 1989: 5, n. 2. See also
        Thráinsson 1986)

According to Platzack, (65) can be considered a V-3 structure, (66) and (67) V-4 structures, (68) a V-5 structure. The type in (65) usually occurs when the adverb *kanske* 'maybe' is present. At first sight it seems to call into question the significance of the exception. *Kanske* is in fact a fossilized adverb (< Aux *kan* 'can' + *ske* 'happen'; cf. Eng. *maybe*, Fr. *peut-être*[175]), which complicates computation of the position within an analysis which takes account of the dynamics of grammaticalization. Platzack (1986: 47, n. 4), however, points out that "occasionally other adverbs may occur in front of the clause, especially in casual speech."

These cases may be compared with the German ones previously mentioned as being problematic for the V-2 theory (cf. here n. 165).

2.4.5.5. German

2.4.5.5.1. Preliminary observations.   German exhibits specific WO patterns compared to other languages of Europe (and, as has been seen, even compared to other Germanic languages), as far as both neutral structures and variations induced by the marked distribution of PFs are concerned. However, to be able to effectively operate a typological comparison with other languages, it is appropriate to distinguish spontaneous or natural phenomena in the spoken language from those which are determined as a result of cultural processes which have influenced the written language. In fact, more than other languages which have been discussed, German shows a sociolinguistic split between the written and the spoken language in matters of WO (cf. also Sections 3 and 6).

420   Rosanna Sornicola

The problem of determining basic order in German, which has characterized the literature of the 1960s and 1970s, has been reformulated in recent developments in generative grammar in terms of V movement from a basic SOV representation. In other words, it is claimed that the order in dependent clauses should be assumed to be the basic one (cf. Haegeman 1991: 537 ff. for an overview of the problem and of the opinions for and against this idea). Such a conclusion is justifiable within a theoretical framework which must account for differences in structure in terms of unitary underlying configurations, in other words, which must trace variants back to underlying invariants. However, such an approach seems more useful for synchronic description than for an explanation of the historical and pragmatic dynamics of the phenomena under investigation. It must, in fact, be stressed that the metacriteria which subtend this model (such as descriptive simplicity) are of a formal rather than historical nature. Furthermore, the very nature of the theory, oriented towards representation of synchronic structures according to essentially deductive and analytic models, makes it difficult to account for linguistic characteristics which are, for the most part, the result of historical processes, in which factors of internal development interact in a complex way with external factors (cf. further on, Section 2.4.5.5.5).

As has been said in Section 1, the perspective of this work is centered on a dynamic model, in which the complexity of variation is considered to be the starting point for an explanation of the typological properties of a language. As far as German is concerned, the outcome of this is an attempt to account for the differences in WO patterns gathered from actual texts in terms of historical motives. This seems all the more indispensable, since German exhibits structural peculiarities of WO which are in part due to the influence of prestigious Latin models (on this problem, cf. Section 2.4.5.5.5).

In the next sections, therefore, an attempt will be made to take into account the structural, pragmatic, sociolinguistic, and historical dimensions.

2.4.5.5.2.   Some structural problems.   A structural description of WO patterns in German must take into account firstly the dichotomy relating to the type of clause, whether it is main or subordinate, and secondly the dichotomy relating to the verbal form, whether it is a structure with a finite V or with an Aux + V participle. As is well known, in main clauses this second dichotomy gives rise to structures such as (69a) or (69b):

(69)   a.   Karl   liest   ein   Buch.
             Karl   read:3SG   INDEF   book
             'Karl is reading a book.'
       b.   Karl   hat   ein   Buch   gelesen.
             Karl   have:3SG   INDEF   book   read:PST.PART
             'Karl has read a book.'

Interaction of syntactic and pragmatic factors   421

Structure (69a) has O in postverbal position, while in (69b), O follows the inflected verbal element (the auxiliary), but precedes the nonfinite verbal element from whose lexical representation it receives its subcategorization. In this sense, comparison of the two structures should be expressed in terms of structural equality – as far as in both (69a) and (69b) right adjacency of O to the inflected element is concerned – and of structural difference with respect to right (cf. (69a)) vs. left (cf. (69b)) adjacency of O to the verbal element which subcategorizes it, rather than in terms of differences in the relative order of V and O (which relates exclusively to a topological representation).

A different question is that which relates to how the configurational character of the language is ascertained. Since the concept of configurationality has undergone redefinition in the course of the last few years, it will be worthwhile establishing the characteristics of German with respect to at least two different conceptions of configurationality. Regarding the typological properties which have just been considered as suitable for the definition of configurational languages (such as the S/O asymmetry, etc.), German may undoubtedly be considered configurational (cf. Abraham 1995: 606). Less clear is the conclusion that may be reached with respect to another parameter, traditionally associated with configurationality, that is, a higher level of positional rigidity of constituents.[176] The problem arises in particular for the *Mittelfeld*, since, as has been said, both the *Vorfeld* and the *Nachfeld* have a very reduced positional space (that is, a space which may accommodate only one structural position) and which has greater constraints on the type of constituent which may occur there.

As far as the *Mittelfeld* is concerned, Abraham (1986: 16, 34–35) claims that "under the current paradigm of configurationality, German is strongly, though not completely, nonconfigurational." While the NPs dominated by V" (DO, IO) have a lower configurational status, prepositional Os (PO) have a stronger one. For example, while the factor "definiteness" may alter the relative order of constituents carrying the GF of O or of IO with respect to the verb,[177] it has no bearing on the position of POs with respect to V (Abraham 1986: 16–18). More generally, only POs for which V is subcategorized have a position close to V which is not influenced by other factors (Abraham 1986: 34). Furthermore, Abraham shows that "from among the parameters of GRs, discourse relations such as topic and focus, and features of subcategorization, it is the latter properties that have the strongest, though not the only constraining force on the linear order" (Abraham 1986: 15).

In any case, in sentences without *Satzklammer* German has a neutral order of GFs:

S + (PO) + DO + IO

while in sentences with *Satzklammer*, a preference is shown for the order:

S + IO + DO + (PO)

422   Rosanna Sornicola

Of particular importance is the principle of "syntaktische Verbnähe, topologische Verbferne" (cf. Helbig and Buscha 1989: 569).[178] In fact, the order of nonverbal elements is conditioned by dependency. When the *Mittelfeld* contains a single complement or specifier, this follows the finite verb. When the *Mittelfeld* contains more elements, their order has mainly to do with the structure of the dependents. However, the position of the complements and of the specifiers cannot be determined with the same certainty as that of the verbs.

2.4.5.5.3.   The pragmatic problem.   Let us now attempt to determine the relationship between basic order, neutral order, and linearization of PFs, especially the function FOCUS: the picture that emerges seems even more complicated. In fact, linearization of FOCUS is not reducible to a unique position nor to a class of positions linked by a common topological property. In analyzing such a question, it is all the more appropriate to bear in mind initially the difference between prosodic correlates of FOCUS, such as nuclear stress, and syntactic and pragmatic correlates.

Abraham (1986: 34) observes that the preferred order S + IO + DO + (PO) "is, to some degree, sensitive to constraints expressed in terms of the discourse functions FOCUS and TOPIC." Although his theoretical framework is different from the one assumed here, some of the conclusions seem to be in line with what has emerged in our analysis of other SVO languages. This applies especially to the relationships identified between nuclear stress, type of constituent, and position in the structure. In German, as in other languages previously examined, for example English[179] and the Romance languages, in sentences with neutral order under specific semantic-contextual conditions, any lexical constituent may carry the nuclear stress.[180] On the other hand, in sentences with marked order, the constituent which is not in situ will necessarily attract the nuclear stress.[181] Again, we have already noticed this property in English and in the Romance languages, which are SVO languages with weak flexibility of constituent orders. This property is interesting from a typological point of view, because it confirms in another way the configurational character of German, from the point of view of the little positional freedom of constituents.

2.4.5.5.4.   The position of constituents in FOCUS.   On the basis of the diversity of WO patterns (without *Satzklammer*/with *Satzklammer*) and the dichotomy between unmarked FOCUS and marked FOCUS, several positions which are characteristic of the constituent in focus may be identified.

We shall examine first of all unmarked FOCUS. As in the SVO languages already examined, the position of FOCUS in sentences without *Satzklammer* and with default stress intonation[182] coincides with the constituent to the right of V, lower down in the VP. Consider again sentence (69a):

(69)   a.   Karl liest ein Buch.

The intonational profile with default stress corresponds to a pragmatic interpretation with broad focus (the sentence is a reply to the question *Was tut Karl?* 'What is Karl doing?'). The constituent in FOCUS is O (*ein Buch*).

In sentences with *Satzklammer* and with default stress intonation, the position of FOCUS is that which immediately precedes the lexical V.[183] As in (69a), also in (69b), repeated here

(69) b.   Karl hat ein Buch gelesen.

the constituent in FOCUS is *ein Buch*.[184] Note that the position to the left of V must be filled by an argument constituent.

Both (69a) and (69b) show application of the FOCUS-last principle, and the difference between them, as far as determining the FOCUS position is concerned, seems to consist in the different shaping of the VPs, that is, the domains in which the principle applies. This has to do with the obvious divergences in the right or left branching properties of the phrase. However, establishment of the FOCUS-last principle is more problematic than it appears at first sight. Consider sentence (69c), containing a temporal adverb:

(69) c.   Karl liest ein Buch / heute.

This sentence is more typical of the spoken than of the written language. Two FOCI can be identified in it: the main FOCUS *ein Buch* and the adverb *heute*, which may be considered an adjunct FOCUS, typical of the P_n position TAIL (cf. Section 2.4.5.3).[185] The FOCUS-last principle cannot be understood in purely linear terms; account must in fact be taken of the constraints imposed by the phrase structure: the FOCUS coincides with the last constituent of the VP domain; in (69c) *heute* being an adjunct, does not belong to the VP and therefore cannot be the FOCUS.[186]

Marked FOCUS positions may be included either in the *Vorfeld* or the *Mittelfeld* or even in the *Nachfeld*. This is a situation which has been variously discussed in studies on German in terms of *Hervorhebung* (highlighting). Note that the type with a constituent moved in the *Vorfeld* may be associated with the function of textual linking. This structure seems to parallel thematic structures with constituents moved in P_1, detected in other language types.[187] If the functional value of this phenomenon consists in a proper focalization, its structural characteristics are basically reducible to constituent movement from canonical to noncanonical positions.[188] As was previously observed, a characteristic prosodic correlate of such movement is the allocation of nuclear stress on the moved constituent.[189]

The occurrence of marked FOCUS in the *Vorfeld* is characterized as constituent movement to P_1:

(70) a.   DAS   habe   ich nicht gesagt.
            DEM.NT.SG have:1SG I   NEG   say:PST.PART
            'That isn't what I said.'

424   Rosanna Sornicola

In this sentence, *nicht* conveys a secondary stress. An alternative prosodic contour can have secondary stress on *das* and primary stress on *nicht*. In this case, *das* is a nonfocal element.

In German – as in English and the Romance languages – alongside this type, left dislocation structures are possible, where the left-dislocated element is a TOPIC. These are characterized by the occurrence of a phrase at the extreme left edge of the sentence (this position has often been distinguished from P_1), with its resumptive pronoun in the body of the sentence, as in example (70b), typical of the spoken language:

(70) b.   Den   Kerl, den   habe   ich zu   oft
            DEF.M.SG.ACC guy   DEM.M.SG.ACC have:1SG I   too often
            gesehen.
            see:PST.PART
            'That guy, I have seen him too often now.'[190]

In the *Mittelfeld* focalization through right movement of the constituent is found. Unlike (70c), which has neutral order (sentence stress on *Blumen*):

(70) c.   Peter hat   Sabine   die Blumen   gebracht.
            Peter have:PRS:3SG Sabine.(DAT) the flower:PL.ACC bring:PST.PART
            'Peter has brought the flowers to Sabine.'

sentence (70d) shows focalization of the constituent *Sabine*, moved from its canonical position:

(70) d.   Peter hat die Blumen SABINE gebracht.
            (cf. Engel 1991: 331)

Example (70d) raises an interesting problem with respect to the assumption that FOCUS in the *Mittelfeld* is to be identified as the constituent which immediately precedes V. In German, as Engel observes, the positions in the *Mittelfeld* are so limited that a constituent moved to the right is in a certain sense "destined" to land in the immediately preverbal position. This situation is difficult to compare with that examined in Basque and Turkish (cf. Section 2.4.5.5.5 for other arguments in favor of this idea).

Focalization also occurs in a different structural configuration from the previous one, that is when constituents with the function S occur in the *Mittelfeld* by means of so-called "inversion":

(70) e.   Ihrem   Bruder   hat   meine   FRAU   helfen
           her:SG.DAT   brother have:3SG my:F.SG.NOM wife:SG.NOM help:INF
           wollen.[191]
           want:PST.PART
           'My WIFE wanted to help her brother.'

This example offers further demonstration of the fact that it is not only position which determines the topology of FOCUS. Keeping the same WO pattern but changing the stress pattern results in a sentence with a different FOCUS:

(70) f.   Ihrem BRUDER hat meine Frau helfen wollen.
           'My wife wanted to help her BROTHER.'

Movement in the *Nachfeld* also implies a process of focalization, which usually produces a stronger highlighting than the one determined by rightward movement in the *Mittelfeld*.[192] Apart from the cases discussed in Section 2.4.5.3, in which the moved constituent is typically a PrepP, constituents of other categorial types may be found only when they are heavy constituents or enumerations:

(70) g.   Er hatte   in Göttingen getroffen   / Friederike,
           he have:PST.3SG in Göttingen find:PST.PART   Friederike
           ihren   Mann   und ihre   beiden   Töchter.
           her:M.SG.ACC husband   and their:F.PL both:PL.ACC   daughter:PL
           'He had met Friederike, her husband and their two daughters in Göttingen.'

By way of a summary, the following schemas may therefore be drawn up for neutral orders (Figure 13) and marked orders (Figure 14).

1   S   V   O
                   F

2   V$_{FIN}$ ...   X$_i$ V$_{PART}$
                   F

Figure 13. Neutral orders

426   Rosanna Sornicola

3   P$_1$ ...   X   (*Vorfeldstellung*)

   F

4   V$_{FIN}$ ...   X$_i$ (...) V$_{PART}$ (*Rechtsverschiebung im Mittelfeld*)

       F

5   P$_1$ ...   P$_n$ (*Nachstellung*)

     F

Figure 14. Marked orders

It should be noted that schemas 3, 4, and 5 primarily deal with marked orders, and not marked FOCUS position. For a complete picture of marked FOCUS positions, it must be added that marked FOCUS is mobile, not only on the basis of what is seen in the schemas 3–5, but also because of the degree of stress mobility. In fact, as has been said, any constituent may be focalized by means of stress.

Schemas 1, 2, and 5 conform to the extremely general "FOCUS-last" principle, which in SVO languages normally determines the unmarked distribution of FOCUS (cf. Section 2.4.1). It is easy to see that 1, 2, and 5 reflect this principle in different ways: 1 and 2 concern unmarked orders, while 5 concerns a marked order. Furthermore, 5 is a special case in that its structure is determined by the occurrence in the *Nachfeld* of constituents not in situ, of a particular categorial nature, and also by the fact that an interaction between the FOCUS-last principle and that of "heavy constituents after" may be noted. It seems, therefore, that 5 is in part a phenomenon specific to German.

Schema 3 conforms to the "FOCUS-first" principle, widely found amongst the world's languages. As has been observed in Section 2.4.1, in SVO languages this principle normally determines a marked distribution of FOCUS. Note that German respects the FOCUS-first principle for marked FOCUS.

In a typological key the most problematic situation is that of 4. It could be embraced by the FOCUS-first principle and in this case German would once again show a different tendency from that of the Romance languages and English which have unmarked FOCUS in P$_n$ and marked FOCUS in P$_1$. But 4 could fall into the situation found in the so-called "discourse-configurational" languages, such as Turkish and Hungarian (cf. Sections 1.7.2, 2.2.1, and 2.3.2). Such an attempt has been made and will be discussed in the following section.

2.4.5.5.5.   Some typological problems.   Leftward movement into the *Mittelfeld* (cf. Section 2.4.5.5.4) has in recent years been compared with the phenomenon of FO-

CUS linked to the immediately preverbal position in languages such as Turkish and Hungarian.[193] In truth, this typological comparison does not seem able to withstand a series of more precise analyses. The apparent resemblance is illusory: it is due by chance to the coincidence of a mere topological parameter, which shows once again that the topological properties alone cannot be characterizing. The German structures and those of Turkish or Hungarian are not only the result of different factors, they are themselves different for several reasons:

(i) Unlike Turkish and Hungarian, there exist in German, as has been seen, more than one characteristic position of marked FOCUS in each of the fields of the sentence. On the other hand, the generalization that in the configuration X_i V, the X_i position is characteristic of marked FOCUS, cannot be extended from the *Mittelfeld* to the *Vorfeld*, in that it is trifling for this configuration.

(ii) Even if the comparison were restricted to the properties of *Mittelfeld* the following differences would be obvious:

In the German *Mittelfeld* it is not possible to topologically identify any position X_i (in the configuration X_i V) which is a characteristic of marked FOCUS. While in Turkish and in Hungarian, as has been seen, all possible categorial constituents may occur in the position X_i of the configuration X_i V, in German S, for example, may occur there only in particular circumstances (through inversion, if P_1 is occupied by a constituent other than S). In other words, the possibility of S occurring in X_i of X_i V has nothing to do with the position itself. Constituents with the GF of O or IO *may,* but they do not necessarily *have* to occur in X_i of X_i V, if they bear marked FOCUS, as is evident from the following examples:

(71) a.   Ich habe   den   Brief meinem   BRUDER
            I   have:1SG DEF:M.SG.ACC letter my:M.SG.DAT brother
            geschickt.
            send:PST.PART
            'I have sent the letter to my BROTHER.'
      b.   Ich habe meinem BRUDER den Brief geschickt.
            'I have sent the letter to my BROTHER.'
      c.   Ich habe   an den   Nagel die   JACKE
            I   have:1SG on DEF:M.SG.ACC hook DEF:F.SG.ACC jacket
            gehängt.
            hang:PST.PART
            'I hung the JACKET on the hook.'
      d.   Ich habe die JACKE an den Nagel gehängt.
            'I hung the JACKET on the hook.'
            (Abraham 1986: 18)[194]

428   Rosanna Sornicola

These examples show, as Abraham (1986) had pointed out, that a constituent often occupies the immediately preverbal position and attracts stress when it is not in situ.[195]

Abraham's generalization concerning a presumed d-configurationality of the *Mittelfeld* seems to be invalidated by the type of structure he examines. The absolute position may only prove to be an epiphenomenon, while the crucial factor seems the constituent not being in situ.

(iii) In the so-called "d-configurational" languages, position X_i of X_i V for marked FOCUS is related to basic SOV order, as was seen in Section 2.2.1. This is not the situation in German, in which the order O V_{nonfin} in main clauses is structurally limited to *Satzklammer*. Furthermore, the German structure seems to have been diachronically originated in written styles (cf. Section 2.4.5.5.6). Synchronic comparison may conceal this fact, but it is nonetheless evident that Hungarian and Turkish structures spring along different routes than the German one. That this is the case is also seen from a striking difference between German on the one hand and Turkish and Hungarian on the other. In the latter languages, the V assumed to be the topological boundary on the basis of which the position of the constituent in FOCUS is determined, is a finite verb, while in German, X_i is defined on the basis of left adjacency to a nonfinite lexical V.

2.4.5.5.6. The sociolinguistic and historical problem.   The typological characteristics of German WO, with strong differences in the relationship between position and PF, reflect the present-day stage of development of a complex process of grammaticalization. Without bearing in mind the historic dynamics of the formation of German WO, it would be difficult to understand the synchronic peculiarities examined so far.

Although description of these dynamics is anything but agreed on,[196] some lines of development do allow for more consistency. The fact that the structure with *Satzklammer* was normalized late and was characteristic of educated, written registers rather than informal and/or semi-educated spoken ones has particular importance for the typology of WO in main clauses.[197] This divergence between the written and spoken language in fact persists today, as shown by the fact that in spontaneous speech as well as in some dialects *Satzklammer* has a lower frequency or is confined to more restricted structural contexts.

The normative and cultivated character of the order with *Satzklammer* must be borne in mind not only for purely descriptive purposes with respect to present-day synchronic properties of WO in German, but also for purposes of typological comparison with other languages. It would, in fact, be misleading to deal with the idiosyncratic properties of German as if they were "natural" characteristics of the language. To limit ourselves to a single consideration, the coincidence of certain prag-

matic properties of WO found in German and languages such as Turkish or Hungarian cannot be considered a real typological resemblance, not just for the reasons discussed in Section 2.4.5.5.4, but also because, in Turkish and perhaps in Hungarian, in main clauses SOV order is inherited, while in German it is artificial, in other words it is the result of language planning from above. Although similar considerations are almost always overlooked in a typological framework, it seems they cannot be ignored in a reliable attempt at crosslinguistic comparison.

2.4.6.   V-initial languages

2.4.6.1.   Irish and Welsh

The V-initial type, which among the present-day languages of Europe is found in Celtic languages such as Irish and Welsh,[198] shows several interesting problems with respect to general WO theory as well as its specific pragmatic implications. A basic VSO order is postulated for both languages,[199] which gives rise to a discontinuous verbal constituent: it is a problem which, as is well known, has created difficulties for traditional constituency models.[200] There is in fact an adjacency relationship between V and S which cannot be broken;[201] it breaks up the VP constituent, as can be seen in the Irish example (72):

(72)   Labhrann   Mícheál Gaeilge le Cáit go   minic.
        speak:PRS.3SG Mícheál Irish   to Cáit PTL often
        'Mícheál often speaks Irish to Cáit.'[202]

A second problem lies in the fact that VSO order deviates from the more general tendency in the world's languages to a basic order with S = TOPIC in P_1.[203] This involves a particular distribution of the PFs in the sentence, which it is worthwhile going into in more detail for a wider typological investigation. In reality, the possibility of alternative orders with S or O in initial position is widely attested throughout the diachronic and synchronic development of the two languages, even though, as will shortly be seen, it poses descriptive and explanatory problems which as yet do not seem to have been clarified.

Two questions are particularly important in this respect:
1. The relationship between position, GF and PF;
2. The typological repercussions of this relationship, with particular reference to the distribution of PFs in the sentence.

2.4.6.2.   Some problems concerning the relationship between position, grammatical function, and pragmatic function

As far as question 1 is concerned, an initial problem lies in the difficulty of satisfactorily determining the relationship between position and GF in the typical cleft

430   Rosanna Sornicola

structures of these languages.[204] In fact, at first sight, it would seem that Irish cleft sentences such as:

(73)   Is   é   Pádraig a   bhuail   an   liathróid.
        be:PRS.3SG PRN.3SG.M Patrick REL.PTL hit:PST.3SG DEF.SG ball
        'It's Patrick who hit the ball.'[205]

could be analyzed as matrix clause + dependent clause, the former having a VS pattern, the latter S (REL.PTL) VO pattern. This analysis, however, comes up against various difficulties: first, the semantic weakness of 'to be' and the particular nature of the identifying relationship obtaining between 'to be' and the NP to its right. These two factors would make it plausible to assume that the verb 'to be' does not occupy an actual position in the representation, but is conversely nothing more that a "pseudo-constituent," or a simple mark of focalization. A second difficulty is that the postcopular NP does not necessarily fill the function of S, but could serve a different function, for example O, in other similar types of construction.[206] A further difficulty then arises from the fact that (73) is close to a structural type, such as (74), which does not have the typical copula + pronoun initial configuration of (73):

(74)   Pádraig a bhuail an liathróid.
        'Patrick (that) hit the ball.'

However, not even a structure such as (74) may strictly speaking be assigned an SVO configuration.[207] V is in fact embedded in a relative structure, which might suggest that the NP constituent in P_1 does not belong to the same clause.[208] And yet the nature of the relationship between Pádraig and the verb looks more like that of a subject and a predicate than that of a head NP and a modifying relative clause.

Besides, it is not clear what the relationship between the type in (73) and that in (74) is. From a structural point of view, (74) could be considered derived from (73) by deletion of the copula (+ pronoun).[209] However, while the type with the copula has had focal value on the postcopular constituent since the earliest documentation, in type (74) the NP in P_1 may be either TOPIC or marked FOCUS. In other words, in both Irish and Welsh the type without the copula may occur without any contrastive value as an all-in-FOCUS or presentative sentence (cf. Section 2.4.6.5).[210] From a semantic point of view, (74) could, if associated with a noncontrastive interpretation, be the result of a process of demarking (73). Frequently, however, this pattern has a contrastive value in both modern Irish and Welsh (cf. MacCana 1973: 110).

On the other hand, sentence (73) cannot be represented as a uniform VSO configuration, since the verb 'to be' which occurs in P_1 is not unambiguously identifiable with V, as there also exists another candidate for this function, which is the verb of the relative clause. Moreover, the second and third positions are each occupied by a constituent which can be a candidate for the S function. In other words, the functions of predicate and S are both split between more than one constituent.[211] Another

difficulty lies in the widespread use of auxiliary verbs in the so-called "periphrastic constructions." In a Welsh sentence like (75):

(75)   (Y) mae   ef   yn   canu.
          PTL be:PRS.3SG he PTL singing
          'He sings / he is singing.'

the verb 'be' (*(y) mae*) and the particle *yn* followed by the verbal noun *canu* form a verbal constituent (the finite verb *mae* carries grammatical information and the nonfinite verb *canu* lexical information) which is split by the S NP *ef*. This poses a problem not only for syntactic analysis (is (75) really to be considered a V-initial structure? If yes, it is obviously different from type (72)), but also for pragmatic analysis. In fact, with constructions like (75) one cannot simply assume that information subsumed under "V" invariably precedes "S".

In short, the impression given is that the presence of elements such as the verb 'to be' in cleft structures or the verb 'do' in structures with verbal nouns (consider the Irish type *B'fhéidir gurb é a chaitheamh amach a rinne sibh*, lit. 'Maybe it is throwing it out you did', analogous to the Welsh type *Gofyn [a] wnes i* '[It was] ask that I did'[212]) do not provide strong grounds for analyzing the syntax of these languages according to standard configurational patterns.

2.4.6.3.   Concurrent SVO types

Whatever the solution to these problems, if for the moment we leave aside the difficulties raised just now and take *Pádraig* in (74) to occupy P_1 position,[213] it may be said that in both Irish and Welsh there exists a whole range of concurrent S/O + relative V types. It is well known that the existence of SVO variants in VSO languages was already foreseen by Greenberg (1963). However, description and explanations of initial S or O types in the languages under consideration are not unproblematic. Perhaps the most obvious hypothesis is that they are the result of processes of topicalization or fronting of a nonverbal constituent, that is, that such structures are the result of movement rules. From this point of view, in both noncontinental Celtic languages and SVO languages the TOPIC function would be associated with P_1, and the difference between them would lie only in the fact that in the former the TOPIC would not be an in situ constituent, while in the latter it would be.

In fact, it is possible to describe these phenomena in terms of topicalization or fronting only if VSO is considered to be the basic order, an assumption which is not altogether uncontroversial. Apart from the problems raised in this respect by diachronic documentation,[214] the picture is further complicated by the presence today, as in previous linguistic phases, of types with initial S/O and nonrelative V in second or in sentence-final position. Although these are marked variants (cf. MacCana 1991: 60), one may well ask how often they occur and what their actual function is in the grammar of insular Celtic languages (for these types, see Section 2.4.6.4 below).

432   Rosanna Sornicola

Now, both the diachronic and synchronic data offer indications that the opposite may be the case. As MacCana (1991: 47) observes, the system of constituent order in insular Celtic is far more complex than was previously thought. The fact remains that the VSO character of insular Celtic, which distinguishes this branch of the Indo-European family from others,[215] has been in continual competition with varyingly structured S- or O-initial types. Let us consider, for example, the so-called "abnormal sentence" (AS) and "mixed sentence" (MS) in Welsh.[216] The former is a structure typical of middle Welsh, frequently found in literature until the beginning of the twentieth century and no longer attested today. In this structure, V is preceded by some phrasal constituent (S, O, or an adverbial) and a preverbal relativizing particle. It is thus embedded in what can be considered a relative clause (or a "quasi"-relative clause [cf. Williams 1980: 168]). In the following example, the NP preceding V has the GF S:[217]

(76) a.   A   i   ddisgyblion a   ddaethant   ato.
           and his disciple:PL REL.PTL come:PST.3PL to.him
           'And his disciples came to him.' (from Williams 1980: 168)

This type has some similarities with the MS, which is represented here by sentences such as (76c) and (81) below. Unlike the AS, however, the MS has been uninterruptedly documented at all periods and in all registers and still survives today. Furthermore, a number of syntactic properties differentiate the two types, for example, the agreement patterns: in the AS, V agrees in person and number with the preverbal S, while in the MS, V is always third singular.[218] AS and MS also have interesting pragmatic differences, in that the first has no FOCUS value on the preverbal constituent, while this is frequently the case with the latter.[219]

2.4.6.4.   Properties of position P_1

From what has been said in Section 2.4.6.3, it follows that in both Irish and Welsh, P_1 has a property worth of mention: it is a position that may well be defined as "polyfunctional," in that it may be occupied not only by V constituents in unmarked declarative sentences but also by:
(a) NPs which are S or O with marked FOCUS function;
(b) NPs which are S or O with TOPIC function, in unmarked declarative sentences. In Welsh, PrepPs and extrasentential AdvPs (adjuncts), which at the pragmatic level are typical backgrounded units, may occasionally be found in this position as well[220] and would normally occur in the "post-field" (i.e., position X in the configuration V S O X).[221]
The functional differences between (a) and (b) are not always clear from the literature, which often merges the two types into one; this seems partly due to the unsatisfactory theoretical tools adopted for the description of pragmatic phenomena. On

the other hand, the relationship between V-initial structures and structures with initial S/O with TOPIC function and relative V is not unanimously agreed upon: while some consider there to be free variation (cf. Watkins 1991: 342), others deny that the two types are synonymous (cf. MacCana 1991: 47).

The following Irish (76b) and Welsh (76c) examples have P_1 occupied by a constituent with contrastive FOCUS value ((76c) is a MS):[222]

(76) b. Seán a   bhí   ann.
          Seán REL.PTL be:PST.3SG there
          'It was Seán who was there.' (cf. MacCana 1973: 110 on this type)
      c. Y   ferch a   rhedodd   i ganol yr heol.
          DEF girl REL.PTL run:PST.3SG to centre DEF road
          '(It is) the girl who ran to the centre of the road.'

In fact, even the debate over the pragmatic properties of MS seems to have suffered from terminological and conceptual confusion regarding the notions of TOPIC and FOCUS (see, for example, the criticism by Shisha-Halevy 1995: 150–155).

As far as the (b) condition for position P_1 is concerned, both Irish and Welsh have a syntactic type with an NP (with GF = S/O) in initial position resumed by anaphoric pronoun which in some respects resembles a left dislocation (cf. especially (78)). A notable property of this kind of structure is the absence of the relative particle which in non-V-initial sentences introduces the configuration containing the main verb. In other words, it is not a type of cleft structure (cf. Watkins 1991: 334, according to whom the type in question is the result of fronting S/O in the V-initial type).

The Irish type is not entirely analogous to the Welsh type: in fact it occurs in Irish only when the NP in P_1 contains a relative clause; the structure could therefore be described as heavy-NP shift to the left.[223]

(77) Na   buachaillí agus na   cailíní a   rabh
      DEF.PL boy:PL   and DEF:PL girl:PL REL.PTL be:PST.REL.3SG
      an   lá saoire   acu,   bhí   siadsan
      DEF.SG day free:GEN with.them be:PST.NRE.3SG ps.prn.3PL:EMPH
      tuirseach ag siúl   na   sráideannaí.
      tired   for walking DEF.PL street:PL
      'The boys and girls who had the day free were tired from walking the streets.'[224]

(78) Fy stumog i   mae   hi   fel crempog.
      my stomach AUX.prn.1SG be:PRS.3SG ps.prn.3SG.F like pancake
      'My stomach, it's like a pancake.'[225]

In (77) and (78) the NPs in P_1 could be considered extrasentential constituents. Watkins (1991: 334) observes that the structure is extremely rare in Welsh, to the

434   Rosanna Sornicola

extent that a native speaker considers it "unnatural, rhetorical, or ostentatious." He finds no attestation of it in his corpus of contemporary literary Welsh.

2.4.6.5.   More on position P_1: A controversial type with NP in P_1

Another NP-initial type found in Irish and Welsh is associated with particular pragmatic and stylistic values. This type is continuously attested in the diachronic development of the two languages. It differs structurally from the types (77) and (78), in that it has the shape of a cleft structure with a relative particle which demarcates the clause with the main verb.[226] On the other hand, it shows formal similarities with MS (cf. (76c)).

The constituent has an explanatory function (as a reply to WH-questions of the type 'What happened?' or as the textual development of a progression having the implicit presupposition 'Why?'). Here are some Irish examples, taken from contemporary novels:

(79)   ['Faoi Dhia, goidé tháinig ort?' ars an t-athair.]
         'In God's name, what happened to you?' asked the father.'
         Mícheál Rua a   bhuail   mé   ars   an   mac.
         Mícheál Rua REL.PTL hit:PST.3SG ps.prn.1SG say:3SG DEF.SG son.SG
         'Mícheál Rua gave me a beating,' said the son.'[227]

(80)   [An lá sin thar éis a theacht abhaile ó chluiche na leaschraoibhe, nar buaileadh tinn mé.]
         'I fell ill that day after coming home from the semi-final game.'
         Slaghdán a   tholg   mé   as   an
         cold   REL.PTL attack:PST.3SG ps.prn.1SG because.of DEF.SG
         allas   agus as   an   gcodladh amuigh.
         sweat and because.of DEF.SG sleeping outside
         'I had caught a cold from sweating and sleeping out.'[228]

The Irish examples (79)–(80) are directly comparable to a well-known Welsh type due to Lewis (1942), who described it in detail:

(81)   y   ffermwyr [a]   adawodd   y   glwyd ar agor.
         DEF farmer   REL.PTL leave:PST.3SG DEF gate   on open
         'The farmer left the gate open.'[229]

Structurally, (81) may be considered a MS rather than an AS.[230] From a pragmatic point of view, the sentence may occur in a context in which someone, walking down a country lane, comes across some stray cattle and asks the reason (this was the context that Lewis imagined). This type, which has a wider occurrence, would, according to Watkins, alternate with the V-initial type. However, a crucial point that

is neglected in the literature seems to be the fact that variations in the suprasegmental parameters may give rise to different pragmatic interpretations. Lewis (1942: 20) had already noted that sentence (81) may have various intonations. Reformulating Lewis's observations, the sentence may be said to have either a neutral prosodic contour or one with prosodic FOCUS on *y ffermwyr*. From a pragmatic point of view, in the former case the sentence would be a reply to 'Why have the cattle escaped?' (an interpretation with the sentence all-in-FOCUS), while in the latter case it would be a corrective sentence with respect to a context with the presupposition 'someone else left the gate open'. Alternatively, it could be used in reply to a question 'Who left the gate open?' (both these last two cases have an interpretation where *y ffermwyr* has FOCUS value, although different kinds of FOCUS are involved).

The pragmatic complexity of this type is further demonstrated by MacCana's observations. Even if the more typical context is that of a reply, MacCana (1991: 65) claims that the most important feature of this type is that of providing "foretokened information or explanation." In fact, the type may occur in textual environments where information is provided unexpectedly, as in the following example taken from a short story:

(82)   Rhyw newydd heb   fod   yn dda sy   gen
        some news   without being PTL good be:PRS.REL.3SG with:1SG
        i,   Mrs Pfennig.
        ps.prn.1SG   Mrs Pfennig
        'I have some bad news for you Mrs Pfennig.'

According to MacCana (1991: 65), the speaker here "is merely providing an explanation for his unexpected coming, or in other words responding to the unspoken question posed by it." MacCana quite rightly concludes that the type under consideration "relates or reacts to a question, statement, directive, or simply a situation that has been expressed, described or implied in the preceding text, and does so most often by supplying relevant information or affirmation" (MacCana 1991: 70). Although comparisons of pragmatic values of structures from different languages are always risky, it can be observed that the values mentioned above have a striking resemblance to those of the SVO order in other VSO languages. In modern Arabic, for example, SVO order prevails in sentences "which describe or provide incidental background or explanatory information about the already defined agents and patients of the text," while VSO order prevails in "event-oriented" narrative progressions (Holes 1995: 205).

A second context in which the type in question appears is narratives, especially the beginning of narratives. The type is found in both ancient and modern folk tales. It is possible that as far back as the Middle Ages it was a technical device of the professional narrator to attract the public's attention. The structures in question serve mainly to introduce a character or a situation.[231] MacCana (1991: 70–71) notes that

436   Rosanna Sornicola

this device is found more frequently and throughout a longer period in Irish literature than he thought in several previous works. It is a device not only for opening stories, but also for introducing a stage, an episode, or a subplot within the more complex structure of the story:

(83)   Ardrí   ro   gab   for   Hérind.
        high.king.SG.NOM PFV take:PST.SG over Ireland.ACC
        'A high-king ruled over Ireland.' (MacCana 1973: 107)

Possibly linked to this function are different but in some way related functions, which Watkins (1991: 346–347) found in Welsh, such as presenting additional information regarding the TOPIC (84), sentence linking (85), proclaiming/announcing (86), (87), and gnomic function (88):

(84)   [Un arall . . . oedd . . . Kate Owen.]
        'Another . . . was . . . Kate Owen.'

        Cadw   siop . . . a   wnâi   ei   theulu.
        keeping shop   REL.PTL do:IMPF.3SG 3SG.POSS family
        '(It was) keep shop . . . that her family did.'

(85)   [A dyna hi'n cyrraedd.]
        'And there she comes.'

        Llygod sy'   n   tynnu ei   cherbyd.
        mice   be:PRS.REL.3SG PTL drawing 3SG.POSS carriage[232]
        '(It's) mice that are drawing her carriage.'

(86)   Madam Sera [a]   fydd   yn ein   difyrru   nos
        madame Sera REL.PTL be:FUT.3SG in 1PL.POSS entertaining night
        Sul.
        sun
        '(It's) Madame Sera that will be entertaining us Sunday night.'

(87)   Y   Parch.   John Evans a   fydd   yn y   pulpud
        DEF respectable John Evans REL.PTL be:FUT.3SG in DEF pulpit
        dydd Sul.
        day sun
        '(It's) the Rev. John Evans who will be in the pulpit Sunday.'

(88)   Pawb   a   chwennych   anrhydedd.
        everyone REL.PTL desire:PRS.3SG honour
        '(It's) everyone that wants honour.'

It is not uncommon for the narrative function to become fused with the explanatory function, as in the interesting example quoted by Watkins and taken up by MacCana (1991: 71):

(89)   [A priest asks one of his parishioners why he has been hay-making on Sunday instead of coming to Mass. The parishioner's reply is:]

Fel hyn   yr   oedd   hi   y
like DEM.NT.SG PreV.PTL be:IMPF.3SG ps.prn.3SG.F PreV.PTL
fi   a' r   wraig oedd   yn taeru   pa
ps.prn.1SG and DEF wife be:IMPF.3SG in insisting INTERR.prn.NT
ddiwrnod o' r   wythnos oedd   hi.
day   of DEF week.SG be:IMPF.3SG ps.prn.3SG.F
'It was like this ... Myself and the wife were arguing what day of the week it was.'

The narrative context is interesting because it once again shows the relationship between fronting and FOCUS for making something a new center of attention: in fact, the type occurs at a change in subject-matter in narrations.[233] Of some interest from a point of view of typological development, at any rate, is the fact that these types with initial NP are undergoing considerable expansion in modern spoken Welsh (MacCana 1991: 72).

It is possible that the pragmatic analysis of the syntactic type discussed in this section requires a description at both the textual and the sentential level. As far as the former is concerned, the structures quoted above function as all-in-FOCUS or presentative sentences (cf. Section 4.2), while at the latter level the NP in P_1 can variously have the PF of a TOPIC (very often carrying the feature [+NEW]), or of a FOCUS.

2.4.6.6.   A few concluding remarks

At this point, it may be concluded that in Irish and Welsh position P_1 may be variously associated with TOPIC or FOCUS, in its various manifestations (of contrast, information progression, call to attention). In fact, such a position is normally exploited as a means of articulating the informative (pragmatic) structure of the sentence.[234] On the other hand, the very multifunctionality of P_1 in these languages shows the complex nature of the relationship between topicalization and focalization. The phenomena analyzed show properties which lend weight to theoretical questions concerning the relationship between TOPIC and FOCUS and their definition. Rather than a clear-cut distinction one should, in fact, think of a gradient.[235] On the structural level, what appears to be a crucial problem is why, in the languages examined, position P_1 requires clefting (relativization) of the sentence in order to be accessible to constituents which would normally be elsewhere in the sentence. To a lesser extent this property is found in other languages with initial S in the basic configuration, especially in French: cf. Section 2.4.4.1.2.

438   Rosanna Sornicola

In any case, these languages pose an interesting problem for typology. According to Tomlin (1986) SVO is a more appropriate order than VSO for conveying information. The V-initial types, in fact, are very clearly in the minority: if one adds together the frequencies of the VSO and VOS types in a representative sample of the world's languages, these amount to 12.19 %, compared with 86.57 % for S-initial types (Tomlin 1986: 21–22). However, the hypothesis that S-initial orders are better adapted to the informative structure needs reconsideration, if only for the existence of empirical argumentation: as in the Celtic languages, in other VSO languages such as the Semitic languages the sentence may be organized according to the needs of the informative structure, with position P_1 utilized for NPs associated with the basic pragmatic functions of TOPIC and FOCUS.[236] This supports what has already been pointed out by Greenberg (1963), that in VSO languages, SVO order is an obvious alternative.

It is not clear whether the situation described in this chapter confirms the hypothesis that the positions which may be taken up by a constituent with marked FOCUS function are those which are not in situ.

With respect to SVO and SOV languages, the following picture is found in non-continental Celtic languages:

(a) Syntactic structure and patterns of order may not be critical to the assignment of PFs. This general property seems to be responsible for the difficulties in distinguishing between marked and unmarked sentences which have been discussed above.

In unmarked sentences are found:

(b) Non-biuniqueness of the relationship between position and TOPIC
(c) Non-biuniqueness of the relationship between TOPIC and GF (T = S/O)
(d) Biuniqueness of the relationship between position and FOCUS
(e) Non-biuniqueness of the relationship between FOCUS and GF (F = S/O)

These properties are represented in Figures 15–17:

P_1 P_2 … P_n

V   S   O

         T   F

Figure 15. Alignment of Positions, GFs and PFs in VSO orders (unmarked) of VSO languages

Interaction of syntactic and pragmatic factors   439
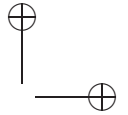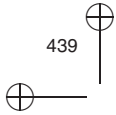
P_1 P_2 . . . P_n

S   V   O

T   F

Figure 16. Alignment of Positions, GFs and PFs in SVO orders (unmarked) of VSO languages

P_1 P_2 . . . P_n

O   V   S

T   F

Figure 17. Alignment of Positions, GFs and PFs in OVS orders (unmarked) of VSO languages

The representations in Figures 16 and 17 imply that both SVO and OVS orders may be unmarked. In fact, a further property may be noted:

(f) No influence of grammatical relations on determination of the TOPIC (TOPIC defined as any nonverbal constituent).

In marked sentences are found:

(g) Non-biuniqueness of the relationship between position and FOCUS.

(h) No influence of GFs on determination of marked FOCUS (any nonverbal constituent may occur in P_1).
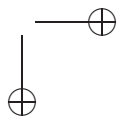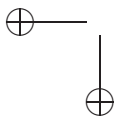
These properties are represented in Figures 18–20:

P_1 P_2 . . . P_n

S   V   O

F

Figure 18. Alignment of Positions, GFs and PFs in SVO orders (marked) of VSO languages

P_1 P_2 . . . P_n

O   V   S

F

Figure 19. Alignment of Positions, GFs and PFs in OVS orders (marked) of VSO languages

440   Rosanna Sornicola

$$P_1 \quad P_2 \quad \ldots \quad P_n$$
$$V \quad O \qquad\ S$$
$$F$$

Figure 20. Alignment of Positions, GFs and PFs in VOS orders (marked) of VSO languages

Figures 16–17 on the one hand and 18–19 on the other show that the same linear configurations can be either unmarked or marked. It seems reasonable to think that in this case it is context and suprasegmental features that determine the actual value of the structure.

In short, Irish and Welsh raise interesting questions for the study of the structural and pragmatic factors of WO and their interrelationship. The assumption that VSO is the basic order, from which S/O initial orders are derived through topicalization or focalization, has the advantage of offering a simple and economic description of the data,[237] but cannot be considered a dynamic model, nor a model which represents the complexity of diachronic development and synchronic variation. On the other hand, functionalist works based on textual and cognitive properties associated with WO patterns may take into account the scarcity of VSO languages, but do not justify their presence. In fact, it is the very presence, albeit minor, of the VSO pattern which requires explanation for an understanding of the type. From this point of view, the really crucial question seems to be: what are the paths which lead to the formation of VSO languages? This question seems all the more justified in that VSO order is not only infrequent amongst the world's languages, but also diachronically unstable and pragmatically variable, as Celtic, Polynesian, and Semitic data show.

In this regard, a more "microscopic" comparison with other languages commonly held to have a basic VSO order is revealing. Some hypotheses of a general nature may be advanced, for which the available empirical data provide support.

One of the main factors in the formation of VSO orders could be a strategy by which cataphorization of NP = S is effected by means of a pronoun in structures such as:

(90)   Pronoun + V + NP

In these structures, the pronominal element with the function S immediately precedes V and develops the properties of a proclitic element; furthermore, it is coreferentially linked to an NP which (immediately) follows V. Evidence of the fact that a cataphoric structure like (90) is involved as one of the routes to the VSO type comes from various VSO languages.

Welsh documentation shows fossilization of personal pronouns in preverbal particles. Consider the pronoun *ef* 'he', generalized to situations where no third person

singular S is found: *ef gwneif* 'I shall make', *ew kuynhiw* 'I shall lament'.[238] A similar development is also found in Welsh for pronouns at other levels of the person hierarchy, which still showed agreement with the verb in Middle Welsh.[239] In Modern Welsh, the forms *e*, *fe*, *fo* frequently occur before the verb in literary prose; in the spoken language, the particles *mi* or *fe*, which cause soft mutation of the V constituent, precede all the inflected verbal forms.[240] MacCana rightly links these types to the AS and considers them to be of fundamental importance for the reconstruction of the history of WO in Welsh. The phenomenon of personal pronoun subjects in preverbal position also has parallels in Breton (cf. Hemon 1984: 70–71; for a Breton-Welsh comparison, cf. MacCana 1973: 119).

In Semitic languages, traces of an old original cataphorization strategy of S may be seen in the prefix conjugation of the verb,[241] which occurs in all the languages of the family and coexists with one with suffixes (on the problems posed by the morphological analysis of the verb with prefix conjugation, see Goldenberg [1998: 16–17, 301–302, 520 ff.]). It is, in fact, very common amongst the world's languages that the conjugation with prefixes may have an etymology going back to pronominal stems.[242]

The picture offered by the Polynesian languages is more problematic for the hypothesis put forward here. In fact, for some of them, the following synchronic rule may be postulated: when S is deictic and first or second person, it is obligatorily realized by a pronoun proclitic to the verb, as in the Tongan example:

(91) a.   Na'a ma 'ave   'e   ho'o telefone.
            PST   ps.prn.1PL=take ERG your telephone
            'We took away your telephone.'

When S is focalized, the clitic pronoun becomes the antecedent in an anaphoric chain with an independent pronoun occurring in postverbal position:

(91) b.   Na'a ma 'ave   'e   kimaua   ho'o telefone.
            PST   ps.prn.1PL=take ERG ps.prn.1PL your telephone

With third person singular pronouns the presence of the element proclitic to the verb is optional rather than obligatory.[243] It is not clear, however, whether structures like (91a) and (91b) supply a real parallel to the Celtic types discussed so far.

From a pragmatic point of view, cataphoric structures like the ones dealt with here can be considered the grammaticalization of afterthought processes. Grammaticalization paths may follow different trajectories, as is clear from a comparison of the Celtic and the Semitic languages. In the latter, for example, the original cataphoric strategy of the prefixed conjugation may have been reanalyzed in terms of aspectual[244] properties of the verb.

To sum up, one of the routes to VSO types may be a reanalysis of S_1 V … S_2 structures (with S_1 = clitic pronoun, S_2 = full NP) into structures in which S_1 has

been reanalyzed as an expletive (or – as in the Semitic languages – as a verbal prefix). This may be considered an "epiphenomenon" of SVO types. On the other hand, in the emergence of the VSO type in Celtic languages an important role has surely been played by the "narrative" strategy of putting V in P_1, which is well documented for ancient Indo-European languages (cf. Gonda 1952; Dressler 1969). The diachronic concurrence of different types of textual processes can perhaps shed some light on the hybrid and puzzling situation of present-day Celtic languages that has been described in Section 2.4.6. Possibly, it can also help to understand the controversial relationship between the instability of the VSO type and the stability of the pragmatic principle (for this polarity, cf. Currie 2000; Poppe 2000).

3.   Word order in sentences with one-argument verbs

3.1.   Monoargumentality and word order patterns

3.1.1.   Constructions with one-argument verbs

In sentences in which V takes one argument "at some level of representation" (see further on), all the languages in the corpus exhibit WO properties which require separate examination. In many languages there exists a real asymmetry between the dominant patterns in structures with two arguments (S, O) and patterns characteristic of structures with one argument. The structural property involved in this asymmetry concerns the relative order of S and V: in sentences with a one-argument verb, both SV and VS are allowed as "labile" variants. As has been said by Daneš, who gave the definition of "labile orders": "In this case, the order of some elements of the pattern on the grammatical level is irrelevant; in utterances based on such a pattern, the position of the respective words vacillates according to non-grammatical conditions" (Daneš 1967: 218).[245] In other words, the often conspicuous oscillations which are found are merely due to semantic and pragmatic factors. This formulation is far from uncontroversial because, as will be seen in Section 3.1.2, there is reasonable evidence in favor of the idea that particular classes of one-argument verbs select one or other order as the dominant one.

The asymmetry between two- and one-argument patterns is typologically conditioned. SVO languages in fact show greater asymmetry than VSO and SOV languages, and this could lead to the assumption that there are structural reasons which contribute significantly to determining the division.

First of all, the criterion for definition of V as "one-argument predicate, at some level of representation" must be explained. It may, in fact, be argued that such a definition is too general for two reasons: (i) it refers to a very abstract property, which does not take into account particular lexical properties of the verbs (in other words, it is under-determined with respect to them); (ii) it is indeterminate with respect to

the level of representation. In fact, due to the very multiplicity of verbal classes and derived constructions[246] to which the verb may belong, the definition proposed does seem to constitute an appropriate general characterization in a typological framework. Individual languages differ among themselves in terms of the subcategorial classes and the semantic and lexical properties of the verbs involved. They vary, furthermore, with respect to the type of construction in which the verb may be found. For example, in most of the languages in the corpus, intransitive verbs of movement or of existence/appearance occur with particular patterns, which differ from those observed with bivalent verbs, while for other verb classes there is greater crosslinguistic variability. In the Romance and Slavonic languages intransitives, reflexive intransitives[247] and reflexive verb constructions exhibit a tendency towards the labile VS order. In the Romance languages such patterns may also occur with transitive verbs constructed impersonally or intransitively or in the passive without an agent (cf. Sornicola 1994; Cennamo 1995).

Obviously, this is due in part to the typological peculiarities of individual languages or groups of languages: the Slavonic and Romance languages both have reflexive intransitive constructions, which are absent from, or found only very marginally, in other languages of the corpus (cf. English *to show itself*).[248]

The VS pattern with the previously defined verb classes or structures has been considered to be one of the devices which allow identification of the "thetic" or "presentative" function in various languages[249] (cf. Sasse, this volume).

Traditionally, locative-existential structures, which show a strong affinity with those just discussed, have also been placed amongst thetic or presentative structures.[250] In languages in which the structural elements of the locative – existential sentences are

((Adv / Loc phrase), 'be / exist / live', S),[251]

S appears in the same position relative to V as it does in structures with intransitive V. This is evident from the following examples:

(92) a. There is a cat in the garden.
      b. C' è   un   gatto   nel   giardino.
           LOC be:PRS.3SG INDEF cat:M.SG in:DEF.M.SG garden:M.SG
           'There is a cat in the garden.'

(93) V nekotorom   gosudarstve žil-byl
      in some:LOC.NT.SG state:LOC.NT live:PST(M.SG)-be:PST(M.SG)
      car'.[252]
      czar(NOM.M.SG)
      'In some land there lived a czar.'

444   Rosanna Sornicola

(94)   Na stene   byla   kartina.
        on wall:LOC.F.SG be:PST:F.SG poster:NOM.F.SG
        'On the wall there was a poster.'

In the descriptive tradition, which has made use of the designations "thetic" or "pre-sentative," emphasis has been given to the semantic and pragmatic values of the structures with VS patterns, while their syntactic structural properties are either over-looked or considered to be secondary.

On the other hand, in the generative literature, especially since the work of Burzio (1986), preference has been given to examination of the syntactic configurations associated with the VS pattern.[253] In Italian, intransitive verbs like *arrivare* 'to arrive' or *parlare* 'to speak' admit both SV and VS structures:

(95) a.   Arrivano   molte   persone.
           arrive:3PL many:F.PL person:F.PL
           'Many people arrive.'
      b.   Molte persone arrivano. 'Many people arrive.'
      c.   Parlano   molte   persone.
           speak:3PL many:F.PL person:F.PL
           'Many people speak.'
      d.   Molte persone parlano.
           'Many people speak.'

Whatever the pragmatic features that determine the SV or VS pattern, the two pairs of sentences (95a) and (95b) are not isomorphic, nor is the relationship between (95a) and (95b) identical to that between (95c) and (95d). *Arrivare* and *parlare*, in fact, belong to two subclasses of intransitive verbs, each behaving in different ways with regard to syntactic properties like auxiliary selection and the possibility of NP undergoing transformation by *ne*-cliticization. The *arrivare* subclass has been variously defined as "unaccusative" or "ergative" verbs.[254] The two subclasses have been assigned different configurational properties. The single argument of sentences with verbs like *arrivare* is generated in O-position at DS level, and from there it can be moved to S-position at SS level; on the other hand, the single argument of sentences with verbs like *parlare* is generated in S-position (i.e., the position Spec of INFL″) at DS level and can be moved to a position to the right of V (i.e., the position of an adjunct to INFL″) at SS level.

In generative models, therefore, the relationship between the surface S of un-accusative verbs and the GF O has been elegantly brought out. This relationship presents a problem crucial to the understanding of some of the VS structures. The generative model was initially elaborated on the basis of Romance languages such as Italian. We shall see, moreover, that interesting confirmation for this comes from ty-pological data relating to other languages (cf. Section 3.6). But both the examination of WO oscillations with one-argument verbs in individual languages (cf. Sornicola

1994, Sornicola 1995b for Italian and Spanish) and crosslinguistic examination of the characteristics of languages which differ in this respect show that the empirical reality is more complex than that hypothesized and requires the setting up of new models.[255]

A contribution to this may be offered by "microscopic" analyses of individual languages. These in fact seem to yield results useful for typological comparison. The fact that the lexical classes of verbs involved in labile patterns are surprisingly similar in many of the languages of the corpus is also a move in this direction. They may be identified and grouped according to a subdivision which was made on the basis of a microscopic analysis carried out on Italian, and confirmed in the comparative study carried out here: it concerns verbs of movement, verbs of saying, stative verbs, and verbs of change of state (cf. Sornicola 1994). Verbs of existence, movement, and saying, in particular, are found to be involved in structures with VS order in all the languages examined here, with the exception of Basque and Turkish (the latter, however, does not exclude them completely: cf. Section 3.4). This raises some interesting but for the moment not easily solvable problems concerning whether such verbs have properties in common and, furthermore, whether such properties are more central in the semantic characterization of the monovalent schema.

3.1.2.   Between the microscopic and the macroscopic

A microscopic investigation carried out on Italian has allowed a clear division to be drawn between two-argument (S, O) structures and one-argument structures (cf. Bernini 1995; Sornicola 1994, 1995b). The former have strong WO stability, even where types of text vary, and moderate variability with respect to the effect of pragmatic functions, while the latter as a whole show labile patterns, tending to gravitate around an equal probability of SV/VS.[256] Deviations from the ideal value of 50 % SV, 50 % VS depend on the type of one-argument construction, on the type of text and on pragmatic factors such as thematicization vs. focalization, backgrounding vs. foregrounding. However, individual data show that certain semantic factors cause a considerable shift from the equi-probable distribution of SV and VS. These are the lexical class of the verb, the *Aktionsart* features relating to it, and the value of the animacy feature of S. Certain verbs that denote a change of state (for example, *apparire* 'to appear', *crescere* 'to grow', *aumentare* 'to increase') are associated with the VS pattern at a much higher frequency than 50 %.[257] Verbs with the feature [+durative] tend to occur with the SV pattern, while verbs with the feature [+punctual] are more frequently associated with a VS pattern.[258] The animacy value of S is very important: in both Italian and Spanish one-argument verbs constructed with an NP (S) [−animate] occur with VS pattern in a high percentage of cases (cf. Sornicola 1995b).

The factor of inanimacy is strictly correlated with the value "eventive," which is often associated with structures with VS order: the entire construction describes an

446   Rosanna Sornicola

event or process without an agent, the constituent with the GF S having semantic
and syntactic properties typical of an O: this is semantically involved in the process,
as a P (recall the natural congruence between elements with the feature [−animate]
and the function P[259]); the position to the right of the verb – as in *brucia la casa* [lit.
'burns the house'] 'the house burns', *maturano le mele* [lit. 'ripen the apples'] 'the
apples are ripening', etc. – may be considered as a syntactic device that is function-
ally equivalent to morphological incorporation of N to V.[260]

Like the more general monoargumentality property, the semantic properties just
described are also found in the languages of the corpus as determinants, in various
ways, of characteristic WO patterns.

Of particular importance is the fact that in SVO and VSO languages, one-argument
structures with labile orders share certain pragmatic properties in addition to a spe-
cific textual property. As far as the former are concerned, depending on the associ-
ated prosodic patterns and the context in which they are used, one or other of the
following semantic interpretations is possible:

(i)   eventive: all the structure is in FOCUS;
(ii)   non-eventive, with focalization of the V constituent or of the S constituent.

As far as the textual property is concerned, in the Indo-European languages with
dominant SVO order, VS order is often associated with narrative progression. This
property is already found in the oldest Indo-European documentation[261] and per-
sists throughout the diachronic development of these languages (cf. Behaghel 1932;
Dressler 1969; Mitchell 1985). In Russian, where VS order may occur with both
monovalent verbs and those with more than one argument,[262] the phenomenon typi-
cally belongs to the so-called "epic style."[263] This textual property may also be found
in literary written registers of the Balto-Finnic languages of the corpus, where it may
be due to the influence of certain Indo-European languages.[264]

Correspondingly, in VSO languages such as Irish and Welsh, SV order with both
two-argument and one-argument verbs occurs in textual progressions of a narrative
nature in various diachronic stages (cf. MacCana 1973: 106–110, and here Section
2.4.6.5).[265]

The picture outlined so far may give an idea of the complexity of the interrelation-
ships between factors on various levels of analysis and the effect they have on the
structures under examination.

In the following sections, WO properties in structures with one-argument verbs
in various languages will be discussed. With respect to preliminary identification
of language groups according to the dominant order in two-argument structures, it
will be shown that there are important correlations between WO properties in two-
argument structures and those in one-argument structures.

3.2. SVO languages

3.2.1. Indo-European languages

Modern Indo-European languages vary with respect to differences in grammatical-ization of the labile variants SV and VS. In some of the Romance languages, such as Italian and Spanish, and in the Slavonic languages there exists a genuine oscillation between the two patterns according to pragmatic conditions, while in the Germanic languages, a tendency towards grammaticalization of SV order may be observed, although this is inconsistent both structurally (inversion phenomena are always possible, when position P_1 is occupied by an expletive constituent, such as German *es*, English *there*, or by a constituent not in situ) and sociolinguistically (in spoken registers and in dialectal varieties a considerable frequency of VS is found). The same considerations apply for a Romance language – like French – in which WO patterns have to a large extent been shaped by normativization from above.

Comparative analysis of Russian and Italian shows numerous affinities and some divergences. In a corpus of spoken Russian, Fougeron has shown structures that may be summarized by the following patterns:

$$S + V', S' + V, V' + S, V\ S', V\ S^{266}$$

The $S + V'$ pattern is characteristic of structures in which S is [+GIVEN] or introduces a TOPIC [+NEW], while V may correct previous information (cf. example (96)):

(96)   ['Are you at home? Good. Has father telephoned?']
        Otec   VERNÚL-SJA.
        father(NOM.M.SG) return:PFV:PST(M.SG)-REFL
        'Father has come back. [He didn't operate. The operation has been postponed.']

(97)   Otec VERNÚLSJA.
        'Father has come back. [He seemed tired, but happy. We understood that our departure had been arranged.]'

The $S'V$ pattern instead is always found in explanatory contexts:

(98)   ['There's a noise outside. Someone asks: "What's happening?" ']
        OTEC vernúlsja.

(99)   ['Give me the bag, let's go.']
        POEZD   idët.
        train(NOM.M.SG) arrive:3SG
        'The train is coming.'

448   Rosanna Sornicola

V′S and VS′ patterns represent, respectively, backgrounding of S (cf. (100), (101)) and focalization of S (cf. (102)). Both, however, have in common the fact that they contradict an utterance or an element in the preceding or following co-text:

(100)   ['He is always doing his work. Never a minute free.']
         IGRAET   Kirjuša.
         play:3SG   Kirjuša:NOM
         'He also plays. [Don't worry. He plays, goes out, everything in its time.]'

(101)   ['You are still writing with the pencil! They've repaired your pen, now
         you can write without spilling ink everywhere.']
         TEČËT   ručka.
         leak:3SG   pen:NOM.F.SG
         'It's still leaking. [I don't know how they repaired it, but even my bag is
         full of ink.]'

(102)   [– 'Will he come, Ljuda? What's up?'
         – 'He's got a headache. And he's bored without his daughter. He's left
         her at his parents.'
         – 'Do they work or are they retired?']
         Rabotaet   OTEC.
         work:3SG   father(NOM.M.SG)
         'Only her father works. [Her mother is retired and nurses her granddaugh-
         ter.]'

Finally, VS pattern is characteristic of eventive structures with S [−animate],[267] and occurs in descriptive textual progressions:[268]

(103)   Idët   dožd'.
         go:3SG   rain(NOM.M.SG)
         'It's raining.'

(104)   Dymit   kostër.
         smoke:3SG   fire(NOM.M.SG)
         'The fire is smoking.'

Almost all the patterns just described have counterparts in Italian, with corresponding types of context as well. The most striking difference concerns the V′S pattern. Structures isomorphic with (100) and (101) in the respective contexts seem to be typical only of spontaneous rather than planned speech (cf. Sornicola 1994).[269]

  It should also be pointed out that oscillation between S′ + V and VS is found in Italian in explanatory contexts such as those of Russian examples (98) and (99). For example, both (105) and (106) would be possible:

Interaction of syntactic and pragmatic factors   449

(105)   PAPÀ è   tornato.
         father be:PRS.3SG return:PST.PART
         'Father has come back.'[270]

(106)   È tornato papà.
         'Father has come back.'

3.2.1.1.   The (X)VS pattern

The Indo-European languages of the corpus all allow VS pattern as a labile order:

(107)   Poco fa è   arrivata   la   nave.
         little ago be:PRS.3SG arrive:PST.PART DEF:F.SG ship:F.SG
         'The ship arrived a few minutes ago.'

(108)   Entre   Gustave.
         enter:PRS.3SG Gustave
         'Gustave enters.'[271]

(109)   Prišla   Galina   Petrovna.
         come:PFV:PST:F.SG Galina:NOM Petrovna:NOM
         'Galina Petrovna came.'

(110)   Beneath these comes the free class of labourers.

(111)   Auf der   Strasse laufen   viele   Jungen.
         on   DEF:DAT.F.SG street   run:3PL many:NOM.PL boy:NOM.PL
         'In the street many boys are running.'

Examples (107)–(111) may be typically associated with contexts where a postverbal S introduces a completely NEW referent, or brings to the foreground a referent already present in the preceding co-text. It is also possible for each of them to have an interpretation in which the entire sentence is in FOCUS. The corresponding structures with related SV order would instead be normally associated with contexts where S is thematized.

   The same textual function as in sentences (107)–(111) may be realized by a VS pattern with verbs of appearance, as in the English example (112) and the Italian example (113):

(112)   Half reclined on a couch appeared Mr. Rochester.
         (from *Jane Eyre* [ed. Nal Penguin, New York, 1960: 123])

(113)   All'improvviso sbucò   dall'   angolo
         suddenly   come:PST.3SG out.from:DEF.M.SG corner:M.SG
         della   strada   un   mendicante.
         of:DEF.F.SG street:F.SG INDEF.M.SG beggar:M.SG
         'Suddenly a beggar came out of the corner of the street.'

450   Rosanna Sornicola

The VS pattern occurs as a preferred or possible option in all the Indo-European languages, not only with verbs of existence, appearance, and movement, but also with eventive structures in which a change of state verb and an S [−animate] appears. It is not unusual for the semantic value of the structure to concern a natural process, such as a physical development (of plants, etc.). In all these cases, some languages (English and German, but also French) have specific syntactic conditions:[272]

(114)   Nella   campagna di mio   padre   cresce   il
        in:DEF.F.SG estate:F.SG of my:M.SG father:M.SG grow:3SG DEF.M.SG
        grano.
        wheat:M.SG
        'In my father's estate wheat grows.'

(115)   Dans ce   jardin   grandissent de   nombreux
        in   this.M.SG garden:M.SG grow:3PL   PRTV many:M.PL
        pommiers.
        apple.tree:PL
        'In this garden many apple trees grow.'

(116)   Vyrosla   repka   sladka.
        grow:PFV:PST:F.SG turnip:NOM.F.SG sweet:NOM.F.SG
        'The turnip grew sweet.'[273]

(117)   Dort gedeihen Aprikosen.
        there grow:3PL apricot:NOM.PL
        'Apricots grow there.'

However, only languages such as Italian and Russian have the possibility of VS order in eventive structures in which S and/or V express a meteorological process:

(118)   Fioccava   la   neve   lentamente.
        fall:PST.IMPF:3SG DEF.F.SG snow:F.SG slowly
        'The snow fell slowly in large flakes.'

(119)   Nad Krakovom   nakrapyval   doždiček.
        over Krakow:INSTR drizzle:IMPFV:PST(M.SG) little.rain(NOM.M.SG)
        'It was drizzling over Krakow.'[274]

A somewhat different use is that of structures with verbs the semantic value of which is reduced to the expression of aspectual (or *Aktionsart*) and temporal features, while the semantic representation of the process is expressed by the noun, as in the following English and Italian examples:[275]

(120)   Whereupon began a stamping, clattering process.

(121)   È   scattato   l'   agguato.
         be:PRS.3SG spring:PST.PART DEF.M.SG   ambush:M.SG
         'The ambush has been sprung.'

Such structures form part of the more general tendency to VS order, but constitute a separate subtype. A related subtype, but one which it is preferable to distinguish from the previous, is that in which the verbal constituent is a verb of movement, whose semantic value is reduced to a mere temporal operator. It is interesting that these structures are often not textually independent; they may often be part of a descriptive progression and function with respect to a previously expressed process, within a broader macrostructure (cf. also examples (108), (109), (112)).

(122)   Then came a deep strong sob.

(123)   Here ensued a pause.

(124)   During and after the final break-up of the Roman empire came times of confusion.

(125)   Es   kommt   ein   Gewitter.
         EXPL.Prn.3SG come:3SG INDEF.SG   storm
         'A storm is coming.'

(126)   [A hysterical subject was freed from most of his symptoms: he felt happier, more energetic and more intelligent and hardly thought of his hypnotist.]
         Suivait   une   seconde   phase, celle   de la
         follow:3SG.IMPF   INDEF:F.SG second:F.SG phase that:F.SG of DEF.F.SG
         passion somnambulique.[276]
         frenzy   of.sleepwalking
         'There followed a second phase, that of the sleepwalker's frenzy.'

(127)   Pošël   dožd'.
         go:PRFV:PST(M.SG) rain(NOM.M.SG)
         'It started to rain.'

English *there*-constructions – with their characteristic VS order[277] – may contain intransitive verbs like *rise*, *come*, *occur*, *spring*:

(128) a.   There rose in his imagination grand vision of a world empire.
      b.   There may come a time when the Western Nations will be less fortunate.
      c.   Not long after this, there occurred a sudden revolution in public taste.[278]
      d.   There sprang up a wild gale that night.

452   Rosanna Sornicola

There is "a rather less common, more literary type of existential clause" than the *there*-construction with the verb *be*.[279]

Finally, in all the Indo-European languages there is oscillation with verbs of saying. In some languages, consistent alternation between the two patterns is found in the written language: SV if the quotation follows, VS if the quotation precedes. In Italian the same tendency may be observed, although with a lesser degree of consistency (cf. Sornicola 1994). In written English, verbs of saying also allow the VS pattern, especially with full (i.e., nonpronominal) NP.

3.2.2.   The Balto-Finnic languages

Also other SVO languages, such as Finnish and Estonian, have inversion of S and V as a labile order with one-argument verbs[280] and occurrence of the structure is determined by semantic and pragmatic factors. The pattern is found with verbs of movement and with locative-existential structures, particularly when S is indefinite (note that an indefinite S is normally marked as a partitive[281]); the prosodic correlate of these syntactic structures is nuclear stress falling on S:

(129)   Kadulla   juoksee   POIKIA.
         street-ADESS run:PRS.3SG boy.PRTV.PL
         'The boys are running in the street.'

(130)   Eteisessä   on   HERRA.
         hall-INESS be:PRS.3SG man.PRTV.SG
         'In the hall there is a man.'[282]

VS order is also found with verbs of saying when S is in FOCUS:

(131)   Sanon   MINÄ.
         say:PRS.1SG 1SG.ps.prn.NOM
         'I say.'[283]

3.3.   VSO languages

Particularly interesting is the situation in the Celtic languages, where the dominant order in sentences with two-argument verbs is VSO (cf. Section 2.4.6). In Irish and Welsh, structures with one-argument verbs and locative-existential structures do not differ from structures with a transitive verb in patterns of order. As in the latter structures, a VS pattern may in fact be found or an SV pattern, related to the pragmatic factors of topicalization or focalization of S:

(132)   Dydd Sul a   ddaeth.
         Sunday   REL.PTL come:PST.3SG
         'Sunday came.'

(133)   Dydd SUL a ddaeth.
'SUNDAY came.' = 'Sunday, and not Saturday, came.'[284]

(134)   Ar   y   môr   yr   oedd   llongau.
on   DET.Art sea.SG preVPTL be:PST ship.PL
'On the sea there were ships.'[285]

This is the same as saying that the Celtic languages, like SVO languages and despite the structural differences between them, have a labile rule of WO with one-argument verbs. More generally, on the basis of what was observed in Sections 2.4.6.5 and 2.4.6.6, the hypothesis may be put forward that Celtic languages exhibit labile rules with both transitive and intransitive structures.

In some dialectal varieties of Irish (in particular in the Munster dialects, but also sporadically in those of Connaught and Donegal) V (. . .) S order is found in narrative style and in structures with intransitive verbs of existence, movement, and appearance, especially if the constituent which carries such a function is "heavy." In other words, an AdvP or PrepP may intervene between V and S (recall that this would not be possible in the dominant order of sentences with two arguments: cf. Section 2.4.6.1).[286] Even though rhythmic factors play a role in this phenomenon, it is possible that it is due to factors more typically connected with grammatical relations (cf. Section 3.6).

3.4.   SOV languages

In SOV languages, the occurrence of VS orders in structures with one-argument verbs is related to focalization of V and backgrounding of S, according to the same strategy which operates in two-argument structures (cf. Section 2.2.1). On the other hand, as in two-argument structures, the constituent in FOCUS occurs immediately before the finite V (cf. Section 2.2.1).

Both phenomena may be found in the following Basque examples:

(135)   Piarres   hil   da.
Peter-ABS die:INF be-PRES.3SG
'Peter is dead.'

(136)   Piarres da hil.
'It is Peter who is dead.'

(137)   Hil da Piarres.
'He is dead, Peter.'

(138)   Zizeronek   dio.
Cicero-ERG say:PRS.3SG
'Cicero says.'

454   Rosanna Sornicola

(139)   Badio   Zizeronek.
         INT.PTL+say:PRS.3SG Zizeron-ERG
         'SAYS Cicero.'[287]

The situation in Turkish shows some affinities with that in Basque, but also some peculiar characteristics, consistent with what was found in transitive structures (cf. Section 2.2.1). With one-argument verbs the pattern:

S (...) V

is possible, as well as:

V (...) S

Each, however, is conditioned by different semantic and pragmatic factors.

The semantic factors concern the incidence of definiteness and animacy features. If S is [+definite], it may occur in both preverbal and postverbal positions:

(140)   Murat dün   Ankara-dan dön-dü.
         Murat yesterday Ankara-ABL return-PST
         'Murat returned from Ankara yesterday.'

(141)   Dün Ankaradan döndü Murat.
         'Yesterday Murat returned from Ankara.'[288]

If S is [−definite] and [+animate], it cannot occur in postverbal position,[289] but may occur in P_1 (although this gives rise to a marked order[290]) or in the immediately preverbal position; finally, if S is [−definite] and [−animate], it is restricted to the immediately preverbal position.[291] Compare examples (142) and (143) with examples (144) and (145):

(142)   Bir kadın   biz-e   doğru   koş-uyor.
         one woman we-DAT toward run-PROG
         'A woman is running toward us.'

(143)   Bize doğru bir kadın koşuyor.
         'A (some) woman is running toward us.'

(144)   Akşamları burada bir   rüzgar es-iyor.
         Evenings   here   one   wind   blow-PROG
         'A wind blows here in the evenings.'

(145)   *Bir rüzgar akşamları burada esiyor.[292]

Interaction of syntactic and pragmatic factors   455

The same restrictions apply to locative-existential constructions.[293]

As far as the pragmatic conditions are concerned, the two structures (140) and (141) differ in that in the first, *Murat* is the TOPIC, while in the second the TOPIC is *dün* and *Murat* is a backgrounded constituent as it occurs in the sentence space bounded by V on the left (postverbal field). S, like any other postpredicative constituent, may not carry nuclear stress.[294] Unlike what is found in SVO languages, the order V (...) S with one-argument verbs can never have the value of focalization of S. When S occurs in the preverbal space, the neutral stress coincides with the constituent immediately in front of the verb, as in transitive structures, in other words, with the constituent in FOCUS:

(146)   Çocuk o   oda-dá   uyu-yor.
          child that room-LOC sleep-PROG
          'That child is sleeping in that room.'

(147)   O oda-dá çocuk uyuyor.
          'A (some) child is sleeping in that room.'[295]

However, stress may coincide with V, and in such a case it determines the interpretation of V as FOCUS. This may happen when it occupies its canonical position as well as when it occupies position P$_1$ (in such a case, V obligatorily carries stress and is a marked FOCUS). On the other hand, contrastive stress may fall on preverbal S, determining a marked FOCUS on that constituent:

(148)   KAMYON   devril-miş.
          truck   roll.over-PST
          'A truck (not a train or a bus) rolled over.'
          'The truck (not the train or the bus) rolled over.'[296]

In Basque and Turkish, therefore, the patterns of order with one-argument verbs cannot be represented with a labile rule, but reflect the same properties as structures with two-argument verbs.

3.5.   Hungarian

Hungarian once again shows behavior only partly akin to Basque and Turkish (cf. Section 2.3.2). As in SVO languages, one-argument verbs whose semantic value is described by Behrens (1989: 116) as "Geschehen/Sich-Ereignen; Sich-Verändern; Gehen/Sich-Entfernen; Kommen/Erscheinen; Entstehen/In-Existenz-Treten; Existieren; Existieren auf eine bestimmte Art und Weise" typically occur with VS order.[297] Verbs of saying may also occur with VS order.[298] However, in Hungarian, as in SVO languages, the verbs in question do not come under the heading of truly grammaticalized WO patterns, but of labile patterns: there is in fact oscillation between VS

456   Rosanna Sornicola

and SV throughout the entire diachronic development of the language.[299] Factors concerning the semantic representation of the verb and/or focalization of the entire construction or of V are fundamental in determining VS pattern. However, only some VS structures may be correlated with phenomena of focalization.[300]

Where Hungarian shows typological convergence with SOV languages such as Basque and Turkish is in the influence of the semantic features of definiteness and referentiality on WO patterns. As in the structures with transitive V, such features are in fact decisive (cf. Section 5.1). Modern Hungarian has a distribution where if S is [+definite] it generally occurs after V, while if S is [−definite] and [−referential], it occurs before V.[301]

3.6.   Between universal tendencies and typological conditions

3.6.1.   Structural factors and statistical regularities in unstable orders

The picture drawn so far raises interesting problems of a general nature. One may in fact wonder in the first place why the tendency to oscillation in the relative order of S and V with one-argument verbs is so widespread amongst the languages of the corpus, even in languages that – due to their specific syntactic and/or pragmatic conditions – do not favor flexibility of WO. In the second place, one may ask oneself if it is to be considered a universal tendency. If the answer to this question were to be in the affirmative, there would remain, however, to explain the relationship which exists between the universal tendency and the typological particularities exhibited by the various groups of languages.

The widespread presence of the labile orders SV, VS amongst the languages of the corpus is presumably due to a property of the argument structure: while bi- (or tri-)argumental structures are in themselves more or less stable with respect to WO, even in languages with morphological case, which, as is well known, have a greater flexibility of WO,[302] one-argument structures are inherently unstable with respect to WO. A serious problem posed by the analysis of these structures is constituted by the fact that, up until now, they have been described according to models constructed for two-argument structures. Comparison between the two types of structure, however, may be useful from a typological point of view in order to construct a correlation.

It is possible that the presence of labile patterns SV, VS is related to a probabilistic rule: On the basis of a statistical examination carried out on Czech texts Uhlířová (1969) has demonstrated that there is a correlation (i) between the number of constituents in a sentence and the number of possible patterns of order it may have (which is entirely expected for a language with flexible order); (ii) between the number of constituents in a sentence and the degree of oscillation between one pattern and another. In other words, increasing the number of constituents increases the rigidity of WO: in this case, a single predominant pattern of order is

found with a rather high frequency, while the remaining patterns have a much lower frequency. Instead, decreasing the number of constituents decreases the rigidity of WO: in this case, there are few patterns, the frequencies of which tend to be similar, until one arrives at an equal probability of VS and SV in sentences with just V and S. By applying Uhlířová's model, similar results have been obtained for the Romance languages.[303]

However, this does not constitute an exhaustive explanation of the phenomenon of labile orders. There are many reasons for thinking that it is the argument structure itself and its relationship with the linear axis which provides further explanation. The presence of labile orders SV, VS has not only to do with the number of constituents of the sentence, but also – more specifically – with the relationship between V and its single argument.

3.6.2.   Some problems with generative models

One solution could be to turn to generative models. The hypotheses described in Section 3.1.1 may in fact offer an explanation of the patterns of order in terms of the GFs of NPs and of their convergent linearization in the VS configuration. The NP constituent which appears to the right of V at SS is either generated in this position at DS, and therefore has the GF of O at this structural level of representation, or it is an S at DS level that is adjoined to the position to the right of V at SS. Both DS representations converge on a single linearization, that is, the V S configuration. In any case, a GF typical of two-argument structures is assigned to it.

Despite the elegance of the model, three significant problems may be identified.

In the first place, it does not take into account oscillations between SV and VS, which determine what, according to Daneš, has been called the labile order. At first sight, this seems to be a mere performance phenomenon, or a phenomenon which could be given an explanation in terms of pragmatic mechanisms, extraneous to structural rules. In the view proposed here, it is rather a *structural* property and as such needs to be dealt with.

The second difficulty concerns a far more important theoretical question, that is, the derivability of grammatical relations from configurational relations in generative models. This aspect highlights a certain circularity in the explanation of SV and VS patterns. In other words, the problem concerning the order with inversion of S and V in some languages raises the final and urgent question of whether it is position which determines grammatical relation, or whether the opposite is the case. It is true that the bifurcation whereby the NP collocated with V, which is S or O at DS level, is determined on the basis of tests independent of WO, such as that of auxiliary choice ('to be' or 'to have') and of *ne* (whether or not the constituent allows transformation with *ne*). However, such tests do not seem solid enough to guarantee identification of the GRs S or O across a wide crosslinguistic range.[304]

458   Rosanna Sornicola

More generally, the problem of identifying the GRs S and O goes hand in hand with the well-known impossibility of establishing absolute criteria for subjecthood and objecthood, which emerged from research in the 1970s. This difficulty is greater for O than for S.[305]

In reality, it is the typological analysis which offers interesting data for reflection. At first sight these seem to confirm, at least partially, the generative model.

In all the languages of the corpus, oscillation between the two patterns of WO is strongly biased in favor of the order with S in what would be the canonical position of O in two-argument structures (and therefore, VS in SVO languages, SV in SOV languages[306]) in conditions which are determined for the most part by the features of animacy and/or definiteness of S. As has been seen, in SVO languages the eventive VS pattern can occur in an entirely consistent manner: in the Romance languages and in the Slavonic languages when S is [−animate]; in Finnish when S is [−definite]. In SOV languages the SV pattern, with S in the immediately preverbal position, occurs when S is either [−definite] or [−animate] (cf. Section 5.1). This could constitute crosslinguistic confirmation of the hypothesis of the unaccusativity of some verbs, all the more so if one bears in mind that in languages such as Finnish, NPs with the feature [−definite] are marked with partitive, the case which also marks O NPs in particular conditions (cf. Karlsson 1987: 79–81).

However, this conclusion is not very convincing for several reasons: (i) the property of unaccusativity essentially concerns verbs, while it is clear that in the phenomena examined, what is at issue are the semantic features of the NP argument of the verb; (ii) although also characterized in semantic terms (cf. Belletti 1988), the model of unaccusativity is formally defined in terms of configurational properties and, derivatively, of the GF O. On the other hand, what seems to be crucial in the phenomena examined are the semantic features [−animate] and [−definite] of N; as is well known, these features have a natural affinity with the semantic role of patient, but this is a semantic, not a syntactic, characterization. In many languages, in fact, a verb of movement whose sole argument is an NP with the feature [+animate] and/or [+definite] appears in structures which it is doubtful that the unaccusative model can deal with.

In any case, only certain structures with a one-argument verb and S [−animate] and/or [−definite] appear to be compatible with the hypothesis of unaccusativity, and not the far greater number defined by tests of auxiliary choice and *ne*. The only thing both sets of verbs have in common is a monoargumental schema. Now, as has been seen, the correlation between monoargumentality and labile patterns is more or less strong in the languages of the corpus.

All this suggests that it is the one-argument structure and the semantic features associated with the arguments, and not the assignment of the GF S or O, which determines the characteristic oscillation of labile SV, VS patterns. This seems to be confirmed by a further result of the enquiry.

From the examination carried out previously, a nontrivial correlation between WO properties in two-argument structures and WO properties in one-argument structures has emerged. As has been seen, while VSO languages can be said to have labile orders in both two-argument and one-argument sentences, SOV languages show a slight difference between patterns of WO in the two structural types. Both types of languages have in common the fact that S and O are found in the same space with respect to V: to the right of V in VSO languages, to the left in SOV languages. On the other hand, SVO languages have the strongest oscillation of WO pattern with one-argument verbs. As is well known, in these languages S and O are found in two different (opposing) sentence spaces with respect to V, and symmetrical inversion of the order of S and O may happen only in non-neutral cases.[307] Now, it should be possible to advance the hypothesis that a higher incidence of labile orders is found in these languages because in one-argument structures the portion of the configuration necessary for codifying GFs is "liberated." Correspondingly, VSO languages exhibit labile orders both in two- and one-argument structures, because the space to the left of V, which may be used for processes of topicalization and focalization in two-argument sentences, is free (in other words, it is not needed for codifying GFs). On the other hand, in SOV languages, where the space to the right of V is free, this space may not be occupied except by backgrounded material in both two- and one-argument sentences, and therefore labile SV, VS orders are not possible.

The following scale of divergence may therefore be established between two- and one-argument structures:

    1) SVO languages   strong divergence
    2) SOV languages   slight divergence
    3) VSO languages   no divergence

To sum up, it may be claimed that the co-presence of labile orders is a tendency operating in a more or less strong manner in typologically different languages. The oscillations between SV and VS orders found in many languages are due to argument structure, which – as far as the property of order is concerned – allows "neutralization" of the opposition between S and O, that is, of the differences between the two GFs. This raises a number of theoretical and typological questions concerning subjecthood. In some languages of the corpus, as in Russian and German, the single argument of a one-argument structure "looks like" an S for the nominative and agreement markers it assumes. In Basque it is codified as an "ergative" or an "absolutive" (see examples (135)–(139)) depending on its thematic relations with the verb. In Italian and Spanish, the only subjecthood feature of this argument in the phrase structure representation is agreement with V. In these languages, the GF of the only argument of V in one-argument structures is not that of a true S, but rather what can be defined a "subjectoid" (see Sornicola 1990 and, for an alternative theoretical treatment, Sornicola 1999). Although these differences are not devoid of interest,

460   Rosanna Sornicola

they seem to be mechanically determined by morphological properties of the various languages; in other words, languages with morphological case and N–V agreement must assign such features.

At a more abstract level, one can pose the question of the extent to which order may be considered a device for the codification of GFs (cf. Section 6).

3.6.3.   Typological conditions

There remains, in fact, the explanation of the relationship between universal tendency and typological conditions. Universality seems to be associated with monoargumentality of V and with the effects that this property has on constituent structure and on order relations. As far as linearization is concerned, statistical regularities, such as those observed by Uhlířová for Czech, may also have an effect on this argumental property.

On the other hand, typological constraints seem to be associated with less abstract (deep) properties, which are related to the differing phrase structures characterizing the various languages. For example, the characteristic constraint of modern Germanic languages and modern French, for which the verb must have an NP with the function S to its left, constitutes a syntactic conditioning of the more general tendency. However, as has been seen, this may manifest itself in VS structures of a particular type, in which the preverbal position P_1 is occupied by an expletive (an adverb, such as *there* in English, an expletive pronoun such as *es* in German, or *il* in spoken French).

It seems interesting that the microscopic analysis carried out on individual languages, such as Italian and Spanish, shows that equi-probability of SV and VS is more likely in spontaneous speech than in written styles (cf. Sornicola 1994, 1995b; Milano 1994).

In fact, the universal tendency may also be observed in the spoken registers of languages which in the written registers show restrictive typological conditions: this is the case in spoken German and some German dialects and in Finnish (the phenomenon of inherited VS order with verbs of saying in Finnish dialects comes to mind).

The VS type with one-argument structures is documented in both colloquial German and German dialects:

(149)   Kommt   da   plötzlich ein   Kerl   herein.
         come.PRS.3SG then suddenly INDET.ART.SG guy.SG in
         'Then suddenly a guy enters.'

(150)   Spricht   der   Maan.[308]
         say.PRS.3SG DET.ART.SG   man.SG
         'The man says.'

(151)   [geːt   diː   doː   nuner]
 go.PRS.3SG DEM.NOM.F.SG there down
 (= NH German: Da geht sie hinunter).[309]
 'There she is, going down.'

As can be seen, in these examples V is in P_1.

In Finnish dialects, the VS pattern is common in sentences with verbs of saying occurring after a reported speech (the so-called "*inquit*-clauses": cf. example (131)) and is considered a native natural development driven by pragmatic features, unlike the inversion due to Swedish influence and typical of the written style.[310]

Diachronic analysis may also bring to light the provisional and sometimes artificial nature of typological conditionings: the VS type with V in P_1 was entirely possible in Old English and in Old French (for a discussion of some problems in the relationship between historical factors and typology, cf. Sornicola 1998).

Among the typological conditionings, the difference between languages which also allow the pattern with inversion of V and S in two-argument structures, and languages which allow it only in one-argument structures, is worth considering. As is well known, the first option is very rare in configurational languages, while it is allowed, albeit in particular styles, in Russian (cf. Section 4.2).

4.   V or sentence in FOCUS

4.1.   V in FOCUS

Most of the languages of the corpus show a specific WO pattern in structures in which FOCUS is placed on the main verb. These are cases where V occurs in P_1. It is interesting that such a pattern is shared by languages with very different structural properties and basic WOs, such as Basque, Turkish, Finnish, Hungarian, Italian, and Spanish. The resulting structures are highly marked in these languages: FOCUS on V is mostly contrastive/corrective, or at least "emphatic";[311] furthermore, V in FOCUS in P_1 always requires the nuclear stress.[312] The data gathered so far support the conclusion that the pattern has a low frequency of use in the languages investigated. In Finnish, for example, sentences with the pattern in question, such as (23e)–(23f), are grammatical, but it is as well to remember that V in initial position is, on the whole, a highly restricted phenomenon.[313] In Hungarian, the pattern (exemplified in (15)) is presumably equally infrequent.[314]

Structurally, there are problems in Italian with respect to V′OS or V′SO orders; sequences such as:

(152)   ??AMA   Mario Lucia.
 love:PRS.3SG Mario Lucia

462   Rosanna Sornicola

(153)   ??AMA Lucia Mario.

are not acceptable, but require separation of the constituents by means of a pause or variation in the pitch in order to work, as in:

(154)   AMA Mario / Lucia.
         'Lucia loves Mario.'

(155)   AMA / Lucia / Mario.
         'Lucia loves Mario.'

In other words, V in FOCUS in P_1 entails a narrowing of the sentence space, which is reduced to only the V position, or possibly the VP position, while S must be confined to the "post-field," usually as an afterthought or backgrounded information.[315] As can be seen from example (155), V may not be separated from S unless the entire structure is segmented (type V + TAIL1 + TAIL2).[316]

   The possibility of V in FOCUS in P_1 is perhaps also allowed in Russian, as can be seen from the following example quoted by Kovtunova (1976: 182):

(156)   POSETIL   laboratoriju   profésor.
         visit:PST.M.SG laboratory.ACC   professor.NOM.SG
         '(Then) the professor visited the laboratory.'

Kovtunova considers this pattern an "expressive variant," typical of the spoken language, whose prosodic schema probably includes two nuclear stresses (V′OS′ schema). Description of the stress schema in (156) brings to mind the fact that, in this respect, Russian has the same property of isolation of S already pointed out for Italian (cf. Section 2.4.4.1.1). Fougeron (1989: 330), however, questions the reliability of the prosodic pattern described by Kovtunova. She observes that in her own corpus of spoken Russian, the few cases of V-initial sentences only appear with nuclear stress on the last constituent (types: VOS′, VO//S′). These structures only occur in narrative progressions, where they signal narrative continuity.[317] The analysis of sentence (156) by two native speakers points to the following characteristics: (a) it may occur in both spoken and written registers; (b) it marks narrative continuity; (c) it may have main stress on V. Judgement (c) is, of course, a mere perception, which cannot be easily accommodated into either Kovtunova's or Fougeron's analysis.[318]

   That the structure under examination occurs in Turkish and in Basque – languages with a basic SOV order – is of great importance from a typological point of view:

(157)   Işte   bil-di-m   ben böyle sıkışıklığ-a
         there   know-PST-1SG I   thus   congested-DAT
         gel-eceğ-in-i   bu   iş-in.
         come-FUT.NOM-POSS-ACC   this matter-GEN
         'There, I knew that this matter was going to be blocked like this.'[319]

In this sentence, "the predicate, which is sentence-initial and stressed, carries the most crucial information in the utterance and hence is foregrounded while the rest of the sentence is backgrounded" (Erguvanli 1984: 61).

Lafitte (1962: 48) observes that in Basque when V is in FOCUS, it occurs in initial position (the example is, however, problematic because V is one-argumental). The structure with *egin* 'to do', which is typical of the dialects of Vizcaya and Guipuzcoa, is also interesting: nonfinite lexical V occurs in VP initial position, followed by *egin* and by the conjugated auxiliary (cf. Rebuschi 1984: 71–73). The structure appears to be associated with two pragmatic conditions: "d'une part, il faut que l'action soit effectivement nouvelle (et donc susceptible de créer un certain effet de surprise chez l'interlocuteur ou le lecteur), et d'autre part, il faut généralement qu'elle s'insère dans une série d'événements se produisant les uns après les autres, sans connexion causale particulière d'ailleurs" (Rebuschi 1984: 72). In some dialects, especially those of eastern Guipuzcoa, the structure has a true contrastive value on V (cf. Rebuschi 1984: 73).

Structures such as (157) in SOV languages with FOCUS linked to the immediately preverbal position contribute to the undermining of the d-configurational model. They constitute empirical data which are meaningful on more than one front: on the one hand, they show the limits of a model which defines PFs on the basis of position; on the other hand, they suggest that P_1 is a powerful position for definition of the PF of marked FOCUS. The fact that the pattern with V in FOCUS in P_1 occurs in languages that are typologically very different shows that pragmatic tendencies may override the structural and typological properties of individual languages. This is further confirmed by the Celtic languages, which have basic VSO order. Even in these languages, where P_1 is occupied by V in neutral sentences, V may occupy P_1 as a result of the characteristic structure of clefting of the verbal noun and the main verb ('to do' or 'to be' occurring in the dependent clause), as in the following Welsh example:

(158)   Edrych ar y   gannwyll yna 'roeddwn   i.
         looking on DET.ART candle.SG there be:PST.IMPF.1SG 1SG.aff.ps.prn
         '(It was) looking at the candle there that I was.'[320]

Such structures are frequent in colloquial Welsh, with the function of focalization of the action expressed by the verbal noun; however, from the frequency of the structure, it would seem that a demarking process is emerging and becoming generalized in the spoken language, that is, the type is becoming a neutral option.[321]

It must be emphasized, in any case, that in many of the languages quoted, the structure with V in FOCUS in P_1 is an alternative to others which do not exploit WO to convey FOCUS on V.

464   Rosanna Sornicola

4.2.   Word order when sentence is all-in-FOCUS

When FOCUS is distributed over the entire sentence rather than coinciding with an argument constituent, interesting structural peculiarities may be observed. It is worth emphasizing once again preliminarily that the property of being all-in-FOCUS pertains to a higher domain of analysis, which is superimposed on the sentential domain (cf. Section 1.8). An interesting theoretical characteristic of the shift of domain is the possible difference between the pragmatic values of the individual constituents at the sentential level and the overall pragmatic value of the structure at the textual level, which may be considered a case of pragmatic noncompositionality.

The pragmatic property of being "all-in-FOCUS" has so-called "eventive" interpretation of the sentence as its semantic correlate and frequently (though not always) "mono-argumentality" of V as its structural correlate (see Section 3; the problem of all-in-FOCUS sentences has especially been dealt with in this volume in Sasse's contribution).

Not all the languages of the corpus use WO variations to encode structures with S all-in-FOCUS. English and German retain the SVO order and make use of prosody: in both languages, in fact, the eventive/all-in-FOCUS interpretation is associated to a characteristic prosodic contour with stress on the argument carrying GF S (cf. Cruttenden 1986: 83; Ladd 1996: 188–189, 199). French makes use of existential (presentative) structures with the cluster of bound morphemes *il y a* in P_1 (cf. Lambrecht 1994: 177 ff.). In Italian WO is only one of the possible correlates of all-in-FOCUS sentences: the nonbasic VS order can be used, but only when V is a one-argument predicate (cf. Sornicola 1995b); as an alternative, an existential (presentative) syntactic type with the whole sentence preceded by *c'è che* (lit. 'there is that') can be used (cf. Bernini 1991, 1995).

The available data indicate that Celtic languages use WO variations as well as the cleft type with the verb 'be' for codifying of the property in question. Apart from the Irish and Welsh structures with initial NP in explanatory contexts, already examined in Section 2.4.6.5, an Irish type with copula + personal pronoun + relative clause, may be quoted:

(159)   Ba   é   a   bhí   cosamhail
        be:PST.3SG 3SG.ps.prn.M REL.PTL be:PST.NREL.3SG alike
        len'   athair   ar lorg a   leicinn.
        with.his father.SG on track 3SG.POSS.M cheek.GEN
        'He looked like his father from the side-view.'[322]

It must be pointed out, however, that in the literature on Irish and Welsh, the discussion of the structure in question as of those examined in Section 2.4.6.5, suffers from an unclear definition of FOCUS.

Establishing the situation in Russian raises problems of a theoretical nature. Fougeron (1989: 306–310, 424–425) found V-initial (VS′O, VOS′, VSO′) orders in his corpus of spoken language. These patterns are typical of narrations (cf. Section 3.1.2) with V always in the past tense (cf. Fougeron 1989: 268):

(160)   V S′ O
[On Saturday we had friends to dinner. Just as I was making the meal, I notice that the lid of the pressure-cooker was not working. I tried turning the screw in every direction, but nothing doing.]

Vzjala   MAMA   kryšku.
take:PST:PRFV.F.SG   mother.NOM.SG   lid.ACC.SG
'Then Mother had a try.' (lit. = 'Mother took the lid.')
[She turned the lid in every direction, with no better result. Eventually I had to get out the cast-iron pan.]

(161)   V O S′
[I've tried everything, she's never happy. With Volodja it's different. Look, for example, she needed an alarm clock, I went to buy one. An ordinary alarm clock. Of course, it wasn't right: it was too big, the tick was too loud, the ringing was too jarring . . .]

Kupil   časy   VOLODJA.
buy:PST.PRFV.M.SG   watch.ACC.PL   Volodja.NOM
'So Volodja went and bought another one.'
[He was congratulated. "Volodja always finds exactly what is needed." Of course. If it had been me who had bought it, she would have said that it was too expensive.]

(162)   V S O′
[Never happy Tatka. Volodja gave her a leather bag, made in Yugoslavia. Well, she doesn't need a bag at all, it's a watch she needs, hers doesn't work any more.]

Kupil Volodja ČASY.
'So Volodja bought her one.'
[That still wasn't right: she wanted one with a round face.]

Fougeron considers examples (160)–(162) to be all-in-FOCUS structures (all rhematic, in her terminology), despite the fact that in each of them the nuclear stress is in a different position. As far as organization of the message is concerned, they would be structures with a single constituent (cf. Fougeron 1989: 417–418). However, it must be pointed out that Fougeron regroups structures (160)–(162) into a homogeneous type on the basis of analysis of a prosodic property.[323] Examination of

466   Rosanna Sornicola

the contextual conditions of the structures in question, instead, could lead to a different regrouping: only (162) seems, in effect, a case of an all-in-FOCUS structure, while in (160) and (161), only the S constituent is in FOCUS.

The prosodic property of establishing a single constituent seems to characterize also all-in-FOCUS structures in languages with less flexible order than Russian (see the discussion of data for English, German, and Dutch in Ladd [1996: 234]).

4.2.1.   A few conclusions on typological properties of all-in-FOCUS structures

The data available therefore allow some general principles to be hypothesized, which may be formulated as follows:

(A) Where a language uses WO to codify all-in-FOCUS sentences, the WO used is a marked pattern in that language.

(B) It may be sufficiently (minimally) described according to the constituent which occupies P_1, such that:

   (a) if P_1 is canonically occupied by the constituent with GF = S, it will be occupied by V in the condition under examination;

   (b) if P_1 is canonically occupied by V, it will be occupied by a constituent with GF (S or O) in the condition under examination.

Principle B (a, b) may be called "mirroring" of the predicative and referential functions. This may explain why only languages such as Russian or the Celtic languages discussed allow use of WO to codify all-in-FOCUS sentences, while other SVO languages either do not allow it at all, as for example English, or allow it only with one-argument verbs, as Italian and Spanish.

5.   The influence of semantic or textual features on word order

5.1.   The influence of the semantic features animacy, definiteness, and referentiality on word order

In some of the languages of Europe, the semantic features of animacy, definiteness, and referentiality associated with N have an important influence on the organizing of WO patterns, to the extent of predominating over GF. In other words, these features are actually grammaticalized with respect to WO in the Slavonic languages and in some Uralic and Altaic languages. We shall discuss here the properties which relate to the languages of our corpus: Russian, Turkish, and Hungarian.

In Russian, animacy is an important factor controlling WO. Regarding a sentence such as:

(163)   Na kogo-to   čto-to   upalo.
         on someone.ACC something.NOM fall:PST.PRFV.NEUT.SG
         'Something fell on somebody.'[324]

any permutation of the constituents would give rise to a less acceptable structure. As Yokoyama observes, "being an animate noun is more important for claiming utterance-initial position than being a nominative subject."[325] In the example, in fact, S, which has the feature [−animate], follows an NP (dominated by a PrepP node) with the feature [+animate].[326] Comparison of the Russian structure with the English translation and with the Italian translation *Qualcosa è caduto su qualcuno*, isomorphic with the English with respect to WO, shows how the behavior of Russian differs significantly from that of languages in which GFs predominate over semantic features.

The features of animacy, definiteness, and referentiality have particular importance for WO in Turkish. As far as the preverbal space is concerned, Erguvanli shows that the features of animacy, definiteness, and referentiality of NPs are the main factors influencing variation in the order of basic constituents.[327] Such factors are independent of each other.

When all the NPs in a sentence have the feature [+definite], there are no restrictions on WO. From a structure such as:

(164) a.   Murat para-yı   bu adam-a   ver-di.
             Murat money-ACC this man-DAT give-PST
             'Murat gave the money to this man.'

with neutral WO, structures (164b)–(164f) may be derived; each of these structures has a marked order, being both nonbasic and pragmatically marked:

(164) b.   Murat bu adam-a para-yı ver-di.   (S IO O V)
            TOPIC *Murat* / FOCUS *para-yı*
      c.   Para-yı Murat bu adam-a ver-di.   (O S IO V)
            TOPIC *para-yı* / FOCUS *adam-a*
      d.   Para-yı bu adam-a Murat ver-di.   (O IO S V)
            TOPIC *para-yı*/ FOCUS *Murat*
      e.   Bu adam-a Murat para-yı ver-di.   (IO S O V)
            TOPIC *bu adam-a* / FOCUS *para-yı*
      f.   Bu adam-a para-yı Murat ver-di.   (IO O S V)[328]
            TOPIC *bu adam-a* / FOCUS *Murat*

However, if one of the NPs has the feature [-definite],[329] it will occur in a different position in the sentence, according to its value with respect to the feature [+/−animate]: if the NP is [+animate] its position is not fixed, if the NP is [−animate] it must occur in the immediately preverbal position.[330]

(165) NP_1 = {[−definite], [+animate]}:
      a.   Bir kadın biz-e   doğru koş-uyor.
            one woman we-DAT toward run-PROG

468   Rosanna Sornicola

b.   Biz-e   doğru   bir   kadın   koş-uyor.
we-DAT   toward   one   woman   run-PROG
'A woman is running toward us.'

(166)   NP$_1$ = {[−definite], [−animate]}:
a.   Akşamları   burada   bir   rüzgar   es-iyor.
evenings   here   one   wind   blow-PROG
'A wind blows here in the evenings'
b.   *Bir   rüzgar   akşamları   burada   es-iyor.
one   wind   evenings   here   blow-PROG[331]

These examples contain an intransitive verb. The situation with respect to sentences containing a transitive verb presents further properties to be taken into account. In general, it is worth noting that independently of the grammatical function (S vs. O) the immediately preverbal position, which in the basic order is occupied by O, is determined by the distribution of definiteness and animacy features. If the S of the transitive structure is {[−definite], [−animate]} and O is {[+definite], [+animate]}, it is S which must be located in the immediately preverbal position:

(167) a.   Ali-yi   ev-de   bir   sürpriz   bekli-yor.
Ali-ACC   house-LOC   one   surprise   wait.for-PROG
'A surprise is waiting for Ali at home.'
b.   *Bir sürpriz Aliyi evde bekliyor.[332]

Erguvanli observes that if both S and O have the feature [+definite], the unmarked order will be OSV, while SOV will acceptable as the marked order. The two sentences:

(168) a.   Ali-yi   ev-de   bu   sürpriz   bekli-yor-du.
Ali-ACC   house-LOC   this   surprise   wait.for-PROG-PST
b.   Bu sürpriz Aliyi evde bekliyordu.
'This surprise was waiting for Ali at home.'

do not have the same pragmatic presuppositions; the second sentence is highly marked and presupposes that some surprise is awaiting Ali at home, while such a presupposition is not attributable to the first sentence, which is pragmatically neutral (Erguvanli 1984: 29).[333]

Of particular importance are constructions in which O is nonreferential or is indefinite nonspecific: in such conditions O has no case marker and forms a single unit with the verb (the situation is similar to that in Mongolian and other Uralo-Altaic languages).[334] However, it then becomes the primary candidate for occupation of the immediately preverbal position:

(169)   O = [−referential]
  a.   Murat kitap ok-uyor.
       Murat book   read-PROG
       'Murat is reading a book.' (= 'Murat is book-reading.')
  b.   *Kitap Murat okuyor.

       O = {[−definite], [−specific]})
  c.   Murat bir   kitap ok-uyor.
       Murat one   book   read-PROG
       'Murat is reading a book.'
  d.   *Bir kitap Murat okuyor.

On the other hand, subject NPs which are [+definite] or indefinite specific may vary their position.[335]

   In Hungarian, in unmarked sentences if O is {[+definite], [+referential]}, its preferred position is after V. If O is {[−definite], [−referential]}, it normally occurs in front of V:

(170)   ˈJános ˈkönyv-et olvasott   a   ˈszobá-ban.
          John   book-ACC   read-PST.3SG ART   room-INESS
          'John read a book/books in the room.'[336]

Behrens observes that this linear property does not apply in sentences with focalization and adds that if O is [−definite] but [+referential], both the preverbal and the postverbal positions are accessible.[337] The positional restraints imposed by the features of definiteness and referentiality also apply to the function S. If S is [+definite] it occurs in postverbal position; if S is {[−definite], [−referential]}, it occurs in preverbal position.

   The interrelationship of the various functions (levels) is further complicated in marked sentences. As far as O is concerned, if the constituent in FOCUS is other than O, the postverbal position of O (independently, it would seem, of the semantic features it bears) is strongly preferred. As far as S is concerned, if the sentence has a constituent other than S in FOCUS, S is postposed to V. The position of both O and S in marked sentences is, therefore, constrained not only by the parameters of definiteness and referentiality, but also by the constituent which bears the function FOCUS (on the whole question, cf. Behrens 1989: 142).

   Behrens observes that if the sentence has no constituent other than O in FOCUS, O with the feature [+definite] may occur either before or after V.[338] It must be noted, however, that SOV with O [+definite] is rarely used and always has a contrastive function.[339] On the other hand, in sentences which have a constituent in FOCUS, as well as in imperatives and in sentences with a negation, the postverbal position of O is strongly preferred.[340]

470   Rosanna Sornicola

It is obvious that this situation differs from that in Turkish in various aspects: (a) although the features of definiteness and referentiality are as important for WO in Hungarian as they are in Turkish, Hungarian does not have the complex hierarchy that these features engender in Turkish, with respect to GFs, for allocation in position P_k or other crucial positions; (b) animacy does not seem to have an effect on WO.

5.2.   Textual relations

In Section 1.8 it was said that the basic model chosen for this study of WO is centered on the sentence (for an explanation of this choice, cf. Section 1.8). We shall now discuss some typological properties which require for their examination a model that takes into account the wider domain constituted by the text, in other words, the effects of context on WO. The discussion will center on the effects of the features GIVEN and NEW as well as certain particular effects induced by the type of textual development.

5.2.1.   The effect of the features GIVEN and NEW on word order

The effect of these features, which are typically anchored to the context, has been widely studied for some time, especially with respect to the Slavonic, Romance, and Germanic languages.[341] As far as these languages are concerned, it may be said that the constituents which carry the feature GIVEN in unmarked sentences tend to be located in the first position in the sentence, while those with the feature NEW tend to be located in the final position. The reverse order NEW . . . GIVEN is connected with marked structures in which the NEW constituent coincides with a marked FOCUS. Furthermore, it has been shown in various recent studies, carried out using different techniques on languages such as Russian and Italian, that the effect of the factors GIVEN and NEW on WO is relatively minimal (cf. Holden and Krupp [1987] for Russian, Sornicola [1994] for Italian).

However, it would be wrong to think that the tendency for "GIVEN first, NEW later" is a pragmatic universal. As we shall see later, the available empirical data show that the properties of the basic order of a language interact in a fixed way with the possible distribution of GIVEN and NEW.

In fact, it should be pointed out straightaway that all the languages in which the above mentioned tendency is found have basic SVO order in one way or another. The Finnish data (cf. Fromm 1982: 142; Tarvainen 1985; Hakulinen 1961), and Estonian data (cf. Tauli 1983: 24–25) are consistent with this picture. On the other hand, GIVEN and NEW have quite different distributional properties in SOV and VSO languages.

5.2.2.   The distribution of GIVEN and NEW in SVO languages

The distribution of the [+GIVEN] or [+NEW] constituents is correlated with topological properties which concern the position of the constituents in the sentence space and the position of the sentence in the context. It seems particularly useful, in view of a typological examination, to appeal to an integrated consideration of the informative and topological aspects employing the concepts of "orientation of the sentence towards the left context" and "orientation of the sentence to the right context."[342] These concepts may be defined in terms of rules which have an iconic nature: the constituents which link, or rather orient, the sentence to the left context (typically GIVEN constituents) tend to be located on the left side of the sentence, while the elements which link, or rather orient, the sentence to the right context (typically NEW constituents) tend to be located on the right side of the sentence. In Russian articulation of the sentence according to left and/or right orientation reaches a higher level of systematicity than in other SVO languages. The fact that WO in Russian and other Slavonic languages shows greater sensitivity to contextual factors compared with Germanic and Romance languages has been recognized for some time in the Praguean literature. It is merely another aspect of the greater flexibility of WO in the Slavonic languages than in the other languages in question, which has been discussed in Section 2. In the characteristic articulation in which the [+GIVEN] constituent occurs in P_1 and the [+NEW] constituent in P_n, the first establishes a link with the preceding (left) context, while the second establishes a link with the following (right) context. In sentences (27) and (30), repeated here as (171) and (172), both with OVS′ pattern, O is [+GIVEN] and signals an orientation towards the left context, while S is [+NEW] and signals an orientation towards the right context:

(171)   ['Why are you making coffee today?']

Kofe   prosit   MÁMOČKA.
coffee(ACC) ask:PRS.3SG mother.NOM.SG
'It's mother who wants it. [She can't do without coffee in the morning.]'

(172)   ['So, you have allocated the roles? Yes.']

Juru   čitaet   SERËŽA.
Jura.ACC   read:PRS.3SG Serëža.NOM
'It's Serëža who will play Jura. [Volodja has rather too soft a voice.]'

It is interesting that in orders in which the [+NEW] constituent does not occur in P_n, but in an earlier position, this constituent does not link the sentence with the right context. Consider the following example, reported by Fougeron (1989: 298), as a variant of (172) with an OS′V pattern:

472   Rosanna Sornicola

(173)   ['So, you've allocated the roles? Yes.']
             Juru SERËŽA čitaet.
             'It's Serëža who will play Jura.'

According to Fougeron (1989: 299–300) the OVS′ type is used in a textually dif-
ferent way from the OS′V type: the first occurs when the following context brings
an explanation of the choice made by the speaker, while the second is preferred in
absence of such a condition.[343]

    The iconic topological property which emerges for Russian on the basis of Fou-
geron's data, and which may be represented as in Figures 21–22, offers a useful point
of reference for identifying the specific characteristics of other SVO languages.[344]
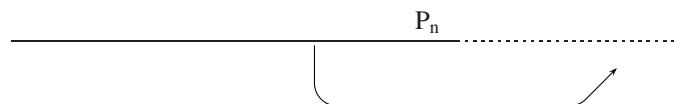
Figure 21. Iconic topological property of the OVS′ pattern in Russian
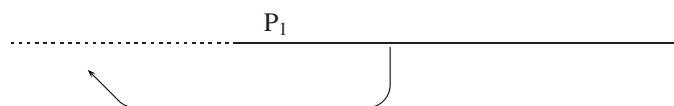
Figure 22. Iconic topological property of the OS′V pattern in Russian

    In fact, the topological properties of the distribution of GIVEN/NEW in SVO lan-
guages seem to divide them into two groups which coincide with the basic division
between languages with high WO flexibility and those with low WO flexibility. This
difference emerges only when considering the interaction of the GIVEN/NEW pa-
rameters with GFs. It is this interaction alone in fact that allows determination of a
clearly defined typology irrespective of general pragmatic tendencies.

    When GIVEN = S, and NEW = O or the entire VP, and there are no marked
focalization conditions, which are always possible as independent conditions with
respect to the distribution of GIVEN/NEW,[345] all the SVO languages conform to the
principle "GIVEN first, NEW later." Differences between them emerge only when

GIVEN and NEW do not coincide with "canonical" GFs (on the natural affinities between GIVEN and S and between NEW and O or VP, cf. Comrie 1981b). It is exactly here that the specific structural conditions of individual languages come into play.

When GIVEN does not coincide with S, but with another GF (for example, keeping within the limits of what is currently under examination, with O), Russian, Finnish, and Estonian allow such a constituent to occur in P_1, this being a proper sentential position, while the Romance and Germanic languages allow O [+GIVEN] to occur in P_1, which is understood as an extrasentential position; in other words, these are dislocation phenomena, varyingly examined for the Romance and Germanic languages (cf. Sections 2.4.4.1 and 2.4.5). Therefore, all SVO languages share the iconic property of the occurrence at the left edge of the constituent which marks orientation of the sentence towards the left, but with the important differences relating to the structural properties of position P_1 we have mentioned.

In the Romance and Germanic languages, on the other hand, when NEW coincides with S, the basic SVO order may be retained with S receiving the nuclear stress, while it has been seen in Sections 2.4.4.1 and 2.4.5 that the possibilities of movement into P_n, or at least rightwards, are severely restricted. Revealing, in this regard, is the comparison between Russian OVS′ structures (*Kofe prosit MAMOČKA* and *Juru čitaet SERËŽA*) and the possible translations into a Romance language, which, as was said in Section 2.4.4.1, would show one of the following alternatives:

(a) S′VO structure;
(b) Cleft structure with postcopular S in FOCUS;
(c) Structure with O moved into an extrasentential P_1 position and S in P_n.

Of course, as far as the structural configuration of the sentence itself is concerned, (c) is not isomorphic with the structures of Russian sentences (171)–(172). Furthermore, as far as textual and topological properties are concerned, (c) does not have regularly the same iconic characteristic of orienting the sentence towards the right context, represented schematically in Figure 21. In fact, in the Romance languages S [+NEW] (or any other NEW constituent) in P_n can only be part of a structure with the pragmatic property of representing the more dynamic informative value at the extreme right-hand edge of the sentence. This is a very general property, characteristic of SVO and, to a certain extent, VSO languages. It is a necessary topological pre-condition in order for a constituent to establish rightward orientation of the sentence, as was also seen in the case of Russian, but it is not a sufficient one. The difference between Romance languages and Russian lies exactly here: in the Romance languages, the NEW constituent in P_n is only optionally an element for establishing a textual relation with the following context; furthermore, position P_n does not alternate with some of the preceding positions which signal lack of rightward orientation of the sentence. In Russian, instead, P_n occupied by a NEW constituent systematically has

474   Rosanna Sornicola

the effect of rightward orientation and is in consistent syntagmatic opposition to a preceding NEW position which does not orient rightwards.

These differences in the pragmatic typology seem to be induced by at least two factors. The first is structural in nature: having greater WO freedom and, hence, a wider range of available positions for the various constituents bearing PFs and informative features, Russian may structure textual properties according to a more articulated grid, while in SVO languages with less free orders, the development of textual properties is codified to a lesser extent by WO. This is more evident on the right side of the sentence than on the left, where the effect of universal pragmatic factors relative to GIVEN and to TOPIC is very strong.

The second factor is of a typically pragmatic nature: it has to do with the different ways GIVEN material is treated in Russian and in the Romance languages, as can be seen from the French and Italian translations of sentence (171):

(174)   C' est maman   qui   en   veut.
         that is   mother.SG REL:SBJ PRTV.CLT want:PRS.3SG
         'It is MOTHER who wants it.'

(175)   Lo vuole   la   mamma.
         it   want:PRS.3SG the mother.SG
         'MOTHER wants it.'

As can be seen, the difference from Russian lies in pronominalization of the GIVEN constituent, which is the more typical structure.[346] In other words, if the constituent with the GF O is [+GIVEN] in the Romance languages (as well as in the Germanic languages), it tends to be pronominalized, while in Russian it may keep its full lexical representation when it is confined to the left side of the sentence.

5.2.3.   The distribution of GIVEN and NEW in VSO languages

In VSO languages, the distribution of GIVEN and NEW conforms to the same general principles of "GIVEN first, NEW later" or of "NEW first, GIVEN later," but with further complications arising from their specific structural characteristics.

In unmarked V-initial structures, if S = [+GIVEN] and V, O = [+NEW], the splitting of the VP by S gives rise to an irregular informative progression NEW – GIVEN – NEW.

On the other hand, S/O initial structures in which – as was seen in Section 2.4.6 –, the initial constituents may be GIVEN or, more often, NEW, may give rise to one or other of the above-mentioned general principles.

The real difference between these and other types of languages, in any case, seems to lie in the frequency with which structures with FOCUS in P_1 give rise to NEW constituents in P_1, in other words, the frequency of informative distributions with "inverse" or artificial order.

Interaction of syntactic and pragmatic factors   475

5.2.4.   The distribution of GIVEN and NEW in SOV languages

SOV languages show a different tendency in the information structure. In these languages it is the postverbal field that is characteristically occupied by GIVEN material (cf. Erguvanli 1984: 44 ff.). If GIVEN coincides with S, there is no alteration to the principle "GIVEN first, NEW later." However, if GIVEN coincides with another GF, it occurs in the postverbal space.[347]

This property, which may be expressed by the principle "NEW first, GIVEN later," has topological and structural implications worth noting, because the GIVEN material does not belong in the strict sense to the sentence domain, which in these languages is delimited by V, but to an adjacent extrasentential space (position P_n = TAIL). It is clear that this situation is very distinct from that found in SVO languages and that in VSO languages.

The picture is, however, complicated by the fact that this postverbal space may be occupied not only by GIVEN material, but more generally by backgrounded material (cf. Erguvanli 1984: ch. 2).

The typology of the spatial distribution of GIVEN/NEW in SOV languages has an interesting symmetry compared to the corresponding typology of SVO languages with a low degree of WO freedom. As was said in Section 2.4.4, in the latter languages, when the GIVEN element does not coincide with S, it occurs in an extrasentential space at the leftmost edge of the sentence, according to the schema in Figure 23.

Extrasentential space	Sentence space
[+GIVEN] X	

Figure 23. Placement of the GIVEN element in SVO languages when GIVEN ≠ S

On the other hand, in these languages if GIVEN = S, the representation shown in Figure 24 holds true.

Extrasentential space	Sentence space		
	P_1	. . .	P_n
[+GIVEN] X			

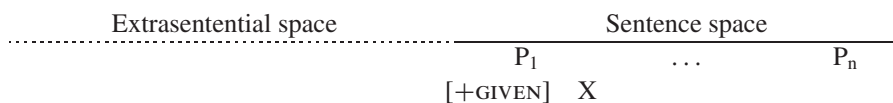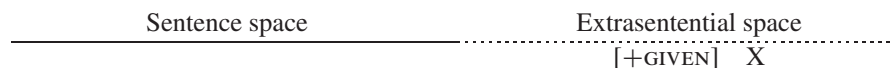*Figure 24.* Placement of the GIVEN element in SVO languages when GIVEN = S

In SOV languages, instead, when the GIVEN element does not coincide with S, it occurs in an extrasentential space at the rightmost edge of the sentence, according to the schema in Figure 25.

476   Rosanna Sornicola

Sentence space	Extrasentential space
	[+GIVEN] X

Figure 25. Placement of the GIVEN element in SOV languages when GIVEN $\neq$ S

On the other hand, if GIVEN = S, it occurs in P_1 in the sentential space (see Figure 26).
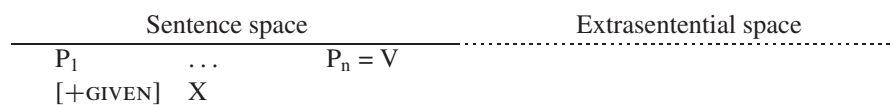
Sentence space	Extrasentential space
P_1 . . . $P_n = V$	
[+GIVEN] X	

Figure 26. Placement of the GIVEN element in SOV languages when GIVEN = S

This symmetry seems connected to the structural characteristics of the two types of language: in both SVO and SOV languages, when the constituent with the feature [+GIVEN] does not have the GF S, it is located on the opposite side of the canonical position for the constituent governed by V.

In short, there are two principles for all of the types examined, "GIVEN first, NEW later" or "NEW first, GIVEN later"; what varies and what may constitute the basis of a textual typology are the various structural configurations with which the two mentioned principles interact, giving rise to different topological properties.

6.   Conclusions

6.1.   Universal tendencies and the typology of word order

6.1.1.   The "written" vs. "spoken" parameter

One of the fundamental problems in the typological determination of linguistic properties is the fact that the results of analyses of one (or more) languages may differ significantly depending on whether the data are drawn from "written" or "spoken" sources. This may have adverse effects not only on the description (for example, on the assignment of a language to a certain type, and sometimes even on the determination of a "type"), but also on the explanation of the factors which favor a given typological regularity. As far as the latter point is concerned, if a particular characteristic is found in the written rather than in the spoken language, there is good reason to suspect that it has come about via a sociolinguistic process "from above," for example, through the influence of prestigious models. Such models may significantly bias the way linguistic regularities are established, especially in highly structured

cultural situations, such as those found in Europe. One need only recall the influence of Latin and/or Greek on the formation of many European literary languages between the Middle Ages and the present day.

In this work, the problem concerning the potential influence of prestigious models on the formation of a literary language has been considered with respect to German in particular (cf. Section 2.4.5.5.6). The German situation is interesting, not only because it shows that the spoken registers of a language may exhibit dominant orders that differ from those of the written language, but also because it confirms how pervasive SVO patterns are in the spoken registers of many modern Indo-European languages. This pervasiveness may be explained with various reasons:

(a) SVO is by far the dominant order in languages which are basically SVO whether they have flexible or rigid constituent order;

(b) It is a "natural" mechanism, which cuts across typological classifications. It may be seen that even languages which are structurally prevented from exploiting SVO order for "topological" reasons, such as Turkish, allow such structures, albeit marginally (cf. Sections 2.2.1 and 5.1);

(c) Distribution of the SVO pattern in neutral sentences in many SVO languages is related to psycholinguistic strategies which characterize information processing in speech planning (cf. Sornicola 1981; Levelt 1989; Miller and Weinert 1998; Miller and Fernandez, this volume). Such psycholinguistic and pragmatic influences should not, however, be overestimated. They act *within an established type* and it is not clear whether they are able in themselves to bring about a change in the basic typological properties of a language. This hypothesis has implications for the principles that model the historical development of WO structures. In fact, if a language has abandoned a characteristic inherent order in developing SVO as the dominant or basic pattern, this may instead be the result of external factors of borrowing in historically determined situations of social and cultural contact. An example of this is the shifting of western Ugro-Finnic languages from SOV to SVO (see Comrie 1981a: 92–93, 121 ff.), but confirmation may also come from other linguistic groups.[348]

As far as possible, then, in typological analysis *it is necessary to compare phenomena of the same register*. This procedure, which the present work has attempted to follow, reflects a *microscopic* criterion of examination of typological conditions, according to what was put forward in Section 1.

6.1.2. The historical conditions

The "written" vs. "spoken" parameter does not seem to be the only one that should be taken into account in a typological analysis. It is clear from what has just been said that the very notion of "typological characteristic" raises several apparently crucial problems. In the first place, its definition cannot be taken for granted. Furthermore,

478   Rosanna Sornicola

it is not clear to what extent it reflects what is "natural" and what is induced from external factors. In examining the data the typologist's view cannot be restricted by overlooking these parameters – on the contrary – the data must be considered in all their complexity and in the details of their origin, that is, in their *historical dimension*. In a typological work, it is in fact only the historical dimension which allows a distinction to be made between what is a genuine universal tendency and what is a particular property of an individual language.

For this reason, another criterion adopted in the present work has been consideration – wherever possible – of the historical conditions with which a given WO pattern is associated. The term "historical" is understood here in the two senses of "linked to a particular condition" and "diachronic."

The first meaning concerns, for example, the influence already noted of prestigious languages, such as Latin and/or Greek, on WO patterns in many European languages. As far as the second is concerned, wherever the existence of reliable studies allows it, examination of synchronic structural properties has been integrated with consideration of the data from older diachronic phases in order to better understand the continuity and discontinuity of particular patterns. For example, the traces of "non-verb-second" properties discussed with respect to some of the Germanic languages are interesting because they seem to confirm the hypothesis that the "V-2" character may be relatively recent and/or an "artificial" development arising from literary models.

More generally, examination of specific historical conditions allows apparently counterintuitive typological considerations to be made. As far as WO in the Indo-European languages of the corpus is concerned, such an examination reveals a considerable affinity in the patterns of the spoken languages (with the only exception of Celtic) to be hypothesized. Indeed, it may be claimed that in these languages, there is a greater divergence of WO patterns in the written than in the spoken registers. This suggests that *such differences are historico-cultural rather than natural*. Regarding this, it is important to stress the impact of the various rhetorical traditions in establishing WO patterns (cf. Scaglione 1972). For example, the different cultural and social value – between the Middle Ages and the present day – ascribed to the practice of *constructio* (parsing) adopted at the highest levels in teaching Latin and, later, modern European languages, has been crucial (for a brief overview of these problems, cf. Sornicola 1995a, 1998 particularly with regard to the influence of the *constructio* on the establishment of WO patterns in English).

6.1.3. Structural conditions

It is now possible to make some tentative conclusions concerning the importance of structural conditions on both neutral and non-neutral (marked) WO patterns. The typological examination carried out does in fact allow some general principles to be

identified, and, at the same time, some current assumptions to be critically reconsidered.

First of all, it is necessary to clarify that here by "structural conditions" the influence is meant of what have been defined in this work as "topological" or "geometrical properties," that is, the properties of organization or structuring of the (neutral or marked) sentence space. It is assumed here that these depend on the relative position of the basic constituents in neutral canonical sentences. Such influence cannot be taken for granted: why, in fact, should it be presumed that the organizational principles of neutral sentences have repercussions on WO patterns in non-neutral sentences as well? It is hypothesized here that the relationship between the principles operating in neutral patterns and the potential occurrence of constituents in certain positions in the non-neutral sentences has to do with two different general strategies which characterize many languages of the corpus. Such strategies have been defined, respectively, as "nondifferential" and "differential" (cf. Section 6.2.2 below). However, before discussing these two strategies, it is appropriate to try and develop the concept of "sentence space" and to determine the relationship between this and the concept of "neutral canonical configuration of constituents." In the next section, some criteria for these two objectives will be made explicit, which should be taken as a heuristic model for the understanding of the typological characteristics of WO.

6.1.4.   Criteria for determination of the sentence space

The concept of "sentence space" is defined in terms of the relationship between constituents, constituent position, and the domain that contains them. It seems relevant, in particular, to consider:

(a) the relationship between constituent and domain;

(b) the relationship between position and domain.

Relationships (a) and (b) are closely connected, since the place that a constituent occupies marks out a position (on some of the difficulties with this, see Section 1.6 and in particular Section 1.6.3).

The concepts "actual position," "possible position," and "potential or virtual position" will now be defined.

> *Actual position*
> This is the position in which a constituent in the *actual* or *real pattern* of a given structure occurs. For example, in the SVO pattern the constituent with the GF S, the constituent V and the constituent with the GF O each occupy actual positions. Similarly, in the OVS pattern each of the three constituents occupies an actual position. This means that in order to have an actual position, a corresponding actual pattern must first of all be defined.

480   Rosanna Sornicola

Possible position

This is the position accessible to the constituents of a sentence structure on the basis of the *possible patterns of order* in a given language. The set of possible positions includes the set of actual positions; in other words, if a position is actual then it is also possible, but not vice versa. In SVO languages with rigid order, any preverbal position (including P_1 and P_k) is accessible to the constituent with the GF O only in marked sentences.

Potential or virtual position

This is the position *potentially available* to a constituent, independently of any specific typological condition. For example, in Turkish the immediately postverbal position is a potential or virtual position with respect to the proper sentence domain. However, as was seen in Section 2.2.1, it can only be filled by an extrasentential constituent. Thus, it can be said that in the proper sentence domain, it is neither a possible nor an actual position because a specific typological condition of Turkish prevents it. It is, in fact, a *potential or virtual position*.

Corresponding to the three notions of *actual position*, *possible position*, and *potential or virtual position* are the three notions of "actual space," "possible space," and "potential or virtual space," defined as follows:

Actual space

This is the space which contains all and only the *actual positions*.

Possible space

This is the space which contains all and only the *possible positions*. Possible space includes actual space.

Potential or virtual space

This is the space which contains all the *potential or virtual positions*. This space includes both actual and possible space.

Note that the concepts of actual space and possible space specifically refer to the *proper sentence domain* and not to the extrasentential domain. The proper domain is defined by V and its arguments, while the dislocated constituents or afterthoughts determine the extrasentential space. The latter may, in its turn, be either actual extrasentential space, possible extrasentential space, or potential or virtual extrasentential space, but this tripartition is less useful and less interesting than it is for the proper sentence domain, since the extrasentential spaces, whether leftmost (that is, positioned beyond the left-hand edge of S) or rightmost (that is, positioned beyond the right-hand edge of S), present weakly defined structural properties; furthermore, the typological conditions governing them are less conspicuous than those which apply within the sentence domain. It may be claimed that actual space, possible space, and potential or virtual space are, in effect, identical in the extrasentential spaces.
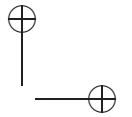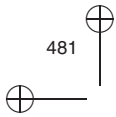
The model which distinguishes between types of positions and types of sentence spaces could have an independent motivation on psychological grounds. Interesting in this respect is the effect that neutral sentences may have. For example, it could be possible to envisage that the structure of the space in the neutral sentence establishes the sentence's external and internal boundaries which the speaker perceives and utilizes more or less actively in production. It is plausible that these boundaries are also valid in marked sentences. In any case, such a hypothesis is interesting only up to a certain point here since the model may and must be defined according to abstract structural properties.

6.1.5.   Some principles of word order organization

On the basis of what has been said so far, some principles which regulate the various organizations of WO emerging from the languages of the corpus may be formulated.

> *Principle I*
> In SVO and SOV languages, the constituent V has the topological property of organizing the sentence space in different modes, each characteristic of a type of structure. It determines (delimits) not only the actual sentence space, but also the possible space, such that:
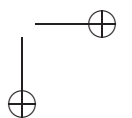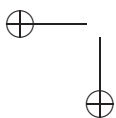
> *Principle II*
> If no constituent may occur in a given virtual position defined with respect to V in the neutral canonical structure(s) of a language, the space identified by that virtual position is not a possible space for any constituent of the proper sentence domain, not even in the marked structure(s) of that language.

In SVO languages, V marks out a space to its right as either actual space or possible space. In such languages, the proper sentence domain is in the majority of cases closed by referential constituents.[349] This gives rise, therefore, to the representation in Figure 27.

| Extrasentential domain | S | V | O | [...] | Extrasentential domain |

Figure 27. Proper sentence domain and extra-sentential domain in SVO languages

As, in fact, was seen in Section 2.4, in all the SVO languages of the corpus any pre- or postverbal position is accessible to the constituent with marked FOCUS; that is, this constituent may occupy the pre- or postverbal space. Generally speaking, in all these languages the configurations with marked FOCUS are distinguished by heavy stress falling on the constituent bearing FOCUS.[350]

482   Rosanna Sornicola

Note that this very general principle applies in languages with morphological case as well as in those without, even though it was noted in Section 2.4.4 that in the latter, movement of O to the preverbal space is not possible in unmarked sentences. This difference, however, seems of secondary importance typologically. Another noteworthy secondary property that varies across SVO languages seems to be the possibility in unmarked sentences for a constituent bearing GF to be moved to the leftmost vs. rightmost sentential position, with the function of textual cohesion towards the left or right contexts, respectively. Again, this is a property differentiating SVO languages with case marking and highly flexible WO, like Russian, from SVO languages without case marking and with weak flexibility of WO (see Sections 2.4.3.2, 2.4.4.1, and 2.4.4.2). It may, therefore, be claimed that properties of a "geometric" nature have greater influence on the set of WO types than do those of a morphological-relational nature, such as the presence of case and a relatively greater flexibility of WO.

It is clear from principles I and II that the constituent V is crucial in both SOV and SVO languages. In the latter in particular, V operates in the sentence as a discriminating element with respect to the GFs. In such languages, in fact, the more general property of asymmetry between S and O with respect to dependency structure is realized by assigning to both S and O a differential position with respect to V. SVO languages have, furthermore, greater possibilities of constituent movement within the proper sentence domain. In particular, it is clear from Figure 27 that:

> *Principle III*
> In SVO languages, all the preverbal space may be exploited for topicalization and focalization.
>
> *Principle IV*
> In SVO languages, the postverbal space is an integral part of the proper sentence domain and may be exploited for both unmarked and marked focalization.

As far as SOV languages are concerned, as has already been observed, V marks the right-hand edge of the proper sentence domain. In fact, it closes the virtual space to its right, which is not a possible space for any constituent in the proper sentence domain. This may be represented as in Figure 28.

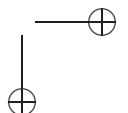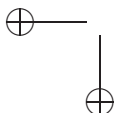| SOV	Extrasentential domain

Figure 28. Proper sentence domain and extra-sentential domain in SOV languages

This property is characteristic of Basque and Turkish and, in general, is widespread throughout many other SOV languages (cf. Section 6.2.4 below).

A model constituted by the following properties may therefore be arrived at.

In the SOV languages of the corpus:

Principle V
The proper sentence domain is weakly delimited on the left by a referential constituent; it is strongly delimited on the right by V;[351]

Principle VI
There is only one extrasentential domain and that is situated immediately after V;

Principle VII
Possibilities of movement within the proper sentence domain are more limited than in SVO languages;

Principle VIII
The preverbal space is structured such that the topicalization and the focalization zones are separated;

Principle IX
The postverbal space is not available for focalization.

Principles V, VI, and IX are also shared by other languages with the same basic or dominant order, while principles VII and VIII are not widely distributed at all among SOV languages (cf. Section 6.2.4 below).

An interesting structural property, more generally characteristic of SOV languages, is that the constituents which bear the GFs S and O are situated in the same part of the space of the proper sentence domain with respect to V, that is, to the left of this constituent. This topological property marks a significant distinction between SOV languages and SVO languages.

It may be hypothesized that in Basque and Turkish it acts in conjunction with the property by which V closes the sentence domain, resulting in the absence of a "differential" strategy for the linear coding of PFs. This hypothesis, although attractive, must be rejected.

Revealing in this respect is comparison with the VSO languages of the corpus, which also have the topological property that the GFs S and O are found in the same part of the sentence space with respect to V, in this case to the right of such constituent. This property may, in fact, be considered "symmetrical" to that of the SOV languages mentioned above. If the hypothesis just advanced had any basis, it would be expected that in VSO languages, the properties governing exploitation of the sentence space would be similar to those of SOV languages. In fact, the only obvious similarity between these language types is that in both, the possibilities of constituent movement under pragmatic conditions are "topologically" more limited than in SVO languages. In VSO languages, movement of S to the canonical position of O is attested as a marginal possibility in some dialects, while the only generally

484   Rosanna Sornicola

$$P_1 \quad \dots \quad P_n$$
$$\begin{Bmatrix} V \\ S \\ O \end{Bmatrix}$$

Figure 29. Relation between position P_1 and GFs in the possible space of VS0 languages (in marked and unmarked sentences)

possible movement of S is to the preverbal position. On the other hand, VSO languages present some interesting structural affinities with SVO languages. Summarizing what has emerged in Section 2.4.6, the following model may be formulated:

> *Principle X*
> The VSO languages of the corpus show different positional properties for codification of PFs in neutral and in marked sentences;
>
> *Principle XI*
> They do not have a unique FOCUS position for neutral sentences and marked sentences;
>
> *Principle XII*
> In most varieties of these languages, in marked sentences, P_1 is the only position for FOCUS.

Furthermore, as was stated in Section 2.4.6.4, in VSO languages, P_1 is a position in which either S or O may occur with the PF of TOPIC in neutral sentences. As a consequence, there is reason to believe that in the languages in question P_1 is a *multifunctional pragmatic position*. In VSO languages, therefore, as in SVO languages, GF and PF movement is possible under pragmatic conditions. On the other hand, these languages have a *characteristic conspicuous position*, valid for both TOPIC and marked FOCUS (cf. Section 2.4.6.4). There would, however, be no justification in assimilating this latter property to that of PF codification, operating in SOV languages such as Basque and Turkish. In fact, the "movement" of constituents to P_1 under pragmatic conditions is a widespread phenomenon among the world's languages. In other words, P_1 is a *generally conspicuous position in many languages*, presumably for reasons related to the serial organization of the utterance.

On the basis of what has been said, the representations in Figures 29 and 30 apply in VSO languages.

Figure 29 represents the fact that P_1 is a possible position for every basic constituent in both neutral and marked structures, while Figure 30 represents the fact that such a position may be occupied by the PF TOPIC or the PF FOCUS.[352]
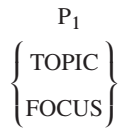
$$P_1$$
$$\left\{ \begin{array}{c} \text{TOPIC} \\ \text{FOCUS} \end{array} \right\}$$

Figure 30. Relation between position P_1 and PFs in the possible space of VS0 languages (in marked and unmarked sentences)

| Extrasentential domain | V | S | O | Extrasentential domain |

Figure 31. Proper sentence domain and extra-sentential domain in VSO languages (Model 1)

$$P_1 \quad \dots \quad \dots \quad P_n$$
$$\left\{ \begin{array}{c} V \\ S \\ O \end{array} \right\}$$

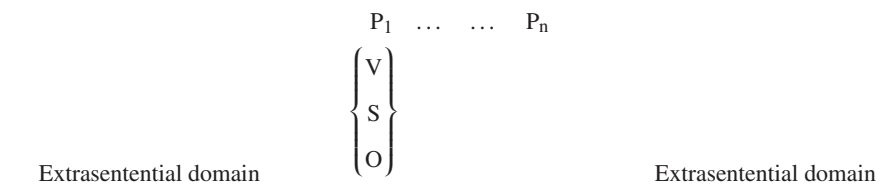| Extrasentential domain | | Extrasentential domain |

Figure 32. Proper sentence domain and extra-sentential domain in VSO languages (Model 2)

The representation of possible space is more difficult. If one accepts the idea that preverbal S/O occupy an extrasentential position, it may be claimed that V signals the left-hand edge of the sentence domain, according to the representation in Figure 31.

An alternative representation would be that in Figure 32, which shows that position P_1 is not typically reserved for V but is "flexible" in that it may be exploited by any one of the basic constituents under different pragmatic conditions.

This representation would, in fact, appear to be preferable, since in Section 3.3 it was claimed that in the VSO languages of the sample *labile rules* of order may be postulated for both intransitive and transitive structures. It suggests that V is not the only constituent which delimits the leftmost edge of the proper sentence domain, but rather that this boundary is signaled by a variable. This would give rise to a further property differentiating VSO and SOV languages. In the latter, the relationship between position and constituent is rigidly defined, while in VSO languages it is not.

The topological properties which have emerged in this work lead one to think that WO phenomena may be considered as more than just the realization of grammatical relations. The topological influences on WO are presumably stronger than some of the regularities which have been traditionally postulated, such as the distinction between languages with pragmatic WO and those with grammatical WO or the in-

486   Rosanna Sornicola

cidence of "morphological case." On the other hand, what has emerged in this work is that there is a relationship between the topological principles mentioned and other means of PF codification, such as the use of clefting processes (as in French and the Celtic languages) and intonation (as in English). These are, as was stated in Section 2.4.4, demarcation phenomena typical of languages in which the linear axis is comparatively less exploited for the codification of PFs. Such phenomena appear to be language-specific.

Another result which has emerged from this work is that the topological properties relevant to bi-argumental structures are different from those which are relevant to monoargumental structures, but that they seem to be correlated in an interesting way (cf. Section 3.6.2).

6.2.   Word order and the asymmetry of pragmatic functions

In Section 2.1, it was noted that grammatical codification of TOPIC and FOCUS by means of WO reveals these functions to be typically asymmetrical. The first is in fact realized by a narrower range of possible positions than the second. This characterization concerns neutral rather than marked sentences. In fact, in all the languages of the corpus, the property of being "what is being spoken about" in neutral sentences is manifested by assigning TOPIC to P_1 or in any case to a preverbal position.[353] Furthermore, in all the languages of the corpus, TOPIC tends to coincide with the GF S in neutral sentences. Minor differences have emerged regarding the possible patterns in which TOPIC coincides with another grammatical function, as in Russian, for example (cf. Section 2.4.3.2). In fact, it has been seen that in this case, TOPIC in P_1 may be a constituent other than S.

The question regarding the linear properties of TOPIC and its association with a given GF in marked sentences is more controversial. As will be recalled, doubts have, in fact, been raised about whether one can always identify TOPIC in marked sentences, since "aboutness" may be – so to speak – "absorbed" by FOCUS (cf. Section 1.8.1). With the exception of Basque and Turkish, the range of possible positions for FOCUS is wider in marked sentences than in neutral sentences.

One may well ask whether TOPIC, like FOCUS, is definable in terms of a differential strategy (cf. Sections 6.1.3 and 6.2.2). It is plausible to claim that, among the world's languages, TOPIC is less characterized than FOCUS in terms of differential movement strategies, a conclusion which seems consistent with the more general asymmetry between the two PFs.

With respect to the codification of FOCUS in particular, the languages of the corpus may be grouped according to two types:

(a)   languages with no difference at all between neutral (unmarked) and marked sentences;

(b) languages in which PF codification is differentiated according to whether the sentence is neutral or marked.

Basque and Turkish belong to the first group, as do languages in which TOPIC is always in the same P_1 position and FOCUS is always in the same immediately preverbal P_k position. The Indo-European languages, as well as Finnish and Estonian, belong to the second group; in these languages, FOCUS has a wider range of possible positional codifications in marked sentences (cf. Figures 3–7 in Section 2.4.1).

Before examining this typological bifurcation more closely, it would be appropriate to look in greater depth at the relationship between position, constituency, GF, and PF, and in light of this, to reconsider the possible codifications of FOCUS which have emerged in the course of this study.

6.2.1.   The relationship between position, grammatical function, and pragmatic function

It should first of all be pointed out once again (cf. Section 1.6.2) that the concept of "movement" is not without theoretical problems[354] and that it should rather be used as a handy metaphor for heuristic purposes. Furthermore, the expression "function movement" must be understood as an abbreviation for a conceptually more complex expression. In fact, the primary manifestation of movement is always of constituents and only in second place may one speak of "GF or PF movement."

In the present work, a four-level model has been employed, in which a different representation corresponds to each level:

>*Representation by position*
>This is the set of n-ple $(P_1, P_2, \ldots, P_n)$, where each P_i is a position.

>*Representation by constituent*
>This is the set of n-ple $(X_1, X_2, \ldots, X_n)$, where each X_i is a constituent.

>*Representation by GF*
>This is the set of n-ple $(\phi_1, \phi_2, \ldots, \phi_n)$, where each ϕ_i is a GF.

>*Representation by PF*
>This is the set of n-ple $(\rho_1, \rho_2, \ldots, \rho_n)$, where each ρ_i is a PF.

In light of what has been said so far, it may be claimed that a direct relationship generally only arises between positions and constituents, and in addition between constituents and GFs on the one hand, and between constituents and PFs on the other. The representation of Figure 33 may therefore be arrived at.

This expresses the fact that the relationship between position and GF and between position and PF is mediated by the representation by constituent. However, this model may be modified on the basis of properties of individual languages. In
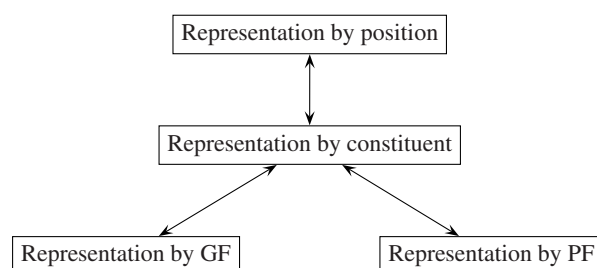
488   Rosanna Sornicola

Figure 33. Relationships between positions, constituents, GFs, and PFs

fact, the relationship between position and GF and between position and PF is differently determined according to the type of language. The same may be said of the relationship between GF and PF.

In theoretical and typological research, the nature of the relationship between GF and PF has been a controversial matter for a while.[355] In particular, the current debate may be summed up in the question of whether it is the GFs which determine PFs or vice versa. The view held in the present work is that any relationship between GF and PF may only be considered in relation to the level of positional codification of the two sets of functions, since, as was said in Section 1, it is assumed that they have independent abstract representations. For this reason, the conclusions arrived at here may be considered to be fundamentally different from those obtained in the framework of the d-configurational model.

In languages where PFs have a conspicuous canonical position, such as Basque and Turkish, there are several GF configurations but only one PF configuration. Variation in the basic order is found only with respect to the GFs (cf. Section 2.2.1). In other words, the GFs but not the PFs may be moved. Furthermore, the relationship between GF and PF is mutually independent and is not subject to any conditions: independently of whatever its GF may be, a constituent acquires a given PF by virtue of occupying a given position, which is unique.

On the other hand, in the SVO languages of the corpus, both GFs and PFs may be moved and both types of function are interdependent, at least as far as FOCUS is concerned. In fact, as will be mentioned in the next section, in order for a constituent to carry (marked) FOCUS, it must occur in a position which is not the canonical one for its GF. This means that a constituent may be positionally determined as FOCUS only after it has been assigned its GF. In neutral sentences, instead, although there is a more or less regular correspondence between GF and PF, it cannot be claimed that PFs are determined from GFs or vice versa.

Interaction of syntactic and pragmatic factors   489

6.2.2.   Codification of FOCUS by means of word order

We may now turn to an examination of the two different strategies for syntactically codifying FOCUS that have emerged from the languages of the corpus. In the first, FOCUS, whether unmarked or marked, is assigned to the immediately preverbal position P_k. We shall call this "Strategy 1."

The second strategy has two closely interrelated properties:

(a) FOCUS is associated with different structural positions in unmarked and marked sentences; in particular, the position of FOCUS tends to coincide with that of O in unmarked sentences;

(b) in marked sentences, FOCUS is not associated with a specific unique position, but with a set of characteristic positions which must be differentially defined as the set of positions that the constituents with the GFs S or O cannot occupy in unmarked sentences.

Regarding this, it should be noted that in marked sentences, both the functions S and O may coincide with FOCUS. In this strategy *it is not an absolute positional property, but a relative or differential property which characterizes FOCUS*, that is, it is *the difference between the canonical position that the constituent has in neutral sentences* (generally unique[356]) *and the position which it occupies in marked sentences* (a position which is not unique, but varies according to the marked pattern). We shall call this "Strategy 2."

The difference between the types of structure resulting from one or the other strategy may be formulated in terms of *absence of movement* vs. *movement of the PFs* (for further details on this, cf. Section 6.2.3 below). It is clear that Strategy 1 is *without PF movement,* that is, these functions *have fixed positions*. Strategy 2 is instead characterized by *PF movement along the linear axis*, that is, these functions *have moveable positions*. An important and consistent characteristic of the second strategy is that constituent movement to a noncanonical position resulting in its bearing the function of marked FOCUS, always co-occurs with the constituent in question receiving nuclear stress. It should be pointed out, however, that in the languages of the corpus with PF movement, principle XIII generally applies, according to which:

> *Principle XIII*
> Movement of a constituent from its canonical position to the anterior (preverbal) sentence space is preferred to movement to the postverbal position.

This principle has sometimes been characterized as "movement to P_1" and has been considered a sort of universal. But it is doubtful whether it is anything other than a fairly widespread tendency among the world's languages.

VSO languages, which behave only in part like SVO languages (cf. Section 3.1.2), require separate discussion. Although WO is used in the codification of PFs, it cannot

490   Rosanna Sornicola

be said that on its own it allows unmarked structures and marked structures to be differentiated. Context and suprasegmental features are instead used for this purpose (cf. Section 2.4.6.6).

The two strategies mentioned constitute a typological parameter which differentiates the SOV from the SVO languages. This typological difference seems of interest; discussion of matters arising from it will be developed in the following paragraphs. They concern the correlation between types of basic order and types of PF codification by WO, the nature of the relationship between the GF O and the PF FOCUS, the potential relationship between morphological type and type of PF codification and, finally, the areal distribution of types of PF codification by WO.

In SOV languages, the position P_k is conspicuous. In SVO languages, instead, the crucial WO property is not the position, but the "change" or difference in position of the constituents which carry GFs. In SVO languages, therefore, every noncanonical position may be said to be a conspicuous position. This property seems to be determined by a principle of differentiation; its importance in the functioning of the languages has already been expressed by the well-known structuralist principle which states that "dans la langue il n'y a que des differences."[357] This is the same principle which can be seen operating in prosodic structures. Phenomena of "intonation" and "accent" cannot, as is well known, be determined on the basis of the absolute or intrinsic characteristics of an element, but rather on the basis of the relative or differential value that they assume with respect to the corresponding value of other elements in the same structure. Similarly, the pragmatic functions of WO may also be realized by characteristics of relative prominence; they may, that is, depend on comparison with another term. As will be seen, this type of codification is more widespread than codification by means of absolute position.

The differential character of WO can be interpreted in two ways. Firstly, in the wider and more common meaning, "order" is in itself a relative concept independent of the type of language (cf. Section 1.6.1). The second meaning is more specific and may be associated with the typological property which has now emerged and which in short may be reformulated thus:

Principle XIV (principle of "differentiation of the codification of FO-CUS")

In languages which conform to Strategy 2, the PF FOCUS in marked sentences may be represented positionally as the set of pairs:

$X\phi$ (P_i, P_j)

such that for each pair, P_i is the canonical position which a constituent X with a given GF ϕ occupies in a given neutral sentence, and P_j is the position which it occupies in a corresponding marked sentence structure.

6.2.3. The relationship between the grammatical function O and FOCUS

A point which appears particularly interesting is the relationship between the GF O and the PF FOCUS. In many languages, the constituent with the function O coincides with FOCUS in unmarked sentences, according to a principle of constituent structure which may be stated thus:

> *Principle XV*
> The constituent with the greater depth of embedding is that which bears FOCUS.[358]

As is clear, this constituency property holds independently of the serial ordering of O with respect to V. In other words, it applies not only in SVO languages, but also in SOV languages. We shall call the relationship between the constituent structure and FOCUS the "constituency property of FOCUS." Note that this is not necessarily connected with the property which allocates prosodic FOCUS, since the latter may coincide with a constituent which is not the most embedded. This is not surprising if one considers that the prosodic structure is not, in principle, isomorphic with the syntactic structure.

Principle XV may be reformulated in generative terms, by assigning to the PF FOCUS a syntactic structure representation in which the constituent X = FOCUS is a projection of V (a feature assigned by V). In other words, X = FOCUS may be considered a constituent governed by V (for discussion of this possibility, cf. Kiss 1995; Kiss [ed.] 1995; Jo 1995).

It is immediately obvious that in languages with a differential strategy, principle XV may apply only to neutral or unmarked structures; only in such structures, in fact, do the most deeply embedded constituents belong to the maximal projection of V (that is, they are governed by V). This principle does not, however, generally apply to structures with marked FOCUS. In fact, these have derived representations in which a given constituent carries FOCUS not because of its original constituency property, but because it occupies a noncanonical position. This may apply not only to O, but also to constituents with the GF S, that is, constituents which occupy a node high in the tree structure and are not governed by V. Therefore, in these cases the fact that the constituent which bears FOCUS is positionally defined as the most embedded constituent is neither a necessary nor sufficient condition. If an element is moved from a less embedded to a more embedded position, the definition in terms of "most embedded constituent" will be trivially true (in the sense that it is not the position which defines the constituent). If instead an element bearing neutral FOCUS is moved from a more embedded position to a less embedded one, the definition is simply not true.

In marked sentences, then, the constituent which carries FOCUS is not necessarily the one with the greatest degree of embedding. Furthermore, in languages with PF

492   Rosanna Sornicola

movement (that is, languages with a differential strategy) assignment of FOCUS to a
given GF does not automatically and exclusively depend on a phrase structure prop-
erty that can be represented by a single phrase structure tree. In fact, to take account
of the *differential character* of the property in question, at least two structures are
needed; the first comprising information concerning the canonical positions of the
different constituents, the second comprising information relative to the "derived"
positions. In any event, in these cases formulation of the relationship between FO-
CUS and GF in terms of assignment of the feature [+FOCUS] to a constituent by V
may be held as unsatisfactory.

The problem is no less thorny for languages without PF movement. Given that
they have a fixed position P_k for FOCUS (both marked and unmarked FOCUS), and
furthermore, that it coincides with the canonical position for O, a plausible hypoth-
esis would be that the property whereby V assigns the feature [+FOCUS] holds not
only in neutral or unmarked sentences, but also in marked sentences. However, it
would be interesting to establish whether the constituent with the GF S which turns
up in position P_k becomes "deeply embedded" only because it occupies the canoni-
cal position of an O constituent. This is not, however, an easy problem to solve out
of a theoretical framework.

The hypothesis that in languages such as Basque and Turkish the position of O
determines the position of FOCUS shall be called "Hypothesis 1." Note that such
hypothesis implies that Principle XV holds not only across languages but across the
distinction of unmarked and marked sentences as well.

But to what extent can this be maintained? The problem just raised requires con-
sideration of the relationship between V and the preverbal position in languages
characterized as agglutinating. Is it really because P_k is a conspicuous position in
terms of constituency that it has been generalized as FOCUS position? After all,
principle XV – which should explain why O-position determines position of FO-
CUS constituent – is not so general. We have seen that it does not hold in marked
sentences of languages with PF movement. It may be irrelevant – for other reasons
– also in languages without PF movement. In fact, in Turkish specific semantic fac-
tors – like animacy and definiteness – are at work in determining what constituent
must fill the immediately preverbal position (cf. Section 5.1). The possibility that
nonconfigurational, semantic factors may determine a fixed FOCUS position P_k will
be called "Hypothesis 2."

6.2.4.   Some possible typological correlations with the parameter "FOCUS in P_k"

6.2.4.1.   Hypothesis 1

Two groups of indications, of a typological and a diachronic nature respectively,
seem to favor Hypothesis 1. They offer, in particular, a unitary explanation of the
situation in Basque and Turkish as well as Aghem, an African language of the Bantu

family. The latter language, in fact, has a basic SVO order and a fixed position for FOCUS coinciding with that of O.[359] A second group of indications is provided by Hungarian. As was said in Section 2.3.2, there is reason to believe that grammaticalization of the immediately preverbal position for FOCUS is relatively recent in this language. Furthermore, and more interesting for the present study, it seems to go hand in hand with the increase in the frequency of occurrence of O in immediately preverbal position.

Acceptance of Hypothesis 1 would lead to the conclusion that the FOCUS position is determined by the GF O. There are, however, certain empirical and theoretical difficulties that such a conclusion must be reconciled with. The question that, above all, must be asked concerns the range of the empirical domain of the property that FOCUS is determined by O. Is it a generalization supported by other SOV languages,[360] or is the property in question restricted to Basque and Turkish? Examination of the wider empirical domain is very important for confirmation of Hypothesis 1. In fact, one would expect that in other languages in which O precedes V in the basic/dominant order the property of codification of FOCUS in immediately preverbal position P_k has developed. As will be seen below, results from examination of data from various SOV languages do not appear to conclusively confirm Hypothesis 1 (cf. Section 6.2.4.1.2). From a descriptive point of view, it may be claimed that Hypothesis 1 is restricted to the few languages in which FOCUS occupies position P_k. However, the scope of the potential correlation between O and FOCUS is of a more general nature. This lies more properly in the theoretical domain. For this reason, a conclusion such as that just advanced would not be satisfactory since it would create more problems than it would solve. In particular, it would raise the question: why has this property developed in a tiny number of SOV languages and not in others?

A secondary, but not negligible, empirical problem is that a model of representation such as that arising from Hypothesis 1 would not be compatible with some of the results which have emerged from studies in pragmatics and language acquisition over the last thirty years. It has in fact been shown that the GF S has ontogenetically developed from the PF TOPIC. Furthermore, over the last decades interesting evidence has been gathered that grammaticalization of PF TOPIC into GF S is also a path in the diachronic development of many languages.[361]

6.2.4.1.1. Some typological properties of SOV languages. Examination of the pragmatic properties of WO in a wider corpus of SOV languages offers some food for thought. Japanese, Korean, Kannada, Burushaski, Somali, and Quechua have been examined. In these languages, the basic SOV order has degrees of "dominance" varying according to the basic constituents in the sentence. Although all these languages clearly favor V in the final position of the proper sentence domain, they allow scrambling phenomena of various types within this domain.[362]

494   Rosanna Sornicola

Two parameters have been considered: the properties of the postverbal field and the position of FOCUS. It turns out that, with the possible exception of only Burushaski,[363] in all these languages the postverbal field is given over mainly to afterthoughts or backgrounded material, just as in Basque and Turkish.[364] As far as FOCUS position is concerned, only Kannada and Quechua appear to exhibit a property linking the FOCUS constituent to a given region in the proper sentence domain.[365] However, in neither case is it the same property as that found in Basque and Turkish. In Kannada the immediately preverbal position for the constituent in FOCUS is a fairly common strategy, although position P_1 is accessible to some constituents in FOCUS.[366] Quechua in its turn offers a different picture which will now be considered in more detail, along with Japanese and Somali. These three languages may be considered representative of three different typologies with respect to codification of FOCUS order in SOV languages.

6.2.4.1.2.   *Japanese, Somali, and Quechua.*   In Japanese, the function FOCUS is codified by many processes acting together. The particles *wa, ga, yo, ne,* and *se* appear to be of particular importance; these mark the constituent that immediately follows them as FOCUS.[367] As far as WO is concerned, Hinds (1986: 150) observes that "any perturbation of neutral sentence order will create emphasis" and that "scrambling or postposing will result in a display of emphasis." Although interesting, these remarks must be treated with caution in typological comparison: the concept of "emphasis" to which Hinds refers is not homogeneous with that of FOCUS used in the present study, nor with the descriptive tools exploited by those studying other SOV languages. This shows clearly, then, that typological comparison is difficult even when based on reliable studies of individual languages. In particular, examination of "postposing" as a means of emphasis is not very convincing[368] or at least it suffers the same pitfalls as Kuno's examination of Turkish criticized by Erguvanli. It is unclear – from the available data – whether Japanese behaves in a different way from Basque or Turkish as far as the structuring and function of the postverbal space is concerned. With due caution, then, Hinds' observations lead one to think that Japanese has free constituent movement in the preverbal space and PF movement. In fact, the examples reported by Hinds (1986: 150) show that the constituent in FOCUS, marked by a particle, may occur anywhere in the preverbal space. In any case, what appears to be important is that in Japanese, the canonical space within which "scrambling" is defined is the preverbal space, a situation similar to that found in Basque or Turkish.

In Somali, which does not have a conspicuous position P_k for FOCUS, the entire preverbal space appears to be given over to constituent movement and to movement of the PF FOCUS. P_1 has particular importance and may be considered a conspicuous position for FOCUS. It should be pointed out, however, that – as in Japanese and Quechua – PFs are identified by the occurrence of particles with topicalizing or focalizing value (cf. Svolacchia et al. 1995).

The situation in Quechua offers interesting, but problematic, comparisons with that in Basque and Turkish. In a study on Huallaga variety, Weber (1989: 15) has pointed out that of the characteristic properties of an SOV language, the less consistently displayed is actually the SOV pattern in main sentences. Weber describes an order which is "fairly free, particularly in nonsubordinate clauses" with considerable variation according to speaker and to speakers' styles within the same dialect (Weber 1989: 15, 402–403). This description agrees with that of Calvo Pérez (1993: 41), according to whom in Cuzco Quechua all six orders SOV, S′OV, S″VO, S″VO′, O′VS, and O′S″V should be possible with varying pragmatic values.[369] The structural and pragmatic analysis is complicated by the presence of particles, such as *-mi/-n*, *qa*, the semantic values of which are not easy to analyze satisfactorily. These particles have been described in the literature as "evidential." According to some, *-mi* and *-n* are probably focalizing particles, and *-qa* is probably a thematizing particle;[370] others, while claiming that this distinction is problematic, nonetheless recognize that *-mi* occurs more naturally with NEW or rhematic information,[371] and that *-qa* often occurs with thematic material. In fact, in many contexts the first particle has the characteristic evidential value of 'learned by direct experience'; the second may indicate contrast in some contexts, while in others it is affixed to the focal element in yes/no-questions.[372]

In any case, Weber's analysis allows the following formula for canonical phrase structure (i.e., that to which most sentences in Huallaga Quechua conform) to be drawn up:

(176) $(\text{X-}qa)0^n$ $(\text{Y-EVD})0^1$ V $\{\text{O-}qa, \text{S-}qa\}$

The formula states that there may be any number of elements with *-qa*, followed by the element with *-mi/-shi/-chi* (i.e., the evidentials), which must precede the main verb or be the main verb. The verb in its turn must be followed by one or two elements with *-qa*, which must necessarily be S or O (Weber 1989: 428). Weber observes that the evidential suffix must not be identified with a particular element, for example, the last thematic or the first rhematic element. He claims that the only conclusion that may be reached is that the thematic material occurs to the right of the evidential suffix and the rhematic material follows the last preverbal *-qa*.[373] This observation would seem to indicate that Quechua behaves like Basque or Turkish as far as closure of the sentence space after V is concerned. Whether this means that a position P_k which is salient may be postulated for Quechua seems more difficult to say.

6.2.4.2.   Hypothesis 2

Hypothesis 2 is also not without problems. According to this hypothesis, it would be semantic features – like animacy, definiteness, and referentiality – which determine the position of preverbal FOCUS. The typological correlation with languages such as Aghem would be lost, but it would be possible to obtain a more finely-grained

description of the properties of a particular language. We therefore find ourselves up against a typical problem concerning the relationship between the microscopic and the macroscopic study: the more an analysis takes the peculiarities of individual languages into account, the more difficult it becomes to make typological comparison, to determine uniform types (for example, is it really possible to classify Turkish along with Basque as far as the property of FOCUS position is concerned? In particular, do the features animacy and definiteness have the same function in Basque as in Turkish?), and to establish universal tendencies. It has to be admitted that the more detail one goes into from the microscopic perspective, the more difficult it becomes to determine *uniform and homogeneous types*. A very general problem is that "sameness of properties" across languages is a problematic notion in itself: what prima facie appears as an identity may turn out to be – on closer inspection – a difference. Regarding the positional properties of German, for example, comparison between German and Hungarian has been seen to be illusory (cf. Section 2.4.5.5.5).

The influence of semantic factors can be seen in another possible correlation, that between FOCUS position and the position of WH-words. As is well known, WH-constituents typically have the value [+FOCUS]. In fact, all the languages which have the property of FOCUS fixed in P_k, such as Basque, Turkish, Mongolian, and some Indo-Aryan languages, also have the property of WH-words canonically occurring in P_k.[374] What appears of some interest is the fact that languages such as Japanese, Korean, and Somali, which do not have a conspicuous position P_k for FOCUS, equally do not confine WH-words in such position.[375] Kannada and Quechua, which, albeit in different ways, have typological affinities with Basque and Turkish, diverge in an interesting way with respect to these languages and between themselves. In Kannada WH-words may occur in P_1, in P_k, or in another position in the proper sentence domain,[376] which is of some interest since these are the same positions which a non-WH-constituent with FOCUS may occupy. In Quechua, instead the typical position for WH-words is P_1, a position which does not coincide with that reserved for non-WH- elements in FOCUS.[377] The correlation between fixed position for FOCUS in P_k and fixed position for WH-words is therefore important, but as far as current knowledge goes, it is difficult to establish which of the two factors has acted upon the other. Hungarian, which has fixed position for FOCUS in P_k, but not for WH-words, offers an argument in support of the claim that WH-words may not be the motivating force in the process.

6.2.4.3.   A third hypothesis

A third hypothesis (we shall call it "Hypothesis 3") would be to postulate a relationship between the various properties of PF codification and morphological properties such as agglutination and inflection. In fact, the type without PF movement occurs in SOV languages which are typically agglutinating, such as Basque and Turkish.

Agglutination could be responsible for the linear property of (left) adjacency to V of the FOCUS constituent.

As is well known, agglutinating languages generally have a lesser degree of freedom of constituent movement than inflecting languages, and this seems to be all the more true in languages such as Basque and Turkish, in which agglutination phenomena favor the occurrence of well-known processes of phrasal association of morphemes. These processes are, among other things, clearly evident within head-modifier structures, such as for example V + subordinate clause, in which the modifier is normally attracted to preverbal position.[378] The position of inflectional case morphemes in NPs such as:

(177)   Ankara ve Izmire   gid-eceğ-im.
        Ankara and Smyrna.DAT go-FUT-1SG
        'I am going to Ankara and Smyrna.'

also confirms the impression that agglutinating morphology, which favors phrasal compounds of elements such that a unique inflectional morpheme applies to two lexical morphemes at the same time,[379] may also be a decisive factor in determining phenomena like the conspicuous FOCUS position P_k in Turkish.

Inflecting languages, on the other hand, show a wide range of constituent movement due to the phenomenon described by Meillet (1912) as "autonomy of the word." In fact, in these languages the word is the typical locus of grammatical relations, while in canonically agglutinating languages such as Turkish, this is not the case. This may well account for the differences which have been pointed out between Turkish and Hungarian, with respect to, for example, the influence of the features animacy and definiteness, and the behavior of WH-words. Although distinguished by agglutinating processes, Hungarian is a less typically agglutinating language than Turkish and therefore has a freer constituent order, in conformity with what is postulated by Hypothesis 3.

However, not even the correlation between morphological type and type of PF codification by position can offer an explanation on its own. Agglutinating SOV languages like Japanese, Korean, and Somali have not developed the property under examination,[380] which has an areal distribution confined to a few regions of the globe.

Notes

498   Rosanna Sornicola

to thank the colleagues in the group of research on "Pragmatic Organisation of Discourse" of the EUROTYP Project of the European Science Foundation. The fruitful discussions with them during the years of the EUROTYP research gave me interesting opportunities for reflection. Giulia Perretti and Giovanni Palumbo assisted me in editing the text at various stages of its preparation. To Giovanni Palumbo I also owe valuable comments and observations. Finally, I want to acknowledge financial support from the Department of Modern Philology of the University of Naples Federico II, without which this piece of research would not have been possible.

2. By basic constituent order, the relative order of S, V, and O is meant, as in the standard meaning taken from Greenberg (1963). The term "constituent" is of course to denote both elements of a categorical nature and those of a functional nature, such as S, O. Although this confusion is often pointed out in the typological literature, the expression has fallen into common use. For terminological and conceptual clarification with respect to the general organization of the present work, cf. Section 1.8.

3. These alterations depend in the first place on the language type, but the factor of "style" also has a role to play.

4. Regarding the importance of this distinction, see Daneš' (1967: 217) observations.

5. In the generative literature this has been represented in terms of movement rules. For an assessment of this point of view, cf. Section 1.6.2.

6. This tradition extends from the work of theorists of the Prague School (cf. Sornicola and Svoboda 1992) to the more recent works of Dik and Comrie (see, e.g., Dik 1978, 1983, 1989; Comrie 1981b). For studies concerned more specifically with WO, cf. Siewierska (1988, 1998) and Connolly (1991).

7. Cf. Payne (1992), Mohanan and Mohanan (1994).

8. For an assessment of the role of psycholinguistic (or cognitive) factors in typological research on WO, see Tomlin (1986: 3–5, 133–139). Chafe (1986) and Givón (1983) have especially emphasized the cognitive constraints on the organization of sentential or textual information (cf. also the various contributions in Givón [ed.] [1983]). For a different line of research, see Vennemann (1972) and Hawkins (1983, 1994).

9. Cf. Holden and Krupp (1987: 269).

10. Cf. the papers in Nuyts and de Schutter (1987).

11. For the influence of Latin and Greek on the syntax of modern languages of Europe, see Blatt (1957), Nykrog (1957), Sørensen (1957), Stender-Petersen and Jordal (1957). Of particular importance were the linguistic effects of Bible translation from either Greek or Latin on European vernaculars during the Middle Ages (see the contributions in Lampe [1969: 338–491]).

12. For problems concerning the relationship between history and typology, cf. Croft (1990: 18–25), in particular p. 24: "This discussion of the sampling problem should make clear to the reader how diachronic considerations enter typology even at its methodological foundations"; Bell (1978: 146): "I think that a case can be made that such research [on language universals] can properly be conceived as sampling language changes, not languages themselves."

13. Cf. Sinor (1990: 16).

14. Cf. Sinor (1990: 16–18).

15. Cf. *PET* (1995–1997: IV, 437b). The question of an exact boundary for Europe arises only in the sixteenth century and acquires truly scientific form only with the birth of modern geography.

16. Interesting observations regarding this question may be found in Sinor (1990).

17. Primary sources were gathered for Russian, German, English, French, Italian, and Spanish.

18. For a discussion of these problems, cf. Croft (1990: 25–26).

19. The methodological problems that this option involves have so far not received sufficient critical attention in the literature on WO: for a defense of this point of view, cf. Abraham (1995).

20. Cf. Greenberg (1963: 76), Siewierska (1988: 8 ff.), and Connolly (1991: 16–19). For a criticism of the notion of "basic WO" as traditionally used in contemporary typology, see Brody (1984). Connolly (1991: 37) convincingly criticizes the notion of "underlying order" of earlier generative models.

21. Cf. Comrie (1981b: Ch. 4) and Croft (1990: 24–25).

22. Until recent years, studies on variation of constituent order in a typological perspective have not been many: see Steele (1978), Holden and Krupp (1987), Connolly (1991), and Hawkins (1994).

23. Cf. Daneš (1967: 218), who compares the concept of "neutral order" to that of "usual order." Cf. also Dryer (1995).

24. Cf. the excellent observations made by de Meij and Marácz (1986: 268 ff., and especially 273) and Comrie (1996). An analysis of the problem may also be found in Kiss (1995). This problem will be readdressed in Section 6.

25. Cf. Daneš (1967: 221).

26. Cf. Daneš (1967: 223). On the problem of variable orders, cf. this chapter, Section 3.

27. See, for example, Dik (1978: 21, 1989: 359 ff.). This idea can also be found in recent generative models.

28. Cf. Dik (1989: 342–370) with a lengthy discussion of this.

29. Or rhematizing if neutral structures are considered as being without FOCUS.

30. Cf. Weil (1879) and for the history of ideas on WO, cf. also Holland (1980: especially Ch. 1). The "prehistory" of ideas on WO goes back to the Middle Ages and even beyond until Classical Antiquity: cf. Scaglione (1972). Of particular interest is the extensive and philosophically based discussion on "ordre naturel" in Enlightenment France: for an overview of the debate, cf. Bossong (1990: 229–269).

31. An interesting examination of the post-Greenbergian movements can be found in P. H. Matthews (1994: 44–45).

32. Cf. de Groot (1989: 10–11 and Ch. IV) and Dik (1989: 363–365).

33. See, for example, the treatment of the GF S in Keenan (1976).

34. On the concept of maximality, cf. Heltoft (1986).

35. Abraham claims that the concepts of theme and rheme should not be considered only in their textual dimension, but should be determined positionally and with respect to the structure in constituents (Abraham 1995: 605). While theme is here synonymous with TOPIC, rheme is not synonymous with FOCUS; the latter in fact is considered to be the stress-bearing constituent and as such is a subconcept of rheme.

500   Rosanna Sornicola

36.   Note that the expression "organized according to T-R" is equivalent here to "d-configur-ational."
37.   On the question, cf. Abraham (1995: 606–607).
38.   In the present work, it is not possible to study the properties of WO in interrogative sentences since they require a detailed examination in themselves.
39.   In some cases (cf. the section on Russian) SFs have also been important.
40.   Cf. Daneš (1967: 217 and Note 2) for the history of the terminological and conceptual pair TOPIC/COMMENT (FOCUS). See also van Oosten (1985), Yokoyama (1986: Part 1), Holden and Krupp (1987: 258–259), Dik (1978: 87 ff., 141 ff., 1989: 266 ff., 277 ff.), Sgall et al. (1986: 175 ff.), Sornicola (1993a), Lambrecht (1994: 117 ff.), Maslova and Bernini (this volume), Miller (this volume).
41.   For an examination of the definitions of "theme" in textual linguistics, cf. Halliday (1967a, 1967b, 1967c, 1968); Dik (1978: 132 ff.).
42.   The reverse relationship is of course not valid.
43.   Cf. Höhle (1982: 99), who observes that FOCUS does not necessarily coincide with a constituent, but may, in particular in presuppositional conditions, be distributed over a string of constituents.
44.   Note that V may well be thematic, but as far as the distinction made between TOPIC and theme is concerned, this is a different property.
45.   Note that in many of the languages of Europe, the category "preposition" is also an element in the set of possible constituents over which the function FOCUS ranges: cf. Italian *Sono stato IN casa* 'I was AT home' (lit. [I] am been IN house) with corrective FOCUS. The typological distribution of this property would, however, deserve further investigation.
46.   Cf. von Stechow and Uhmann (1986) for an overview of the relationships between prosodic and phrasal FOCUS in the generative literature.
47.   For a recent proposal regarding differentiation in this sense, cf. Abraham (1995: 615 ff.). For problems concerning the prosodic analysis, cf. Ladd (1996: 160 ff., 198 ff.) and Cruttenden (this volume).
48.   Cf. Dik (1978: 131 ff.), who opposes identification of FOCUS with the constituent which receives main stress, and Comrie (1984), with a clear formulation of this point of view.
49.   For Turkic languages, cf. Comrie (1981a: 79–80); of particular importance for Turkish is Erguvanli's (1984) study; for Mongolian, see Poppe (1954). For the Uralic languages, see Comrie (1981a: 92, 122–124), who emphasizes how important a property the SOV/SVO feature is in differentiating the Eastern Uralic languages – some of which, such as Mari (Cheremis), have been exposed to a strong Turkish influence – from the Western ones. Some of the latter in fact (especially the Balto-Finnic languages) have come under the influence of Indo-European languages (cf. Sauvageot 1973: 9–94). For Basque, cf. de Rijk (1978, especially 109) and Ortiz de Urbina (1995). The literature on Hungarian is extensive: cf. Section 2.3.2. For the general problem of the preverbal position as locus of FOCUS placement, see Kim (1988).
50.   For a comprehensive study of these characteristics in the Uralic and the Altaic languages, see Bese et al. (1970). For the correlation between the position of WH-words

and the position of FOCUS in the Altaic languages, cf. Comrie (1981a: 79–80, 124); for Turkish in particular, see Erguvanli (1984: 34–35). On the position of WH-words in Basque, cf. de Rijk (1978: 103) and Ortiz de Urbina (1995: 100). For Hungarian see, among others, Kiss (1987: 56), Horvath (1986: 44 ff.), de Groot (1989: 25). The agglutinating characteristic is more problematic for the Uralic languages (cf. Comrie 1981a: 177–220), which move to a greater or lesser extent away from the type in question, "while still being in general closer to this type than to the fusional type represented by the older Indo-European languages" (Comrie 1981a: 117). For the general view of the relationship between SOV order and agglutination, see Lehmann (1973).

51. For an introduction to Indo-Aryan languages and their numerous dialectal varieties, cf. Masica (1991: 8–23); for some problems in identifying such languages, cf. Masica (1991: 27 ff.).

52. Cf. Masica (1991: 335), Bhatt (1994).

53. Cf. Masica (1991: 395, especially ex. (463)), altogether similar to the Turkic structures with respect to the positional properties of the WH-constituent.

54. Cf. Masica (1991: 332–333, and 395 for WO); for the complex question of morphological properties, cf. Masica (1991: 212–216). On WO in Indian languages, see also Swarajya Lakshmi and Mukherjee (1996).

55. Cf. Delbrück (1878: 14, 1888: 16, 1900: 82–83) and more recently Lehmann (1974: 30 ff.), C. Watkins (1976), Verpoorten (1977). For a historical overview of the problem, cf. Holland (1980: Ch. 1).

56. Cf. Delbrück (1900: 259–260) and Lehmann (1974: 121). The parameter in question is therefore presumably an innovation in modern Indo-Aryan languages.

57. The situation in Somali and Quechua is further complicated by the fact that in both languages, the constituent with the feature [+FOCUS] is classified according to its combination with certain particles as well as by its position and prosodic properties. For Somali, cf. Svolacchia et al. (1995); for Quechua, cf. Muysken (1995).

58. Cf. Austerlitz (1970). Moreover, not even this relationship is implicational: Cushitic languages, which are SOV, are highly synthetic (cf. AWL 1994: 274). Modern Indo-Aryan languages, which are SOV, as previously mentioned, are not uncontroversially considered agglutinating.

59. Horvath adds that the principle spells out a necessary but not sufficient condition for the position of WH-phrases.

60. Georgian presents an interesting case; D is found (cf. Aronson 1990: 184 [sentence no. 16], 186 [sentence no. 44]) but not A (cf. Section 2.3.1). The situation in Armenian is the same (cf. Comrie 1984). However, as is well known, this does not deny the validity of the relationship *A entails D*.

61. The considerable degree of dialectal fragmentation found in Basque must be taken into consideration. The data made use of here are from literary Navarro-Labordino, a variety that enjoys considerable prestige and that served as the base upon which the linguistic standard was established.

62. Lafitte (1962: 47) translates (2) as 'le vase, c'est le père qui l'a jeté.'

63. According to Lafitte (1962: 48), in exceptional cases, inversion of Aux and V participle may be a device for focalizing the constituent which precedes the verbal block: cf.

502   Rosanna Sornicola

Aitak du aurdiki 'c'est le père qui l'a jeté,' *Aitak untzia du aurdiki* 'c'est le vase que le père a jeté.' Lafitte observes that "ici le relief est très accusé, même violent." See, however, the observations of Rebuschi (1984: 77) on this point; he claims that inversion of Aux and participle V is anything but exceptional. On the basis of an analysis conducted on actual texts, he maintains that the phenomenon has no contrastive value at all.

64. Examples (3a)–(3f) are taken from Erguvanli (1984: 33–34). For a discussion of Turkish WO, see also Banfi (1998).

65. As observed by Erguvanli (1984: 37).

66. Cf. Comrie (1981a: 79).

67. In other words, word order in a sentence would conform to an increase in the so-called communicative dynamism: cf. Firbas (1965) and for a more recent formulation, Firbas (1992). This theory has been adapted in the framework of American functionalism by Chafe (cf. latterly Chafe 1994: Ch. 13).

68. Regarding this suggestion, see for example Desző (1973: 254).

69. This description does not seem to take into account the difference between broad and narrow FOCUS: O + participle is broad FOCUS, while O must be narrow FOCUS. On the other hand, there are other ways of focalizing V: in Vizcayno and Guipuzcoano, the verb *egin* 'to do' is inserted immediately after V and before Aux (cf. Rebuschi 1984: 71).

70. Indirect O raises specific questions and the problems associated with it which arise in Basque and Turkish are not considered here. Indirect O, in any case, has not been held to be a basic constituent.

71. Cf. Comrie (1981a: 222). For an analysis of WO problems in Georgian from a generative perspective, cf. Boeder (1989).

72. The problem in general, as well as the results reported here, are discussed by Vogt (1971, 1974), where a critical examination of other literature may also be found.

73. The examples are taken from Vogt (1974: 52–53). The linguistic remarks and translation into English are due to Marcello Cherchi.

74. I quote this word in the form given by Vogt (1974: 72) (note that in Standard Georgian it has the form /gamarʒvebuls/).

75. In Georgian folklore a *dev* is a type of demon. I would like to thank Marcello Cherchi for supplying me with this clarification.

76. Note that Vogt exemplifies each of the two structures with only a single occurrence from his corpus.

77. Cf. Harris (1981: 14–18).

78. The two different types of prosodic contour are described by Kálmán et al. (1986: 130–131). "Eradicating stress" is defined as "a main stress that is not necessarily stronger than a normal main stress, but which 'eradicates' all subsequent stresses, and thus, cannot be followed by any more [*sic*!] main stress (except for multiple contrast)" (Kálmán et al. 1986: 132).

79. On the other hand, in sentences which have a focused constituent, the postverbal position of O is strongly preferred if the sentence is imperative or contains a negation (Behrens 1989: 142).

80. These sentences are reported by Kenesei (1986: 144), without a detailed pragmatic analysis. I integrated the pragmatic interpretation which I put forward here on the basis of a native speaker's judgements.

81. This characteristic has been ably described by de Groot (1989), according to whom Hungarian has two special positions in the clause, P_1, or the position of the constituent with TOPIC function, and P_0 or the position of the constituent with FOCUS function.

82. The examples are taken from Kálmán et al. (1986: 133, 140). The HOCUS position may also be occupied by a pre-verb, cf. Kálmán et al. (1986: 133, 140), Tompa (1985: 227) (with different terminology and descriptive devices). Cf. also Behrens (1989: 148–149), according to whom the position in front of the verb in nineteenth century texts is to be analyzed as being without FOCUS for [−referential] O; cf. also, however, what Behrens (1989) has to say in Note 182.

83. Komlósy (1986: 218) claims that in a sentence such as:

(i)   Mari   lát-t-a   János-t   a   kert-ben.
        Mary   see:PST.3SG.SBJ.3SG.DEF.OBJ   John-ACC   ART   garden-INESS
        'Mary saw John in the garden.'

(with main stress on *látta*) the FOCUS position is completely empty.

84. Komlósy (1986: 220). Hungarian, however, has structural properties which differ from those found in Turkish with respect to the placement of the FOCUS constituent and this can be clearly seen from two further characteristics: (a) differences in the effects of semantic features in determining the constituent occupying the immediately preverbal position (cf. Section 5.1); (b) differences in the distribution of subordinate clauses. De Groot observes that when the FOCUS function coincides with a finite subordinate clause, the latter may not occur in front of the verb, but after it; in this case the immediately preverbal position will be filled by a "dummy element" which represents the subordinate clause and which indicates the focal property of the latter element (de Groot 1989: 25–26).

85. It should be emphasized that only structures in which the positions $P_1 \ldots P_n$ constitute the *proper domain* of the sentence, and not structures with extrasentential constituents (right- or left-dislocated, as in the schema of Fig. 7), are considered here.

86. From a historical point of view, the suggestion, defended by Meillet in many works, that there was originally a close relationship between "autonomy" of the word and freedom of WO in Indo-European is important; later, the historically documented Indo-European languages probably developed specific patterns of WO. For an interesting examination of hypotheses concerning WO in Indo-European languages between the nineteenth and twentieth centuries, cf. Holland (1980: Ch. 1).

87. For the development of the WO patterns from Latin to the Romance languages, cf. Meillet (1906: 8–9) who, in conformity with his more general ideas, saw in these languages a loss of word autonomy. The influence of the hypothesis that word autonomy correlates with freedom of constituents in the sentence is also found in Wartburg (1950: 104) with respect to free word order in Old French (constrained, however, by the rules of the everyday language). For the diachronic development of French, cf. Marchello-Nizia (1995). For the diachronic development of English, cf. Jespersen ([1922] 1969: 344–345, 355–357).

504   Rosanna Sornicola

88. For German, cf. Section 2.4.5.5; for Icelandic, see Pétursson (1978: 138 ff.), Kress
     (1982: 263 ff.), and in a generative framework, Sigurðsson (1989: 5–6). The situation
     in Modern Greek is also interesting: despite conservation of a case morphology, the
     modern language has moved away from the old language (cf. Dover 1960; Schwyzer
     1950: 691 mentions a "habitual" or usual order, which limited the structural freedom
     of Classical Greek), in reducing the range of structural possibilities and in choosing
     SVO as the predominant pattern (cf. De Simone Brouwer 1921: 217). Thumb (1912 :
     200–201) maintains that "modern Greek vernacular has, on the whole, maintained the
     ancient Greek freedom in the order of words, i.e. all kinds of combination are possible
     in the sequence of the composite parts of the clauses." However, the counterpart of this
     potential freedom is "a recognized normal sequence of words, so that any deviation
     from the same lends a special emphasis on the irregular member" (Thumb 1912: 201).
     Cf. also Kalitsunakis (1963: 183 ff.). For a recent account in a generative grammar
     framework, cf. Tsimpli (1995).
89. In a corpus constituted of a sample of 2254 sentences taken from Ælfric's Homilies
     (the last phase of Old English, circa 1000), Kohonen (1982) found 54 % SVO order
     and only 13 % SOV order. Cf. furthermore Mitchell (1985: II, 963 ff.).
90. Dik (1989: 351) is of the same opinion.
91. On the diachronic and typological problems relating to the basic SVO order in these
     languages, cf. Comrie (1981a: 92–93, 121 ff.); the literature on WO in Finnish is quite
     large: cf. Hakulinen (1961, 1979: especially 499 ff.) and van Steenbergen (1989) who,
     on the basis of experiments, shows the configurational nature of the language. Vilkuna
     (1989: 9) observes that in a simple sentence consisting of S, V, O and some further
     complement or adjunct "all permutations are grammatically possible." In both Estonian
     and Finnish, specific grammatical restrictions on the complex form of the verb may be
     observed (cf. Kippasto 1989: 86 with a comparative study of Estonian and Votian;
     Vilkuna 1989: 9).
92. Cf. Sauvageot (1973: 193–194).
93. Cf. Tauli (1983: 24–25), although he uses the term "emphasis" and not focalization.
94. Both examples are taken from Tauli (1983: 54).
95. All the examples are taken from Hakulinen (1961: 315). I have indicated with an apos-
     trophe in front of the constituent the fact that this is the one that is stressed.
96. The suggestion that all deviations from basic SVO order ("default") are marked has
     also been put forward more recently by Vilkuna (1989: 42).
97. Hakulinen (1961: 316) observes, in fact, that the beginning and the end are "the most
     stressed positions in the sentence."
98. Vilkuna has adapted to the d-configurationality framework an idea put forward by
     Karttunen and Kay (1985), according to whom the discourse functions contrast and
     topic probably have the following fixed positions:

     (i)   Contrast, Topic, . . .

     Vilkuna, however, proposes specific descriptive differences with respect to the d-con-
     figurationality model put forward by Kiss for Hungarian.
99. Cf. Vilkuna (1989: 37).

100. Vilkuna (1989: 42). The pragmatic analysis is, however, perplexing: if a T constituent, defined as the intersection of the concept of OLD information and ABOUTNESS (cf. Vilkuna 1989: 38, 79–84), appears in these configurations, how is it possible that the examples quoted are "all-NEW sentences"?

101. On p. 42 of Vilkuna (1989) "default T" is defined as "the grammatical subject, if present."

102. Vilkuna also mentions "a little K," which occurs in the V-field, but it refers to something else.

103. Vilkuna (1989), however, often refers to this author (on p. 41, his 1976 work is quoted).

104. I refer to the suggestion that because a constituent is a TOPIC, it must have referential value. In fact, this concept is only indirectly alluded to on p. 40 (Vilkuna 1989), in the wake of work by Hakulinen and Karlsson, which talks of T in terms of a "nominal constituent." Vilkuna herself observes that T is a function "normally" realized by the constituents NP and Adv (Vilkuna 1989: 38).

105. For an extensive and lucid account of FOCUS function values, cf. Dik (1989: 288 ff. and in particular 282–285).

106. On the way such a model may apply to Finnish, see Tarvainen's (1985: 363) observations. Vilkuna herself has reservations over this.

107. For Bulgarian, cf. Beaulieux (1933: 359–360) and
Guentchéva (1994); for Polish, Meillet and De Willman-Grabowska (1921: 203–204); for Serbo-Croatian, Meillet and Vaillant (1924: 288). The literature on Czech from the Prague School is vast (see the fundamental work by Mathesius 1941–1942; Daneš 1967; Havránek 1968; Uhlířová 1969; Svoboda 1984; Sgall et al. 1986: 175 ff., 194 ff.); the literature on Russian is also extensive (cf. Adamec 1966 and Krylova and Khavronina 1976 for an approach in terms of Praguean functionalism; see also Benoist 1979, 1988 and the works of Nikolaeva, Kovtunova, Sirotinina, Lapteva, quoted in Yokoyama 1986 and Fougeron 1989, who may be referred to for both the bibliography and a critical examination). In a different perspective, studies based on spoken texts have recently shown that basic SVO order may be postulated for various Slavonic languages. For Russian, Holden and Krupp (1987: 267) claim that "there is strong support . . . for the notion that there is a 'basic' SVO sequencing of constituents which affects native speaker's acceptability judgements of sentences regardless of their contextual associations, i.e. independent of the distribution of GIVEN/NEW information, or theme-rheme, etc." In a similar vein, Urošević et al. (1986: 178) point out that in Serbo-Croatian "OSV is the next most frequent order in speech after SVO, while V-initial orders are the least frequent." In the generative literature, Babby (1978), making the case for Russian, and Tajsner (1990) for Polish, both oppose the view that SVO is the basic order (Babby claims that there is no independent basic WO in Russian other than that determined by pragmatic conditions).

108. The two studies use terms and concepts which are variously drawn from the Slavonic linguistic tradition (theme and rheme, etc.). I will endeavor as far as possible to adapt the data and the analyses presented in them to a different theoretical framework, such as the one presented here. Cf. Yokoyama (1986: 173–175) for the history of the problem of WO in Russian and the impact of the Praguean literature has had on it.

506   Rosanna Sornicola

109.   Cf. Meillet and Vaillant (1924: 288). Dik (1989: 364) claims that Serbo-Croatian has
       a P_1 CLT P_0 V X structure (where CLT is a "strongly defined clause-second position
       used for a variety of clitics"). This model has, however, been drawn up according to
       the view that the GFs S and O are nonexistent in Serbo-Croatian; the general substance
       of this view has been criticized here in Section 1.

110.   Cf. Gvozdanović (1981: 127), who reports the following orders as having marked FO-
       CUS: *Petar knjigu čita, Knjigu čita Petar, Knjigu Petar čita* 'Peter reads a book.' On
       the position of interrogative pronouns, cf. Gvozdanović (1981: 125).

111.   Cf. Fougeron (1989: 206–207) and the numerous references reported there.

112.   Fougeron's monograph is basically rooted in the framework of the Praguean School.
       The concept of "theme" is defined as "known" or "GIVEN", that of rheme as "NEW".
       The limitations of these concepts are pointed out (cf. Fougeron 1989: 4–5), but no
       alternative formulation is drawn up. A key concept in the work is that of "noyau
       d'information," borrowed from Perrot and defined as "l'élément qui est reconnu par
       nos auditeurs comme le porteur de l'information principale du message" (Fougeron
       1989: 108); on the prosodic level, this element always carries the sentential stress. The
       concept of "noyau d'information" partially overlaps with that of FOCUS, an expres-
       sion never used in the monograph. It all depends on the textual and the prosodic dimen-
       sions: although the author examines in detail the various positions and constituents in
       which the core of information falls, this is done from an entirely empirical perspective.
       Amongst the various consequences of this approach, I will merely point out the lack of
       distinction between all-in-FOCUS structures and structures with FOCUS on a single
       constituent. The work is, nonetheless, of considerable value for the richness of its ex-
       perimental and contextual analyses. The theoretical categories with which Yokoyama
       (1986) has worked seem to be more sophisticated; Yokoyama's work shows the com-
       plexity of the relationship among suprasegmental factors, constituent structure, and
       semantic and pragmatic factors.

113.   Cf. Yokoyama (1986: 198). The author observes furthermore that the neutral, context-
       free variants correspond to what she defines as Type I utterances. Nichols (1985: 190–
       191) proposes to give up the terminological and conceptual distinction of neutral vs.
       non-neutral intonation in favor of level vs. nonlevel prosody.

114.   Fougeron (1989: 207). For the textual distribution of these structures, see Section 5.2.

115.   Fougeron (1989: 305–308).

116.   The example is taken from Fougeron (1989: 288).

117.   Fougeron's pragmatic interpretation of sentence (27) does not seem entirely convinc-
       ing. According to the author, the prosodic characteristics (that is, the lack of the two
       prosodic indicators of theme: a rise in the fundamental frequency of the stressed sylla-
       ble and the possibility of a virtual pause between the constituents) allow the structure
       to be interpreted as entirely rhematic. This analysis, which pays attention mainly to
       prosodic factors, is furthermore based on a factor to which too much emphasis has
       perhaps been given: replies given by a group of hearers, who saw a strong cohesive-
       ness in such a structure (cf. Fougeron 1989: 294). The impression of cohesiveness,
       however, may have nothing to do with the rhematic property being distributed over the
       entire sentence.

118. The example is taken from Fougeron (1989: 291).
119. Fougeron (1989: 291–292) observes that its role in the message and its prosodic characteristics "permettent de considérer que le composant c'est le thème." Fougeron notes furthermore that in structures of this sort, the complement of V "introduit un *élément nouveau*, qui situe l'information contenue dans la séquence SV" (Fougeron 1989: 289).
120. Fougeron (1989: 292).
121. Examples (29) and (30) are taken from Fougeron (1989: 298).
122. The example is taken from Fougeron (1989: 419).
123. Note that in Russian, as in other languages of Europe with high WO flexibility, movement of a constituent from its canonical position does not in itself constitute a precondition for focalization of that constituent, but may result in an "all-in-FOCUS" structure. This property seems to regard V-initial structures in particular: cf. Section 4.2.
124. Cf. Srpová (1990: 408), who in the contrastive analysis of Czech and French WO observes that a WO variation in Czech may correspond to such a variation in French as well, but this often results in prominence either by prosodic means or by means of a grammatical construction of the type 'It is X who . . .'.
125. Examples (32)–(36) are taken from Fougeron (1989: 427–437).
126. Fougeron herself (1989: 440–441) observes that sentences such as (32)–(36) may encode two types of opposition: a phrasal opposition, in which the constituent which carries the nuclear stress determines a contrast within an adversative construction; or a paradigmatic opposition, for which the element in question determines a contrast with respect to other potential elements.
127. This characteristic seems to coincide with a property mentioned by Fougeron (1989: 362), that it is the position of the stress and not the type of constituent on which it falls that determines the overall melodic pattern.
128. The data and the interpretations offered by Fougeron are not always consistent. This could make the conclusion arrived at here problematic. For example, regarding C S V′ structures (where C stands generically for "complement"), Fougeron observes that C, separated from V (it normally would be postposed) and isolated from the rest of the sentence, is obligatorily "assigned prominence" at the prosodic level (Fougeron 1989: 425). In Fougeron's terminology, such structures are neither entirely rhematic (that is, all-in-FOCUS sentences) nor are they sentences segmented into unmarked theme and rheme. In any case, it is not clear why in other cases (cf. the example quoted here as (28)) a complement separated from V and isolated from the sentences should be described as "theme" and the whole structure considered divided into theme and rheme. These inconsistencies are probably due to the nature of the conceptual tools with which Fougeron has worked.
129. Yokoyama's data reported as (25)–(26c) are crucial to this.
130. There exist a few limited structural exceptions as well: see the discussion of example (42) further on.
131. It may be that the difference between Italian and Spanish on the one hand (where an alternative possibility is to mark the moved constituent with an identifying construction) and French on the other lies in the different prosodic properties of these languages.

508   Rosanna Sornicola

132.   The interpretation in which the constituent *Mario* in (42) coincides with a TAIL is
        not considered here. Out of context, (39) would be ambiguous (that is, it would allow
        both the interpretation determined by the OVS functional structure as well as the inter-
        pretation determined by the SVO functional structure); (40) instead could not in any
        circumstances ever be interpreted as a SOV structure.

133.   As would be the case in *LUCIA, ha aiutato Mario*; *LUCIA, Mario ha aiutato*; *Ha
        aiutato MARIO, Lucia.* If the fact that the nuclear stress falls on *Mario* is taken into
        account, no virtual pause could occur in (42) (*Ha aiutato LUCIA, Mario* would be
        possible, but this is another type altogether).

134.   On the semantic interpretation of structures with leftward movement of O without a
        resumptive pronoun in Italian, cf. Berretta (1998).

135.   Cf. Sornicola (1988); Weinert and Miller (1996).

136.   Parentheses denote the optionality of the pause. On dislocations in Italian, see Cinque
        (1977, 1990: Ch. 2), Sornicola (1981), Berruto (1986), Benincà (1988), Maslova and
        Bernini (this volume).

137.   The only exception is (43), which allows a contrastive interpretation of 'Mario' in a
        prosodic contour in which there is no pause before Mario.

138.   On the problem of how to compute the position of the postcopular constituents in
        FOCUS in cleft sentences of this kind, see the observations on Celtic languages in
        Section 2.4.6.2.

139.   This, of course, applies to structures containing verbs with two arguments; as to one-
        argument structures, see Section 3.

140.   In Italian an isomorphic type *Mario è LUCIA che ha aiutato* [lit. 'Mario it-is LUCIA
        that he-has helped'] is possible, although rare. The Romance types are also interesting
        because they add weight to the hypothesis put forward in Section 2.4.6.2 concerning
        the relationship between the identifying structure and WO in Celtic languages.

141.   Note that French has the usual S topicalization possibilities: *Son père, il m'a donné
        le livre*. The types in question have been dealt with in different ways by Lambrecht
        (1981), Barnes (1985), and Blanche-Benveniste (1990).

142.   The French syntactic types with isolation of constituents have been dealt with also,
        among others, by Tesnière (1959: 175), Bossong (1981), Heger (1982), Jacob (1990:
        78 ff.). The typological characteristics of French with respect to these types become
        clear when compared with those of the Slavonic languages: cf. Srpová (1990).

143.   Structures such as those found in the Romance languages with postverbal S in FOCUS
        (cf. (42)) may be rendered in English by assigning the nuclear stress to S placed in
        typical preverbal position (cf. Daneš 1967).

144.   O in P_1 may instead sometimes have a "highlighting" value, as in the following exam-
        ple from literary prose, where O is NEW: "This morning the village school opened.
        I had twenty scholars. But three of the number can read: none can write a cypher.
        Several knit and a few sew a little. They speak with the broadest accent of the dis-
        trict. At present, they and I have a difficulty in understanding each other's language.
        Some of them are unmannered, rough, intractable, as well as ignorant; but others are
        docile, have a wish to learn, and evince a disposition that pleases me. I must not forget
        that these coarsely-clad little peasants are of flesh and blood as good as the scions of

gentlest genealogy; and that the germs of native excellence, kind feeling, are as likely to exist in their hearts as in those of the best-born. My duty will be to develop these germs: surely I shall find some happiness in discharging that office. *Much enjoyment I do not expect in the life opening before me*: yet it will, doubtless, if I regulate my mind, and exert my powers as I ought, yield me enough to live on from day to day" (*Jane Eyre* [ed. Nal Penguin, New York, 1960: 361]).

145. Italian and Spanish may also make use of simple WO variation or clefting in order to encode narrow FOCUS.

146. Note that a cleft construction such as *It is MARY that I love* does not have the same semantic representation as structure (48): some speakers in fact claim that the cleft structure expresses a stronger contrast than (48).

147. However, cleft constructions in English are more frequent in spoken than in written registers (cf. Breivik 1986: 817).

148. The data gathered for Russian by Holden and Krupp (1987), as well as the data gathered for Serbo-Croatian by Urošević et al. (1986) show this unequivocably. Both Russian and Serbo-Croatian data were obtained from comprehension tests. However, Fougeron's (1989) data for Russian, drawn from a corpus of spontaneous spoken language, also lead to similar conclusions.

149. For the origin of Diderichsen's ideas, cf. Heltoft (1986) and in particular Henriksen (1986: 75–76). A summary of Drach's ideas and a picture of the scientific influences on him may be found in Etzensperger (1979: 28–51). The influence of Ries and also of various structuralist models on both Diderichsen and Drach clearly emerges from the works of Etzensperger and Henriksen (cf. Etzensperger 1979: 26–27; Henriksen 1986: 75–76). Moreover, regarding the idea of a sentence schema comprising empty positions, Etzensperger puts forward the hypothesis that Drach may have been influenced by some of Bühler's psychological experiments (Etzensperger 1979: 33 and n. 20 for further observations).

150. For Diderichsen, the verb was "the element which indicated the relationship of the conceptual content of the sentence to the reality" (Henriksen 1986: 76), while for Drach "ein geschlossener Satz kommt erst zustande, wenn ein Verbum finitum ihn trägt und seine Vorstellungsinhalte als wirklich und gültig beglaubigt" (cited in Etzensperger 1979: 35).

151. These traditions are well represented in the rich bibliography cited by Behaghel (1932: 1–3, 10–11).

152. This position and the properties which defined it were identified for the old Indo-European languages by Wackernagel (1892). However, the unstressed nature of the verb in the main clause had already previously been recognized for Vedic (cf. Delbrück [1888] 1976: 35–36 with bibliography; Delbrück 1900: 41–42, where the main V of independent sentences is ascribed to a category of words that may or may not have stress; furthermore it is assigned an intermediate stress value between typically unstressed elements like particles and fully stressed words). Wackernagel's idea also had an important impact on the study of WO in the Germanic languages (cf. Wackernagel 1892; Scaglione 1981: 110 ff.). It provided an explanatory hypothesis for the syntactic change from the order with the verb in main clause final position, docu-

510   Rosanna Sornicola

mented in the oldest phases of the Indo-European languages, to the order with the verb in medial position, according to many already characteristic of the oldest phases of the Germanic languages (for an overview of the various opinions between the end of the nineteenth century and the first decades of this century, cf. Scaglione 1981: 108–126).

153. Cf. Behaghel cited in Etzensperger (1979: 165), Drach cited in Etzensperger (1979: 35).

154. Note that in some grammatical treatments, such as the recent generative one, the V-2 phenomenon is defined on the basis of declarative sentences and WH-interrogatives: cf. Haider (1986), Holmberg (1986), Platzack (1986).

155. The examples are taken from Platzack (1986: 27). Cf. also Dooley Collberg (1988: 132–133).

156. The German examples are taken from Helbig and Buscha (1989: 564–565). Cf. furthermore Haegeman (1991: 522–523) for Dutch examples, and Pétursson (1978: 138) for Icelandic examples.

157. In Danish and Swedish only some types of adverbs may intervene between Aux and nonfinite V in the structure Aux + nonfinite V in main clauses; in both languages the characteristic structure of main clauses may be represented as:

(i)   Initial position   Finite V (S)   Sentential adverb   Nonfinite V
           O/complement/S   Other adverbs

(cf. Allan et al. 1995: 492; Holmes and Hinchliffe 1994: 503). In German and Dutch instead the characteristic structure with *Satzklammer* is found (cf. Engel 1991: 304–305; Haegeman 1991: 523). The presence vs. absence of a structure with *Satzklammer* is therefore a phenomenon which differentiates contemporary Germanic languages. The asymmetry between the position of finite V in main clauses (P_2) and in subordinate clauses (P_n) also marks a distinction between the modern Germanic languages: again, German and Dutch have in common alternation between V-2 order in main clauses and orders with finite V in final position in subordinate clauses (cf. Scherpenisse 1986), whereas the Scandinavian and Icelandic languages do not have such alternation (cf. Pétursson 1978: 139; Platzack 1986: 28; Allan et al. 1995: 492–493, 497; Holmes and Hinchliffe 1994: 503, 509). Moreover, Danish and Swedish (but not Icelandic: cf. Sigurðsson 1986) show a different asymmetry between the order of constituents in main clauses and subordinate clauses: in the latter the following structure is found:

(ii)   Conj   S   Sentential adverb   Finite V   Nonfinite V   O/complement
           Other adverbs

It is hardly worth nothing that the structural similarity between German and Dutch on the one hand and Swedish and Danish on the other is related to historical factors.

158. The example is taken from Haegeman (1991: 522).

159. The example is taken from Platzack (1986: 27). Cf. Pétursson (1978: 138) for contemporary Icelandic examples.

160. For this formula, cf. Haegeman (1991: 522). Cf. Dik (1980: 158) for a representation within the framework of his functionalist model.

161. These are not the only definitions of *Mittelfeld*, *Vorfeld*, *Nachfeld* in German grammatical tradition. For an overview of the history of these notions, see Etzensperger (1979).

162. A concise presentation of Engel's model, its deviations from other topological models, and the problems it poses can be seen in Etzensperger (1979: 194–203).

163. For an examination of the type of element which may occur in the *Vorfeld* or in the *Mittelfeld*, cf. Engel (1991: 310 ff., 320 ff. respectively). Some constituents may not occur in the *Vorfeld*, for example the *Abtönungspartikeln* and adverbs, such as *sehr* 'very', *außerordentlich* 'extraordinarily', which regularly occur as adjective modifiers (cf. Engel 1991: 312, 313). The expletive *es*, whose distribution is confined to the *Vorfeld*, may not occur in the *Mittelfeld* (cf. Engel 1991: 309).

164. Cf. Engel (1982: 223).

165. It has been observed that this constraint has an apparent exception. Sentences such as *Mit dem Ball ins Gesicht hat er mir geworfen* lit. 'with the ball into the face has he me thrown', 'Throw the ball in my face he did', *Gestern am Strand hat er mir ein Geheimnis gesagt* lit. 'Yesterday at the beach has he me a secret said', 'Yesterday on the beach, he told me a secret' seem to have two constituents in P_1. Scherpenisse (1986: 53–55) explains this structure as the effect of two conditions that do not violate the constraint of the single constituent in P_1: (a) the string in P_1 must be continuous; (b) it must be dominated by a unique node (which is not necessarily a maximal projection). Such structure is described by Scherpenisse as an "adverb cluster"; it is associated to a characteristic prosodic property, that is, the fall of a heavy stress on the rightmost constituent of the cluster. Other syntactic types with more than one constituent before V are less easy to analyze in that they cannot be reduced to constituents which are dominated by a single node. They are characterized by the occurrence of particular adverbs before V, namely, *allerdings* 'certainly', *sogar* 'even', *nur* 'only', *nicht* 'not': *Graf Pocci allerdings hat amüsante Sachen geschrieben* 'Earl Pocci certainly wrote amusing things', *Nicht das wollte ich eigentlich erzählen* 'It was not that I really wanted to tell' (cf. Tarvainen 1985: 382); however, these two examples seem to be different, as in the first NP and Adv are in fact discontinuous constituents, while in the latter, negation is merely an operator with scope on the constituent in P_1. Cf. Sigurðsson (1989: 5, n. 1) for the exclusion of structures with *ja* from these more problematic cases.

166. Engel (1991: 316) observes, however, than any "complement" which is a heavy constituent may occur in the *Nachfeld*.

167. For the constituents which may occur in the *Nachfeld*, cf. Engel (1991: 316–318). Pronominal complements and modal adverbials have a distribution which excludes the *Nachfeld* (cf. Engel 1991: 316).

168. This function is ably described by Weinrich (1993: 83). For frequency in the spoken language of structures with circumstantials of time and place, cf. also Engel (1991: 316).

169. For structures which require the *Rahmenbildung*, cf. Helbig and Buscha (1989: 567). See also the more recent description of German WO patterns in Zifonun et al. (1997: 1495–1680).

512   Rosanna Sornicola

170. Cf. here Section 3.
171. Cf. Haider (1986: 56), who observes that in colloquial German nominative and accusative pronouns may be omitted in A position. Haider shares Huang's (1984) opinion according to which an empty operator appears in position A in such cases, which implies that V is always in P_2. Cf. also Cardinaletti (1994).
172. An opinion held by researchers in various frameworks is that the development of obligatory pronouns goes hand in hand with the V-2 type (cf. Haiman 1974; Dik 1980: 157), but this view seems to suffer from an excessive teleologism.
173. Note, however, that according to Platzack (1986: 46, n. 2) in Swedish the phenomenon under examination is less common than in Icelandic.
174. Especially in coordinate clauses introduced by *und* (cf. Behaghel 1932: 30–37).
175. Cf. Platzack (1986: 47, n. 6).
176. Abraham (1986: 16) defines the configurational option as that in which "the underlying word-order as defined by grammatical criteria is subject to linear variation as required by textual organization" and "movement according to textual parameters results in marked linearization."
177. Cf. *Ich gab dem Mädchen ein Buch* 'I gave to the girl [+definite] a book [−definite]', *Ich gab das Buch einem Mädchen* 'I gave the book [+definite] to a girl [−definite]'.
178. According to Abraham, the PO argument, with basically locative value, "contributes the most to the meaning of the predicate and as such is the verb's closest argument" (Abraham 1986: 16; cf. also the summary on pp. 29–30). This description is not, however, very convincing: it cannot be upheld with strictly semantic criteria and may be regarded as circular.
179. Cf. Daneš (1967: 226, 228) and the discussion here in Sections 2.4.4.2 and 2.4.4.3.
180. Cf. *Karl hat dem Kind das Buch geschenkt, Karl hat dem KIND das Buch geschenkt, Karl hat dem Kind das BUCH geschenkt, Karl hat dem Kind das Buch GESCHENKT* (cf. Höhle 1982: 86, with examination of the various pragmatic conditions associated with such sequences). For the distinction between "lexical items" (N, Adj, V, Adv) and "grammatical items" (Art, Prep) with respect to stress, cf. Cruttenden (1986: 82). In any case, in semantic-pragmatic conditions of corrective contrast, constituents such as articles and prepositions may also attract nuclear stress. For the relationship between "normal" order and allocation of nuclear stress, cf. Höhle (1982: 85–86, 91–92), who examines the phenomenon within a textually-oriented approach.
181. Cf. Höhle (1982: 125–126).
182. This is the so-called intonation with "normal stress," which presents problems for a definition in prosodic terms: cf. Cruttenden (1986: 81, 94–95), who prefers to use the concept of "broad focus." Abraham (1995: 615) makes a distinction between "grammatikalischer (default) Fokusakzent" and "markierter oder kontrastiver (semantischer) Fokusakzent." Cf. also Jacobs' (1988: 120) distinction between "neutrale Akzentuierung" and "normale Akzentuierung."
183. Cf. Höhle (1982: 85–92), Abraham (1986: 17–18), and Jacobs (1988: 107), who assumes a principle of "FOCUS-last": "Wenn das MF (= Mittelfeld) genau ein F_i-Satzglied K enthält, so folgt K auf alle anderen Satzglieder." Counterexamples may be explained by the interaction of other principles (cf. Jacobs 1988: 107–108).

184. Cf. Abraham (1995: 616, 624 n. 58). Semantic and textual factors such as definiteness vs. indefiniteness of the head of NP also play a role in determining the FOCUS. For the intonation pattern with stress on O, see Cruttenden (1986: 148).

185. Note that in the written language, the neutral position of the temporal adverb would be immediately after V: *Karl liest heute ein Buch*.

186. This situation is found in other languages (for English, cf. Cruttenden 1986: 84; for Italian, cf. Sornicola 1993a), and seems to be connected with a general property of SVO languages.

187. Cf. Engel (1982: 229, 1991: 330, 332).

188. Cf. Engel (1991: 31).

189. On the relationship between constituent movement and stress in German, Engel (1991: 332) observes: "Meist ist die stellungsbedingte Hervorhebung zugleich von stärkerer Betonung begleitet; in solchen Fällen erweist sich die Intonation als Konsequenz der Stellung." Of course, highlighting may also occur with purely prosodic means, that is, a given constituent may keep its canonical position and convey the nuclear stress.

190. The example is taken from Scherpenisse (1986: 45). On the problem of left dislocation, cf. Altmann (1981, especially chs. 4–7) and Scherpenisse (1986: 55).

191. In neutral intonation, the sentential stress coincides with the verb.

192. Cf. Engel (1991: 316 ff., 332).

193. Cf. Scherpenisse (1986: 64–65), Abraham (1986, 1995: 606–607). Jacobs (1988: 108–109) has an interesting problematic formulation of the question.

194. Note that when the *Mittelfeld* contains a PrepP and a DO, the order PrepP + DO (with DO occupying position X_i) is possible only when DO is [+ definite]: cf. Abraham (1986: 18).

195. In sentences with *Satzklammer* which have non-neutral order, nuclear stress is linked to the immediately preverbal position. This conclusion is reached by Abraham for sentences with *Satzklammer* which have one of the following WO patterns:

    (i)   S IO DO V

    (ii)   S DO IO V

    (iii)   S DO PrepP V

    (iv)   S PrepP DO V

    The first and third patterns show an unmarked (more frequent) order, while the second and fourth show a marked (more rare) order. Now, precisely the first and third patterns also have a wider range of accentual options than that of the second and fourth patterns. The latter only admit a single option in which the nuclear stress falls on the constituent that is not in situ, which occurs immediately in front of V, as can be seen from examples (71a)–(71d). Being so defined, such a property does not seem to be related to the determination of a characteristic position for FOCUS.

196. For an overview of the problem, cf. Scaglione (1981). One of the most debated questions has been that of the influence of Latin patterns on German WO (a clear and succinct examination can be found in Wells 1987: 205–259). Rather than thinking of an influence of Classical Latin on German WO, it seems more convincing to accept

514   Rosanna Sornicola

Burdach's (1914) claim of an influence of Chancery Latin based on Humanistic Italian models (Fleischmann 1973 opposes this idea). In any case, as far as main clauses are concerned, studies carried out on texts show that in the oldest documents of the Germanic languages the dominant order was already SVO (cf. Braunmüller 1982: 138–139), which was, however, in competition with other orders. These fluctuations persist throughout the course of many centuries. The sixteenth century texts still exhibit WO structures and, more generally, syntactic structures different from those grammaticalized in the present-day phase (cf. Wells 1987: 253–254).

197. The structure with *Satzklammer* is already found in Old High German (in Notker), but is not grammaticalized. This situation still prevails in the sixteenth century (cf. Wells 1987: 258). Ebert's (1980) investigation of the Nuremberg texts has shown the importance of the role of the Chancery: while up to the fifteenth century, chancery texts differ only slightly from those of individual writers of the region with respect to the extent of *Satzklammer,* from the sixteenth century the divergence becomes greater. Ebert has shown a distinct influence of administrative usage on individual writers in Nuremberg. It is interesting that less educated writers tend to use the frame less frequently. Ebert hypothesizes that the *Satzklammer* had to a greater or lesser extent caught on in the spoken language of the educated circle comprising the city's administrators, without necessarily reaching the dialectal or colloquial spoken registers of Nuremberg. For the influence of the press on grammaticalization of the structure with *Satzklammer*, cf. Wells (1987: 259).

198. The situation with respect to Breton is more controversial, cf. Timm (1989, 1991).

199. For Irish, cf. Stenson (1981: 40–41), McCloskey (1983), Ó Siadhail (1989: 205–206). For Welsh, cf. MacCana (1973, 1991), Williams (1980), Rouveret (1990), Tallerman (1991). See also the contributions in Hendrik (1990) and Borsley and Roberts (1996).

200. For a critique of such models, cf. Anderson and Chung (1977) and the papers in Gazdar et al. (1983), in particular, on the possibility to postulate a VP phrase in VSO languages. On the other hand, recent proposals to eliminate the linear representation in the phrase structure rules of generative grammar may overcome such difficulties. The problem of the VP constituent in VSO languages has been recently re-examined by various contributions in Carnie and Guilfoyle (2000). It is clear that these difficulties do not arise in functional models, in which linearization rules are simply the realization of nonlinear functional representations (cf. Dik 1989, especially 333 ff.).

201. In Irish, no constituent may be placed between V and S (cf. Stenson 1981: 41) except for a small number of parenthetic expressions, such as *muise* 'indeed', *mh'anam* 'certainly' (cf. Ó Siadhail 1989: 217).

202. The example is taken from Ó Siadhail (1989: 205).

203. Cf. the statistics in Tomlin (1986: 17–22).

204. In this sense, Dik's (1980: 155 ff.) observations are to be treated with caution. Cf. Timm (1989, 1991: 278), Delanoy (1990) for Breton.

205. The example is taken from Stenson (1981: 40–41).

206. Cf. Stenson (1981: 40–41).

207. In fact V is embedded in a relative clause (cf. Stenson 1981: 40–41).

208. In both Irish and Welsh, the relative clause is generally introduced by a particle *a*

(of different origin in the two languages); in the oldest phases of Irish this particle was not present, but the V morphology had special properties: cf. Pedersen ([1909–1913] 1976: vol. II, 217–223), Lewis and Pedersen (1937: 236–245). The difficulties of analyzing the Celtic structures in question have been pointed out in different ways by various scholars: see, for example, MacCoisdealbha (1998: especially chs. 2 and 3), Shisha-Halevy (1995: 149–156, 2000: 80).

209.   Note that the pronoun is optional in type (73).

210.   For this structural type, cf. MacCana (1991: 73), Evans ([1964] 1989: 140–141). For Middle Welsh, see also Shisha-Halevy (1995: 153–157).

211.   On the problem regarding the structural representation of cleft sentences, cf. Sornicola (1988) and the bibliography included therein.

212.   For the Irish example, cf. Ó Siadhail (1989: 305); for the Welsh example, Watkins (1991: 345).

213.   This choice, which is not without potential criticism, could be justified by the fact that what counts here is the relative order of V and its arguments.

214.   For an examination of this problem, see the discussion in MacCana (1991: 60).

215.   For the V-initial type in ancient monuments of Indo-European languages, see Delbrück (1900: 58–61), Marouzeau (1938: 49 ff.), Schwyzer (1950: 689 ff.), Gonda (1952: 26–33), Dressler (1969), Verpoorten (1977: 38–46), Holland (1980: Ch. I). Gonda (1952: 72 ff.) maintains that V-initial position in Sanskrit has the function of a "mise en relief" as well as an affective value; in other terms, V-initial sentences are – in his opinion – structures with marked FOCUS. On the other hand, Dressler (1969: 3 ff.) considers V-initial position "normal" in most ancient Indo-European languages for sentences in narrative textual progressions. Dressler's hypothesis allows a consideration of V-initial sentences as a marked pattern in that it cannot be found in all textual environments (see here Section 1.6.1). As to the diachronic explanation of the Irish type, see Watkins (1963), Wagner (1959: 152 ff., 205 ff.) (who believes in the influence of a Hamito-Semitic substratum), and Wagner (1967). The V-initial structure of the noncontinental Celtic languages is unanimously considered an innovation (see MacCoisdealbha 1998: 1–2).

216.   The two types show structural similarities (in both the "fronted" constituent is followed by the relativized verb) and differences (in the AS the verb of the relative sentence agrees with the head of the "fronted" NP, while this is not the case in MS). Besides, MS is a marked sentence (in fact, its fronted constituent is a marked FOCUS), while AS is not. For a summary of the differences between the two types, see Fife and King (1991: 84–85). Note that MS and AS also show a different diachronic development: the first has always been documented as a marked order, while the latter was the prevailing unmarked structure of Middle Welsh (see MacCana 1991). On the problem of WO in Middle Welsh, see E. Poppe (1989, 1990, 1991a, 1991b: 14–34), Willis (1998: 7 ff., 51 ff.).

217.   For examples of constituents with a different GF in the preverbal position, see Willis (1998: 4).

218.   Cf. Willis (1998: 168); see also Williams (1980: 4–5) for a discussion of other syntactic differences concerning the form and the distribution of the negative marker.

516   Rosanna Sornicola

219. Cf. Williams (1980: 168), Willis (1998: 4). See also
     Shisha-Halevy (1995) for a different pragmatic interpretation based on Middle Welsh
     data.
220. Cf. Watkins (1991: 342), who points out that fronting of sentential adjuncts with the
     function of setting the scene is a recurring phenomenon in fairytales.
221. Position X of the post-field in the configuration VSOX is indeed typical of adjuncts in
     both Welsh (cf. Watkins 1991: 334) and in Irish (cf. Stenson 1981: 41–42; Ó Siadhail
     1989: 205–206).
222. Most specialists consider the MS to be always emphatic. Contrary to this view, Row-
     lands (1980: 220) considers that "emphasis is not a matter of degrees that vary with
     context" and that "MS need not be emphatic at all, though it is grammatically an em-
     phatic structure" (quoted in Fife and King 1991: 87).
223. The phenomenon has been described in Old Irish by MacCoisdealbha (1998: Ch. 4);
     cf. furthermore MacCana (1973: 100). For modern phases, see Stenson (1981: 47).
224. The example is taken from MacCana (1973: 98).
225. The example is reported by Watkins (1991: 334). For the form *mae*, cf. Pedersen (1976:
     vol. 2, 426).
226. The analysis in the literature of the syntactic and pragmatic properties of the types
     mentioned so far is not very clear. The distinction between the cleft type (76b) and
     noncleft types (cf. (78)) is defended by Watkins (1991: 332 ff.) for Welsh. In the sec-
     ond group he places both the AS and the MS. According to Watkins, if one leaves
     aside certain structural differences and differences in the semantic/pragmatic value,
     the two types may be grouped together by important properties, relating to subordi-
     nation, negation, and some aspects of agreement (Watkins 1991: 332). Watkins main-
     tains that the most characteristic property of the cleft type is the relativization marker
     which precedes V (Watkins 1991: 341), even though a multiplicity of factors may
     make recognition of this problematic: for example, there is a difference between spo-
     ken and written registers with respect to the relative pronoun *a*, which is regularly
     cancelled in spoken Welsh, present in literary prose and only occasionally cancelled
     in informal written language (Watkins 1991: 334). Moreover, even when deletion has
     been ascertained the soft mutation of the initial V segment is unequivocable proof of
     the underlying presence of the pronoun (Watkins 1991: 334, n. 8): for discussion of
     further problems, cf. Watkins (1991: 335–341). Similar to Watkins' proposal is that
     of MacCana (1991: 62 ff.) who arrives at similar descriptive results, although his ex-
     amination of the semantic and pragmatic functions of fronting phenomena seems to
     be more articulated. Contrary to Watkins and MacCana, however, Fife and King claim
     that there is no single cleft type in Welsh, but that a distinction should be made be-
     tween AS and MS on the basis of their respective syntactic and distributional proper-
     ties (Fife and King 1991: 84–85): in short, AS may represent a topicalization structure,
     in which the fronted constituent has all the characteristic functions of a TOPIC, both
     sentential (those of adding salience to the TOPIC, of establishing the theme of the
     sentence, of promoting informative progression) and discoursive (those of giving co-
     hesion to the discourse, of TOPIC shift, and the stylistic function of suspense building
     devices, anticlimax strategies, etc.; cf. Fife and King 1991: 132). The MS instead is

probably structurally and functionally a cleft sentence, its primary pragmatic function being in fact that of contrast. However, the two authors' hypothesis seems to suffer from two kinds of problems: direct resort to universal pragmatic processes to explain two historically determined types; a none too clear definition of the very pragmatic processes resorted to. If fronting in the MS is associated with a phenomenon of contrastive FOCUS, fronting for topicalization as in the AS refers just as equally to a "focused structure" (cf. Fife and King 1991: 122: "topicality is one sort of focusing device"). The two authors had started out with the observation that "the information structuring concept most relevant to the Welsh data is what we call focus. This term refers to the degree to which an item is made more prominent in the communicative structure of the sentence" and add "one very common form of marked focus is what the literature calls topicalization" (Fife and King 1991: 95). For a criticism to Fife's views, cf. Shisha-Halevy (1995: 152).

227. The example is taken from MacCana (1973: 106).

228. The example is taken from MacCana (1973: 106).

229. The example is taken from Lewis (1942: 20).

230. The hypothesis that it may be a survivor of the AS had been put forward by Lewis (1942). This idea has been convincingly refuted by Watkins (1991: 331) and by MacCana (1991: 47).

231. For Irish, cf. MacCana (1973: 106–110, 1991: 70–71).

232. *Sy'n* is the contracted form of *sydd* 'who is (are)'+ *yn*, the particle preceding the verbal noun in the progressive construction.

233. MacCana (1991: 71) speaks of "a common functional element that exists between the several subcategories of relativized V-second sentence," but does not make the nature of such an element any clearer.

234. This last point has quite rightly been emphasized by Fife and King (1991: in particular 142–143).

235. MacCoisdealbha (1998: 111) had held that topics constitute an intermediate category between proper contrast and simple intraclausal thematic structuration, such as passivization. For a view of TOPIC and FOCUS as a graded scale, cf. Sornicola (1993a).

236. Jongeling (1991: 107) considers VSO the basic WO of Classical Hebrew, while SVO is the order conveying emphasis or contrast. However, Modern Hebrew has SVO as basic order (see Glinert 1989: 413). In fact, the change from VSO to SVO can already be observed in the internal diachrony of Biblical Hebrew: see Givón (1977b). Modern Standard Arabic has more than one dominant order, though V S O can be considered basic (cf. Agius 1991; Ingham 1991 for dialectal varieties). On basic VSO order of Polynesian languages and their competing non-V-initial orders which are due to topicalization processes, cf. Chapin (1978) (for the Easter Island language), Anderson and Chung (1977) (for Samoan and Tongan). Jongeling's work, which also typologically compares Semitic languages and Welsh, has sound critical observations on Tomlin's ideas: "The most important objection is that Tomlin does not explain why there are still V S O languages left" (Jongeling 1991: 110, n. 29).

237. Regarding this, see the observations of other authors who have dealt with VSO lan-

518   Rosanna Sornicola

      guages: cf. Anderson and Chung (1977: 23). Jongeling (1991: 106) and Ingham (1991: 727) are of the same opinion.

238.   The morphology of the verb exhibits a first person singular suffix, the etymology of which is a first person singular pronominal clitic.

239.   Cf. MacCana (1973: 117).

240.   Except for the present and imperfect of *bod* 'be', which are instead preceded by *y* (*i*) (at least in literary Welsh): cf. Watkins (1991: 340 and n. 19). On the use of *ef* (*e*) as a referential pronoun and as an expletive, cf. Willis (1998: 175). For the forms *fe*, *fo* see ibidem.

241.   Brockelmann (1913: 171) makes explicit reference to this phenomenon, pointing out that contemporary speakers of Semitic languages are no longer aware of it.

242.   In all the Hamito-Semitic languages, with the exception of Egyptian, an extremely archaic stem *ya is attested in the third person masculine singular prefix of the prefix conjugation, having a personal and determinative value (cf. Garbini 1984: 81 and the bibliography ibidem n. 1). Similarly, the prefix *t-* represents a reduced form of the stem of the second singular personal pronoun: cf. Garbini (1984: 224).

243.   Cf. Anderson and Chung (1977: 7–9) for a more detailed discussion. The "clitic placement" rule which they postulate is a typical synchronic rule formulated in terms of descriptive economy: "First, a clitic pronoun copy of a pronominal S is created and positioned before the verb. Then, if the pronoun subject is unemphatic, it is deleted, leaving the clitic copy as the only residue of the original subject." Such a rule may have no significance for the formulation of diachronic hypotheses.

244.   In the Hebrew of Early Biblical poetry, the conjugation with suffixes is already well established as a means of expressing the past; it is the form most used in narrative progressions and usually expresses punctual aspect, connected to an action already completed, while the conjugation with prefixes usually expresses durative aspect, connected to an action which the speaker considers to be incomplete (cf. Sáenz-Badillos 1993: 58–59). In Late Biblical Hebrew this distinction has already been grammaticalized in the verbal system, with the suffixal conjugation used only for the past and the prefixal conjugation only for the future (cf. Sáenz-Badillos 1993: 129). For the extent to which aspectual and pragmatic factors affect the VS order in the diachronic development of Hebrew, cf. also Givón (1977a).

245.   See also further on, Daneš (1967: 218–219, 223–224).

246.   I refer to the fact that even verbs in passive or pseudo-intransitive constructions may belong to the situation being described. Now, while an intransitive verb is inherently so (the property is represented in the subcategorization frame of the lexeme), a passive, reflexive, or pseudo-intransitive is intransitive by derivation (the property has nothing to do with the subcategorization frame of the lexeme).

247.   Cf. for Russian *vernulsja otec*, lit. 'is returned father' 'father has returned'; Italian *se ne è andato il gatto*, lit. 'Refl.Prn.2sG from there is gone the cat', 'the cat has left'.

248.   For an examination of this typological resemblance between the Slavonic and Romance languages, cf. Cennamo (1993).

249.   The literature is vast and I will limit myself here to reference to Ulrich (1985: ch. 1); for an examination of the history of the concept "thetic" in the German linguistic

tradition, cf. also Haberland (1993) and Sasse (1987). For the concept of "presentative", cf. Hetzron (1971, 1975). There are numerous papers dedicated to examination of the structures in questions: cf. Lambrecht (1994: 39–43, 177–181) with particular reference to the Romance languages and English; Bernini (1995) for Italian; Abraham (1995: 626–627) for German; Erdmann (1990) for English; Kirsner (1979) for Dutch; Fougeron (1989: ch. 5 and 6) for Russian.

250. The literature on this structure is also vast. For a general view of the problem, cf. Lyons (1977: 723) and especially Sasse (1987, 1995a). See Euler (1982) for a comparative analysis of this structure in ancient Indo-European languages; for modern Indo-European languages, see Ulrich (1985: 82–114). Interesting observations on the function of this structure in spoken Italian can be found in Berruto (1986) and Bernini (1991, 1995). In many Indo-European languages, locative-existential structures may be used as temporal circumstantials that function as introducers of referents in narrative texts, for example fables: cf. Yokoyama (1986: 228–229) for Russian. It is not unusual for presentative and locative-existential structures to be interchangeable: cf. Gadet (1992: 71, 76–77), Lambrecht (1994: 169–170).

251. Of course a separate discussion is required for languages in which the structural elements are ((Adv loc), 'have', O), such as in the Spanish structures *hay* + NP and French structures *il y a* + NP, where O appears in postverbal position.

252. The example is quoted from Yokoyama (1986: 282).

253. The idea of different types of intransitive verbs was originally presented by Perlmutter (1978).

254. Verbs like *arrivare* require 'be'-auxiliary and their single NP can undergo *ne*-cliticization, while verbs like *parlare* require 'have'-auxiliary and their single NP cannot undergo *ne*-cliticization (cf. Burzio 1986: 16 ff.; and more recently, Graffi 1994: 235–237, 275–277). For the terminological problems, see Haegeman (1991: 298–312) and Graffi (1994: 236 ff.).

255. For a discussion of this problem, cf. Sornicola (1994).

256. Such a probability applies in spoken rather than written texts: cf. Sornicola (1995b) for an examination relating to Italian and Spanish.

257. In some verbs in Italian, such as *bastare* 'to suffice', *servire* 'to serve', *interessare* 'to interest', etc., the NP which controls agreement occurs consistently in O position: cf. Sornicola (1990).

258. Cf. Sornicola (1994: 36–37).

259. Cf. Comrie (1981b: Sect. 6.2).

260. Morphological incorporation of the nominal root with the feature [−animate] to the verbal root in functionally eventive structures is a phenomenon which is found in some native American languages (cf. Sasse 1987).

261. Cf. Gonda (1952: 73) and the bibliography quoted therein.

262. Cf. Fougeron (1989: 260, 305 ff.). Note that in such a case, the verb is always in the past tense.

263. Cf. W. K. Matthews (1960: 51) and Yokoyama (1986: 284 ff.). See also here Section 4.1.

264. Cf. Hakulinen (1961: 318–319).

520   Rosanna Sornicola

265. Note, however, that structures with one-argument verbs are in the majority in Mac-Cana's (1973) data.

266. Cf. Fougeron (1989: 106–140), from where all Russian examples quoted here are taken. The apostrophe indicates that the constituent carries nuclear stress. The VS pattern has a characteristic intonational profile: basic stress is strongest when it falls on the S constituent (cf. Fougeron 1989: 150). Furthermore "le palier mélodique, situé dans la zone des fréquences élevées et accompagné d'une quasi-stabilité de l'intensité (relativement forte), crée une sorte de 'retardement de resolution' et maintient une certaine tension" (Fougeron 1989: 151).

267. An interesting subclass of eventive structures characteristically occurring with the VS pattern is that of "meteorological verbs"; for further examples from other languages, see here Section 3.2.1.1.

268. Cf. Fougeron (1989: 138). Fougeron observes that the combination of the syntactic pattern and the prosodic pattern contribute to the informative balance of the two constituents, that is, neither prevails in the structure of the sentence.

269. In these styles the variants V' / S, with S in TAIL position (afterthought), are allowed.

270. In this context, a structure with the same WO but with neutral prosodic contour, that is, *Papà è tornato*, would also be possible. In any case, as far as the comparison between Russian and Italian is concerned, it should be observed that all the structures described by Fougeron are from her corpus of spoken language, which doesn't exclude the fact that in Russian, there may also be oscillation between SV and VS. Yokoyama's data, in fact, give reason to think that such an oscillation exists (cf. Yokoyama 1986: 184–194).

271. Cf. Lerch (1934: 258).

272. In the Germanic languages, VS is allowed only if V does not occupy position P_1 (cf. Section 2.4.5). But cf. Section 3.6.3 for examples of VS order with intransitive V in P_1 in colloquial and dialectal Modern German. The structures belonging to this group have been described by Bossong (1984, 1987) under the label of "paradigmatic inversion."

273. The example is taken from a fable text, reported by Yokoyama (1986: 285–286). It is a narrative sequence.

274. The example is taken from Yokoyama (1986: 184).

275. Cf. Sornicola (1995b).

276. The example is taken from the French translation of H. F. Ellenberger's, *History of the Discovery of Unconscious* (Ellenberger 1994: 187).

277. *There* is often described as the "grammatical" S in present-day grammars. Obviously, however, the whole structure has a split of syntactic properties of subjecthood between *there* and the postverbal NP.

278. The examples are from Quirk et al. (1985: 1408).

279. Quirk et al. (1985: 1408).

280. For Estonian, cf. Tauli (1983: 159), who observes that when the structure comprises the constituents S, V, and Adv, "all possible combinations occur, but with different stress and frequency."

281. Cf. Karlsson (1987: 77 ff.).

282. Examples (38) and (39) are from Fromm (1982: 143).

283. This example is from Hakulinen (1961: 320). These types of inversion are probably inherited in Finnish (the example with the verb of saying is also found in dialectal varieties) and probably differ from the type of inversion due to interference from Swedish (cf. Hakulinen 1961: 318–319). As to the stress pattern, nuclear stress is on S (cf. Hakulinen 1961: 319). On the problem of patterns of order in existential sentences, cf. also Vähämäki (1984: 446), who claims that statistically existential sentences in Finnish have the pattern XVS. Vilkuna (1989: 10) records cases of V-initial with intransitive verbs of movement in the translation of Russian fables and considers them to be exceptional, although possible in particular styles.

284. Both examples are taken from Fife and King (1991: 145–146). Cf. also MacCana (1991: 67) who claims that SV order appears in the presence of "expressive force" distributed over the entire sentence.

285. Cf. Watkins (1991: 341); the locative adjunct occurs in P_1 in the structure.

286. Cf. Ó Siadhail (1989: 217).

287. All the Basque examples are taken from Lafitte (1962: 47–48).

288. Cf. Erguvanli (1984: 44–45).

289. Cf. Erguvanli (1984: 47).

290. Cf. Erguvanli (1984: 17).

291. Cf. Erguvanli (1984: 17).

292. Examples (140)–(143) are taken from Erguvanli (1984: 15–16).

293. Cf. Erguvanli (1984: 6 and 69, n. 3).

294. Cf. Erguvanli (1984: 44, 68–69).

295. Erguvanli (1984: 129).

296. Cf. Erguvanli (1984: 130), with an interesting discussion on the relationship between stress and definiteness.

297. Cf. also Sasse (1995b) and Sasse (this volume).

298. Cf. Behrens (1989: 117), who does not include these verbs in the same group as the preceding ones, since she claims that SV/VS oscillation in this case is essentially due to textual factors of dialogue progression. In VS order S is, however, always rhematic. Furthermore, Behrens observes that in cataphoric uses of S in modern Hungarian, S goes in front of V, while it is postposed to V when preceded by quotation of a direct discourse.

299. Cf. Behrens' (1989: 116–142) examination.

300. In Behrens' data, the percentage of structures with focalization (of the entire sentence or one of its constituents) varies according to the period and to texts.

301. Cf. Behrens (1989: 116). However, in Old Hungarian S [−definite] and [−referential] could also occur in postverbal position. The picture is further complicated by the fact that in these old diachronic phases, preverbal S [−referential] always occurs in the characteristic immediately preverbal position.

302. This is very clear from the fact that, in these languages, it is equally possible to identify a "dominant" order: cf. the discussion on Russian in Section 2.4.3.2.

303. Cf. Sornicola et al. (1994). This work, however, does not present a wide statistical base like Uhlířová's.

304. The choice of the auxiliary 'to be' or 'to have' is a parameter which may undergo

522   Rosanna Sornicola

considerable alteration in the diachronic development of a language (for English, cf. Rydén and Brorström 1987); it is, furthermore, too closely linked to specific characteristics of individual groups of languages (languages such as Chinese which do not have auxiliaries of this type come to mind). Even more restricted typologically is the diffusion of structures with *ne*.

305.   Cf. Keenan (1976); Plank (1984); and the papers in Plank (ed.) (1984).

306.   The situation of VSO languages is, in a way, uninteresting in this respect: see further on in this chapter.

307.   Albeit with differences between languages with more flexible orders and languages with more rigid orders, as has been said in Section 2.

308.   *Maan* is a Thuringian form for *Mann*.

309.   Examples (149) and (150) are from Behaghel (1932: 38), who quotes (149) as possible in his own speech. Interestingly, he also quotes examples with VS order for structures with a transitive verb, from both colloquial and dialectal German varieties. The two structures may convey a value of surprise (I owe this interpretation to Erich Poppe; his judgement is that they are both non-neutral sentences). Example (151) is a structure described by Russ (1989: 258) for the dialect of Palatinate.

310.   Cf. Hakulinen (1961: 319–320). The inversion due to the Swedish influence is structurally similar to that of Modern Standard Germanic languages, that is, it respects the V-2 constraint.

311.   This is presumably a universal tendency, since in many of the world's languages unmarked narrow FOCUS tends to coincide with a referential constituent. This seems to be true not only of FOCUS as a PF, but also of its prosodic correlates (see Ladd 1996: 188–189). However, as far as the latter are concerned, it does seem that determining the canonical allocation of FOCUS as stress is more complex: see Cruttenden (1986: 147–150). Ladd (1996: 167 ff.) has an interesting discussion on the independence of prosodic structure from the structure of PFs and on its typological variation across languages.

312.   Proof of this is supplied by numerous data: cf. Hakulinen (1961: 315) for Finnish, Kenesei (1986: 144) for Hungarian, and Erguvanli (1984: 60) for Turkish. For Italian, see Section 2.4.4.1.1.

313.   Cf. Vilkuna (1989: 10). Similar considerations apply to Italian (cf. the percentages on initial V reported in Section 2.4.4.1.1).

314.   Behrens' (1989: 118–142) data on VS structures in Hungarian of the fifteenth and nineteenth centuries may also serve as indirect proof of this. VS structures with V in FOCUS form less than 10 % of the total VS structures of the corpus. These data are to be treated with caution since they concern the relative order of S and V, not the absolute position of V in the sentence.

315.   More rarely, in certain contexts S in the post-field may have a value other than that of adding information, a similar situation to that described by Erguvanli (1984: 60–62) for Turkish.

316.   Regarding these structures, as well as those described in Section 2.4.4.1.1, it should be pointed out that structures with pronominal objects proclitic to V are preferable to these; that is, respectively *Lo AMA Mario / Lucia* and *Lo AMA / Lucia / Mario*.

317. This function can be seen in Russian since the earliest documentation (cf. Gonda 1952: 73; W. K. Matthews 1960: 51).
318. The fact that speakers may perceive prosodic prominence on constituents which is not confirmed by experimental analysis is a well-known phenomenon in experimental phonetics.
319. Cf. Erguvanli (1984: 61), who observes that the unmarked order would be *İşte ben bu işin böyle sıkışıklığa geleceğini bildim*.
320. The example is taken from Watkins (1991: 345). The neutral (unmarked) structure would be *Roeddwn i yn edrych ar y gannwyll* 'I was looking at the candle' (I owe this example to Erich Poppe, pers. comm.).
321. Watkins (1991: 344) observes that the type is the usual option in colloquial Welsh for explanation of the action carried out.
322. The example and the description of the structural and pragmatic characteristics of this type are from MacCana (1973: 110).
323. According to a sample of native Russian speakers, it would be impossible to introduce a pause between the constituents (cf. Fougeron 1989: 310).
324. The example is taken from Yokoyama (1986: 272).
325. Yokoyama (1986: 273).
326. For a study of the influence of the feature animacy on WO in Russian, cf. Maslova (1995).
327. Erguvanli (1984: 44) criticizes the structural conditions set down by Hankamer (1971) for the extraposition rule which should generate structures with the verb in nonfinal position in Turkish.
328. Cf. Erguvanli (1984: 33–34).
329. On the definition of "definite," cf. Erguvanli (1984: 17–18).
330. Erguvanli (1984: 33).
331. Cf. Erguvanli (1984: 15–16).
332. Cf. Erguvanli (1984: 28–29).
333. The prosodic properties of the second sentence are also interesting: after *bu sürpriz* there must be a pause (intonation break): cf. Erguvanli (1984: 29).
334. Whether this may be considered a real phenomenon of incorporation in Turkish has been discussed but seems to be highly controversial: for a summary of the arguments for and against this idea, cf. Erguvanli (1984: 24–26).
335. See the examples reported on p. 22 and also the discussion on p. 28 (Erguvanli 1984). On the whole question of the effect of semantic features on word order in Turkish, cf. also Erguvanli-Taylan (1987) with discussion of additional data.
336. Cf. Kenesei (1986: 145).
337. Behrens (1989: 142).
338. Cf. Behrens (1989: 142).
339. Behrens (1989: 14).
340. Behrens (1989: 142).
341. I shall limit myself to mention of the Praguean tradition (cf. Firbas 1992; Sornicola and Svoboda 1992) and the European and American traditions influenced by it: cf. Dik (1989); Chafe (1994).

524   Rosanna Sornicola

342.   These concepts are taken from Fougeron's (1989: 300–301) study of Russian.

343.   Fougeron provides data from which it can be inferred that the final or nonfinal position of the constituent bearing the nuclear stress (which, note, may have any GF) is decisive in the orientation of the sentence, respectively rightwards or leftwards (cf. Fougeron 1989: 300–301, 362–367). Fougeron also observes that sentences such as *Juru Serëža čitaet* are linked to the left context. The description of these phenomena in Russian given by Yokoyama (1986: especially 190–191) is different. The author, in fact, establishes a relationship between the linear parameters of WO, those concerned with stress, and the informational parameters GIVEN/NEW. The distribution GIVEN/NEW is in effect related to the type of intonation: with intonation type I, without sentence stress, there is a progression from GIVEN to NEW (cf. example (26a) reported in Section 2.4.3.2), while with intonation type II, in which there is sentence stress, there may be a progression from NEW to GIVEN (cf. example (26b) in Section 2.4.3.2). The situation in which the NEW constituent occurs between two GIVEN constituents with intonation type II (cf. example (26c) in Section 2.4.3.2) is worth mentioning. Such a case would be more problematic in the Romance and Germanic languages.

344.   In Figure 21, the dotted line represents the right-hand context of the sentence. In Figure 22, the dotted line represents the left-hand context of the sentence.

345.   This is the so-called "inverse" or "artificial" order variously examined in the functionalist tradition. The inverse order is always accompanied by nuclear stress falling on the NEW constituent in P_1: for Russian, cf. Yokoyama (1986: 190–191).

346.   Structures with full NPs (which would, of course, be left dislocations of O) such as French *Le café, c'est maman qui le veut* and Italian *Il caffè lo vuole la mamma* would not be impossible in spontaneous speech with its redundant features; however, with respect to the context quoted for the Russian example sentences (171)–(172) are far more acceptable.

347.   The same property is encountered in Basque (cf. Lafitte [1944] 1978: 46–48) and in Cheremis (Mari) (cf. Lewy 1922: 168).

348.   For a discussion of the change of order from Latin to Romance languages, cf. Sornicola (1995a, 2000).

349.   Leaving aside, for simplicity, situations such as those in the Germanic languages in which the sentence is closed by a particle.

350.   Due to their greater freedom of constituent order, Finnish and Russian show certain peculiarities in this respect, which have been discussed in Section 2.4.3.

351.   The weak leftward delimitation of the sentence domain in these languages is related to the fact that no true dislocation structures (i.e., structures with the dislocated constituent anaphorically resumed in the "body" of the sentence) may occur. Thus, it is perhaps possible to envisage an "open" preverbal field.

352.   Note that such a representation has been claimed for similar S/O initial structures in Semitic languages such as Arabic (cf. Holes 1995: 203 ff.) and Hebrew (cf. Jongeling 1991: 107 ff.). However, the difficulties discussed in Section 2.4.6.1 must be borne in mind about the determination of position P_1 in VSO languages. It would, in fact, be quite reasonable to believe that the constituents with the GF S/O that occur before V occupy not P_1 but rather an extrasentential position.

353. This generalization does not apply to the Celtic languages, although – as has been seen – even in these languages topicalization processes may exist which place S or O in P_1 in neutral sentences.

354. In addition to the literature quoted in Section 1.6.2, cf. also the interesting criticism by Marantz (1984) within the generative framework.

355. Cf. Abraham (1995: 606–607).

356. This, of course, does not apply to Russian which, as was said in Section 2.4.3.2, does not have a unique position for FOCUS in neutral sentences.

357. Cf. CLG (1916: 172).

358. Cf. Sornicola (1986), Cinque (1993), Ladd (1996).

359. Cf. Watters (1979), Horvath (1986: 124 ff.). Note that Aghem is described as a language with rigid constituent order, which leads one to think that such a parameter contributes to determining the property under examination. Although Basque and Turkish are not languages with rigid constituent order, like many agglutinating languages, however, they present phenomena with relative rigidity of constituents.

360. The structure under examination has, in fact, very rare counterparts in SVO languages. The situation in Aghem, an SVO language with FOCUS fixed in the immediately postverbal position, appears to be rather peculiar.

361. For a new start of these old trends of research, see, for example, Gruber (1967) and the contributions in Li (1976).

362. For Korean, cf. Sohn (1994: 86, 209–210), who notes, however, that the "complement must always immediately precede the verb." For Kannada, cf. Sridhar (1990: 137). The case of Japanese and Quechua will be examined in more detail in Section 6.2.4.1.2.

363. If the description given by Tiffou and Pesot (1989: 70–71) can be relied upon – and it is not at all clear that it can be – the two authors report examples of the constituent with the GF ERG in sentence-final position with the value of FOCUS.

364. Cf. Sridhar (1990: 138) for Kannada. For Somali, examples (10c) and (10f) from Svolacchia et al. (1995) show the same property. Sohn's (1994: 195) description of Korean, while on the one hand confirming that the postverbal position is reserved for afterthoughts, on the other hand shows that such a position may also be the locus of "emphatic" constituents; the relevant examples, however, seem dubious and the same reservations may be had as with respect to certain descriptions of Japanese (cf. below).

365. Cf. Sohn (1994: 232) for Korean; he describes a multiplicity of possible positions for FOCUS in the proper sentence domain, a situation similar to that in Japanese. For Burushaski, cf. Tiffou and Pesot (1989: 70–71).

366. Cf. Sridhar (1990: 138–139). He notes, moreover, that "the initial placement of the object constituent is only for discourse continuity (old information) and not for emphasis" (Sridhar 1990: 138). However, the examples quoted by him seem to point at a scenario in which position P_1 could be filled by an IO with FOCUS function (cf. ex. (510)), but not by an O with such PF.

367. Cf. Hinds (1986: 148–150).

368. Cf. Hinds' (1986: 150, 166) examples.

369. Calvo Pérez denotes with "'" the TOPIC and with "''" the FOCUS (in his terminology, "el validador").

526   Rosanna Sornicola

370.   Cf. Calvo Pérez (1993: 41).

371.   Weber quite rightly distinguishes between the concepts of NEW or rhematic element and FOCUS. He observes that in some studies on other varieties of Quechua, as opposed to that which he investigated, the particles *-mi* and *-shi* have been identified as markers of the FOCUS element; according to him this is not always true in Huallaga Quechua (Weber 1989: 427).

372.   Cf. Weber (1989: 408–409, 419).

373.   Weber (1989: 429). Cf. also Muysken (1995: 385), who claims that Weber's formula expresses more than a frequency pattern.

374.   The problem has already been mentioned in Section 2.2 (see also the bibliography cited therein and Poppe [1954: 173] for Mongolian).

375.   In Japanese, the interrogative element generally occupies the same position in the sentence as it does in the corresponding declarative (with the only exception of interrogative sentences that have undergone a process of clefting): cf. Hinds (1986: 39–40). In Quechua, the interrogative element is usually (but not always) fronted to sentence-initial position (cf. Weber 1989: 19). In Somali, as far as may be deduced from Svolacchia et al.'s (1995: 73–75) examples, it is in one of the positions in the preverbal space.

376.   Cf. Sridhar (1990: 8).

377.   Cf. Weber (1989: 327–328).

378.   For an examination of this situation and the few exceptions to it due to models borrowed from Persian, cf. Erguvanli (1984: 72–117).

379.   Matthews (1991: 216), who cites the example reported here, has called attention to the significance of such a process in a typological framework.

380.   For Altaic languages, cf. Comrie (1981a: 77), who moreover observes that within this family, the validity of the parameter under examination has been verified in detail only for certain Turkic languages. For
        Indo-Aryan languages, cf. Section 2.2 and the references therein.

References

Abraham, Werner
    1986   Word-order in the middle field of the German sentence. In *Topic, Focus and Configurationality*, Werner Abraham and Sjaak de Meij (eds.), 15–38. Amsterdam and Philadelphia: John Benjamins.

    1995   *Deutsche Syntax im Sprachvergleich. Grundlegung einer typologischen Syntax des Deutschen*. Tübingen: Narr.

Adamec, Přemysl
    1966   *Porjadok slov v sovremennom russkom jazyke*. Prague: Academia.

Adams, Jim N.
    1976   A typological approach to Latin word-order. *Indogermanische Forschungen* 81: 70–99.

Agius, Dionisius A.
    1991   Precedence of VOS over VSO in Modern Standard Arabic. In *Semitic Studies in Honour of W. Leslau on the Occasion of his 85th Birthday*, Vol. 1, Alan S. Kaye (ed.), 39–55. Wiesbaden: Harrassowitz.

Allan, Robin, Philip Holmes, and Thomas Lundskaer-Nielsen
   1995   *Danish: A Comprehensive Grammar*. London: Routledge.

Altmann, Hans
   1981   *Formen der "Herausstellung" im Deutschen*. Tübingen: Niemeyer.

Anderson, Stephen R., and Sandra Chung
   1977   On grammatical relations and clause structure in verb-initial languages. In *Grammatical Relations*, (Syntax and Semantics 8), Peter Cole and Jerrold M. Sadock (eds.), 1–25. New York: Academic Press.

Aronson, Howard I.
   1990   *Georgian. A Reading Grammar*. Columbus, OH: Slavica Publishers.

Austerlitz, Robert
   1970   Agglutination in Northern Eurasia in perspective. In *Studies in General and Oriental Linguistics Presented to Shirô Hattori on the Occasion of his Sixtieth Birthday*, Roman Jakobson and Shigeo Kawamoto (eds.), 1–5. Tokyo: TEC.

AWL = Moseley, Christopher, and R. E. Asher (eds.)
   1994   *Atlas of the World's Languages*. London: Routledge.

Babby, Leonard H.
   1978   Word-order, case and grammatical relations in Russian. *Folia Slavica* 2: 35–45.

Banfi, Emanuele
   1998   Tra pragmatica e sintassi: alcuni aspetti della frase interrogativa nel turco (di Turchia). In *Atti del XXX Congresso Internazionale della Società di Linguistica Italiana (Pavia 26–28 settembre 1996)*, Paolo Ramat and Elisa Roma (eds.), 385–395. Rome: Bulzoni.

Barnes, Betsy K.
   1985   *The Pragmatics of Left Detachment in Spoken Standard French*. Amsterdam: John Benjamins.

Beaulieux, Léon
   1933   *Grammaire de la langue bulgare*. Paris: Champion.

Behaghel, Otto
   1932   *Deutsche Syntax, Vol. IV: Wortstellung, Periodenbau*. Heidelberg: Winters.

Behrens, Leila
   1989   *Textanalyse und Sprachwandel. Historische Syntax des Ungarischen: ein empirisches Modell*. Frankfurt a. M.: Peter Lang.

Bell, Alan
   1978   Language samples. In *Universals of Human Language, Vol. I: Method and Theory*, Joseph H. Greenberg, Charles A. Ferguson, and Edith Moravcsik (eds.), 123–156. Stanford, CA: Stanford University Press.

Belletti, Adriana
   1988   The case of unaccusatives. *Linguistic Inquiry* 19 (1): 1–34.

Benincà, Paola
   1988   L'ordine degli elementi nella frase e le costruzioni marcate. In *Grande grammatica italiana di consultazione, Vol. I: La frase. I sintagmi nominale e preposizionale*, Lorenzo Renzi (ed.), 115–225. Bologna: Il Mulino.

Benoist, Jean-Pierre
   1979   *Les fonctions de l'ordre des mots en russe moderne. Romans et nouvelles de Gorki*. Paris: Institut d'Études Slaves.
   1988   Remarques sur l'ordre des mots "grammatical" en russe moderne. Place du determinant accordé. Étude diachronique et synchronique. *Revue d'Études Slaves* 60 (3): 649–656.

Bernini, Giuliano
   1991   Frasi relative nel parlato colloquiale. In *La lingua degli studenti universitari*, Cristina Lavinio and Alberto A. Sobrero (eds.), 165–187. Florence: La Nuova Italia.

528   Rosanna Sornicola

1995   Verb-subject order in Italian: an investigation of short announcements and telecast news. *Sprachtypologie und Universalienforschung* 48 (1/2): 44–71.

Berretta, Monica
1998   Valori pragmatici diversi dell'ordine O V (O V S / O S V) nell'italiano contemporaneo. In *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza (Palermo, 18–24 settembre 1995)*, Vol. 2, Giovanni Ruffino (ed.), 81–90. Tübingen: Niemeyer.

Berruto, Gaetano
1986   Le dislocazioni a destra in italiano. In *Tema-rema in italiano*, Harro Stammerjohann (ed.), 55–69. Tübingen: Narr.

Bese, L., L. Dezső, and J. Gulya
1970   On the syntactic typology of the Uralic and Altaic languages. In *Theoretical Problems of Typology and the Northern Eurasian Languages*, L. Dezső and P. Haidú (eds.), 113–128. Amsterdam: Grüner.

Bhatt, Rakesh Mohan
1994   Structure of the Kashmiri clause. In *Theoretical Perspectives on Word-order in South Asian Languages*, Miriam Butt, Tracy Holloway King, and Gillian Ramchand (eds.), 31–65. Stanford, CA: Center for the Study of Language and Information.

Blanche-Benveniste, Claire
1990   *Le français parlé. Études grammaticales*. Paris: Éditions du CNRS.

Blatt, Franz
1957   Latin influence on European syntax. In *The Classical Pattern of Modern Western Civilization: Language*, (Acta Congressus Madvigiani, Travaux du Cercle Linguistique de Copenhagen 11), Franz Blatt (ed.), 33–69. Copenhagen: E. Munksgaard.

Boeder, Winfried
1989   Verbal person marking, noun phrase and word-order in Georgian. In *Configurationality. The Typology of Asymmetries*, László Marácz and Pieter Muysken (eds.), 159–184. Dordrecht: Foris.

Borsley, Robert D., and Ian Roberts (eds.)
1996   *The Syntax of the Celtic Languages*. Cambridge: Cambridge University Press.

Bossong, Georg
1981   Séquence et visée. L'expression positionnelle du thème et du rhème en français parlé. *Folia Linguistica* 15: 237–252.
1984   Diachronie und Pragmatik der spanischen Wortstellung. *Zeitschrift für romanische Philologie* 100: 92–111.
1987   Wortstellung, Satzperspektive und Textkonstitution im Iberoromanischen, dargestellt am Beispiel eines Textes von Juan Rulfo. *Iberoromania* 19: 1–16.
1990   *Sprachwissenschaft und Sprachphilosophie in der Romania*. Tübingen: Narr.

Braunmüller, Kurt
1982   *Syntaxtypologische Studien zum Germanischen*. Tübingen: Narr.

Breivik, Leiv E.
1986   Some remarks on cleft sentences in present-day English. In *Linguistics across Historical and Geographical Boundaries. In Honour of Jacek Fisiak*, Vol. 2, Dieter Kastovsky and Aleksander Szwedek (eds.), 815–826. Berlin and New York: Mouton de Gruyter.

Brockelmann, Carl
1913   *Grundriss der vergleichenden Grammatik der semitischen Sprachen, Bd. 2: Syntax*, Berlin: Verlag von Reuther & Reichard.

Brody, Jill
1984   Some problems with the concept of basic word-order. *Linguistics* 22: 711–736.

Burdach, Konrad
1914   Über deutsche Erziehung. *Zeitschrift für den deutschen Unterricht* 28: 657–678.

Burzio, Luigi
    1986   *Italian Syntax*. Dordrecht: Foris.

Calvo Pérez, Julio
    1993   *Pragmática y gramática del Quechua Cuzqueño.* (Monumenta Lingüística Andina 2). Cuzco: Centro de Estudios Regionales Andinos Bartolomé de Las Casas.

Cardinaletti, Anna
    1994   *La sintassi dei pronomi. Uno studio comparativo delle lingue germaniche e delle lingue romanze*. Bologna: il Mulino.

Carnie, Andrew and Guilfoyle, Eithne (eds.)
    2000   *The Syntax of Verb Initial Languages.* Oxford: Oxford University Press.

Cennamo, Michela
    1993   *The Reanalysis of Reflexives: A Diachronic Perspective*. Naples: Liguori.
    1995   Transitivity and V S order in Italian reflexives. *Sprachtypologie und Universalienforschung* 48 (1/2): 84–105.

Chafe, Wallace
    1986   Cognitive constraints on information flow. In *Coherence and Grounding in Discourse*, Russel S. Tomlin (ed.), 21–51. Amsterdam: John Benjamins.
    1994   *Discourse, Consciousness and Time*. Chicago: Chicago University Press.

Chapin, Paul G.
    1978   Easter Island: a characteristic V S O language. In *Syntactic Typology*, Winfred P. Lehmann (ed.), 139–168. Austin, TX: University of Texas Press.

Cinque, Guglielmo
    1977   The movement nature of left-dislocation. *Linguistic Inquiry* 8: 397–411.
    1990   *Types of Ā-Dependencies*. Cambridge, MA: MIT Press.
    1993   A null theory of phrase and compound stress. *Linguistic Inquiry* 24: 239–297.

CLG = Saussure, Ferdinand de
    1916   *Cours de linguistique générale*. Charles Bally and Alfred Sechehaye (eds.). Paris: Payot.

Como, Paola
    1995   Fumetto, giallo e romanzo d'autore: un'indagine sull'ordine delle parole. Unpublished M.A. thesis in sociolinguistics, Università di Napoli "Federico II".

Comrie, Bernard
    1981a   *The Languages of the Soviet Union*. Cambridge: Cambridge University Press.
    1981b   *Language Universals and Linguistic Typology. Syntax and Morphology*. Oxford: Blackwell.
    1984   Some formal properties of focus in Modern Eastern Armenian. *Annual of Armenian Linguistics* 5: 1–21.
    1996   *Sprache und Sprachen*: *Universalien und Typologie*. In *Deutsch-typologisch*, Ewald Land and Gisela Zifonun (eds.), 16–30. Berlin and New York: Mouton de Gruyter.

Connolly, J. H.
    1991   *Constituent Order in Functional Grammar: Synchronic and Diachronic Perspectives*. Berlin and New York: Foris.

Croft, William
    1990   *Typology and Universals*. Cambridge: Cambridge University Press.

Cruttenden, Alan
    1986   *Intonation*. Cambridge: Cambridge University Press.

Currie, Oliver
    2000   Word order stability and change from a sociolinguistic perspective: the case of Early Modern Welsh. In *Stability, Variation and Change of Word Order Patterns over Time*, Rosanna Sornicola, Erich Poppe, and Ariel Shisha-Halevy (eds.), 203–230. Amsterdam: John Benjamins.

530   Rosanna Sornicola

Daneš, František
    1967   Order of elements and sentence intonation. In *Intonation*, Dwight Bolinger (ed.), 216–232. Harmondsworth: Penguin.
de Groot, Casper
    1989   *Predicate Structure in a Functional Grammar of Hungarian*. Dordrecht: Foris.
Delanoy, Arnaud
    1990   Breton and the word-order concept. *Word* 42 (2): 219–221.
Delbrück, Berthold
    1878   *Die Altindische Wortfolge aus dem Çatapathabrāhmaṇa*. Halle: Verlag der Buchhandlung des Waisenhauses.
    1900   *Vergleichende Syntax der indogermanischen Sprachen*. Vol. III. Strassburg: Trübner.
    1976   *Altindische Syntax*. Darmstadt: Wissenschaftliche Buchgesellschaft. Original edition Halle 1988.
de Meij, Sjaak, and László K. Marácz
    1986   On question sentences in Hungarian. In *Topic, Focus and Configurationality*, Werner Abraham and Sjaak de Meij (eds.), 253–276. Amsterdam and Philadelphia: John Benjamins.
de Rijk, R. P. G.
    1978   Topic fronting, focus positioning and the nature of the verb phrase in Basque. In *Studies on Fronting*, F. Jansen (ed.), 81–112. Lisse: Peter de Ridder Press.
De Simone Brouwer, Francesco
    1921   *Grammatica della lingua greca*. Naples: Tipografia della Reale Università.
Dezső, László
    1973   Topics on syntactic typology. Paper presented at the Symposium on Typology, Prague 14–17 August 1973. [Also translated in Italian as: Problemi di tipologia sintattica. In *La tipologia linguistica*, Paolo Ramat (ed.), 239–261. Bologna: Il Mulino, 1976.]
Diderichsen, Paul
    1936a   Prolegomena til en metodisk dansk Syntax. *Forhandlinger paa det ottende nordiske Filologmode i København den 12–14 August 1935*, Copenhagen: Schultz. [Reprinted in *Paul Diderichsen. Helhed og Struktur. Udvalgte Sprogvidenskabelige Afhandlinger*, Anders Bjerrum et al. (eds.), Copenhagen: Gad.]
    1936b   Om den danske Sætnings Struktur. *Selskab for nordisk filologi*, *Aarsberetning for 1936*: 5–6.
Dik, Simon C.
    1978   *Functional Grammar*. Amsterdam: North Holland.
    1980   *Studies in Functional Grammar*. London and New York: Academic Press.
    1983   *Advances in Functional Grammar*. Dordrecht: Foris.
    1989   *The Theory of Functional Grammar*. Dordrecht: Foris.
Dik, Simon C., Maria E. Hoffmann, Jan R. de Jong, Sie J. Dijang, Harry Stroomer, and Lourens de Vries
    1981   On the typology of focus phenomena. In *Perspectives on Functional Grammar*, Teun Hoekstra, Harry van der Hulst, and Michael Moortgat (eds.), 41–74. Dordrecht: Foris.
Dooley Collberg, Sheila
    1988   GPSG and the verb-second phenomenon in Swedish. In *Papers from the Tenth Scandinavian Conference of Linguistics*, Vol. 1, Victoria Rosén (ed.), 132–146. Bergen: University of Bergen.
Dover, K. J.
    1960   *Greek Word-order*. Cambridge: Cambridge University Press.
Drach, Erich
    1963   *Grundgedanken der deutschen Satzlehre*. Frankfurt a.M.: Diesterweg. Original edition Frankfurt 1937.

Dressler, Wolfgang
    1969   Eine textsyntaktische Regel der indogermanischen Wortstellung. *Zeitschrift für verglei-chende Sprachforschung (KZ)* 83 (1): 1–25.
Dryer, Matthew S.
    1995   Frequency and pragmatically unmarked word order. In *Word Order in Discourse*, Pamela Downing and Michael Noonan (eds.), 105–135. Amsterdam: John Benjamins.
Ebert, Robert P.
    1980   Social and stylistic variation in Early New High German word-order: the sentence frame (Satzrahmen). *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)* 102: 357–398.
Ellenberger, Henry F.
    1994   *Histoire de la découverte de l'inconscient*. Paris: Fayard.
Engel, Ulrich
    1982   *Syntax der deutschen Gegenwartssprache*. Berlin: Schmidt.
    1991   *Deutsche Grammatik*. 2nd revised edition, Heidelberg: Julius Groos Verlag.
Erdmann, Peter
    1990   *Discourse and Grammar. Focussing and Defocussing in English*. Tübingen: Niemeyer.
Erguvanli, Eser
    1984   *The Function of Word-Order in Turkish Grammar*. (University of California Publications in Linguistics 106). Berkeley and Los Angeles: University of California Press.
Erguvanli-Taylan, Eser
    1987   The role of semantic features in Turkish word-order. *Folia Linguistica* 21 (2–4): 215–227.
Etzensperger, Jürg
    1979   *Die Wortstellung der deutschen Gegenwartssprache als Forschungsobjekt*. Berlin and New York: Walter de Gruyter.
Euler, Wolfram
    1982   Es war ein König. Eine Einleitungsformel mit indogermanischer Grundlage. In *Sprachwissenschaft in Innsbruck*: *Arbeiten von Mitgliedern und Freunden des Instituts für Sprachwissenschaft der Universität Innsbruck aus Anlass des fünfzigjährigen Bestehens des Instituts im Jahre 1978 und zum Gedenken an die 25. Wiederkehr des Todestages von Hermann Ammann am 12. September 1981*, Wolfgang Meid, Hermann Ölberg, and Hans Schmeja (eds.), 53–68. Innsbruck: Das Institut.
Evans, Simon D.
    1989   *A Grammar of Middle Welsh*. Dublin: Dublin Institute for Advanced Studies. Original edition 1964.
Fernandez, M. M. Jocelyne
    1986   Word-order and enunciative particles in the Sami ("Lappish") dialogue. A contrastive and panchronic approach. *Finnisch-Ugrische Mitteilungen* 10: 113–125.
    1994   *Les particules énonciatives dans la construction du discours*. Paris: Presses Universitaires de France.
Fife, James, and Gareth King
    1991   Focus and the Welsh "abnormal sentence": a cross-linguistic perspective. In *Studies in Brythonic Word-order*, James Fife and Erich Poppe (eds.), 81–153. Amsterdam: John Benjamins.
Fillmore, Charles
    1968   The case for case. In *Universals in Linguistics*, Emmon Bach and Robert T. Harms (eds.), 1–88. New York: Holt, Rinehart & Winston.
Firbas, Ian
    1964   On defining the theme in functional sentence analysis. *Travaux Linguistiques de Prague* 1: 267–280.

532   Rosanna Sornicola

1965   A note on transition proper in functional sentence perspective. *PhP* 8: 170–176.

1992   *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.

Fleischmann, Klaus

1973   *Verbstellung und Relieftheorie. Ein Versuch zur Geschichte des deutschen Nebensatzes*. Munich: Fink.

Fougeron, Irina

1989   *Prosodie et organisation du message*. Paris: Klincksieck.

Freeman, R. A.

1937   *Syntax of Old Icelandic*. Leeds: Chorley & Pickersgill.

Fromm, H.

1982   *Finnische Grammatik*. Heidelberg: Winter.

Gadet, Françoise

1992   *Le français populaire*. Paris: Presses Universitaires de France.

Garbini, Giovanni

1984   *Le lingue semitiche. Studi di storia linguistica*. Naples: Istituto Universitario Orientale, Dipartimento di Studi Asiatici.

Gazdar, Gerald, Ewan Klein, and Geoffrey K. Pullum (eds.)

1983   *Order, Concord and Constituency*. Dordrecht: Foris.

Givón, Talmy

1977a   On the V S word-order in Israeli Hebrew: pragmatics and typological change. In *Studies in Modern Hebrew Syntax and Semantics: The Transformational-Generative Approach*, (North-Holland Linguistic Series 32), Peter Cole (ed.), 153–181. Amsterdam: North-Holland.

1977b   The drift from VSO to SVO in Biblical Hebrew. In *Mechanisms of Syntactic Change*, Charles Li (ed.), 181–254. Austin, TX: Texas University Press.

1983   Topic continuity in discourse: an introduction. In *Topic Continuity in Discourse. A Quantitative Cross-language Study*, Talmy Givón (ed.), 1–41. Amsterdam and Philadelphia: John Benjamins.

Givón, Talmy (ed.)

1983   *Topic Continuity in Discourse. A Quantitative Cross-language Study*. Amsterdam and Philadelphia: John Benjamins.

Glinert, Lewis

1989   *The Grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Goldenberg, Gideon

1998   *Studies in Semitic Linguistics. Selected Writings*. Jerusalem: The Magnes Press & The Hebrew University.

Gonda, J.

1952   *Remarques sur la place du verbe dans la phrase active et moyenne en langue sanscrite*. Utrecht: Oosthoek's Uitgevers.

Graffi, Giorgio

1994   *Sintassi*. Bologna: Il Mulino.

Greenberg, Joseph H.

1963   Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, Joseph H. Greenberg (ed.), 73–113. Cambridge, MA: MIT Press.

Gruber, Jeffrey

1967   Topicalization in child language. *Foundations of Language* 3: 37–65.

Guentchéva, Zlatka

1994   *Thématisation de l'objet en bulgare*. Bern: Lang.

Gvozdanović, Jadranka
  1981   Word-order and displacement in Serbo-Croatian. In *Predication and Expression in Functional Grammar*, Machtelt Bolkestein et al. (eds.), 125–141. London: Academic Press.
Haberland, Hartmut
  1993   Thetic/categorical distinction. In *Encyclopedia of Language and Linguistics, Vol. 9: Syntax*, R. E. Asher and J. M. Y. Simpson (eds.), 4605–4606. Oxford and New York: Pergamon Press.
Haegeman, Liliane
  1991   *Introduction to Government and Binding Theory*. Oxford: Blackwell.
Haider, Hubert
  1986   Verb-second in German. In *Verb-second Phenomena in Germanic Languages*, Hubert Haider and Martin Prinzhorn (eds.), 49–76. Dordrecht: Foris.
Haiman, John
  1974   *Targets and Syntactic Change*. The Hague: Mouton.
Hakulinen, Auli, Fred Karlsson, and Maria Vilkuna
  1980   *Suomen tekstilauseiden piirteitä: kvantitatiivinen tutkimus*. (Publication 6) Helsinki: University of Helsinki, Department of General Linguistics.
Hakulinen, Lauri
  1961   *The Structure and Development of the Finnish Language*. (Indiana University Publications, Uralic and Altaic Series 3). Bloomington, IN: Indiana University Press.
  1979   *Suomen kielen rakenne ja kehitys*. Helsinki: Otava.
Halliday, Michael A. K.
  1967a   Theme and information in the English clause. In *System and Function in Language. Selected Papers*, Gunther Kress (ed.), 174–188. Oxford: Oxford University Press.
  1967b   Notes on transitivity and theme in English, Part 1. *Journal of Linguistics* 3 (1): 37–81.
  1967c   Notes on transitivity and theme in English, Part 2. *Journal of Linguistics* 3 (2): 199–244.
  1968   Notes on transitivity and theme in English, Part 3. *Journal of Linguistics* 4 (2): 179–215.
Hankamer, Jorge
  1971   Constraints on deletion in syntax. Yale university dissertation.
Harris, Alice C.
  1981   *Georgian Syntax. A Study in Relational Grammar*. Cambridge: Cambridge University Press.
Havránek, Bohuslav
  1968   Quelques problèmes de l'étude diachronique de la structure syntactique, surtout en slave. *Travaux Linguistiques de Prague* 3: 9–16.
Hawkins, John A.
  1983   *Word-order Universals*. New York: Academic Press.
  1994   *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
Heger, Klaus
  1982   "Il la lui a donné, à Jean, son père, sa moto." Neue Überlegungen zu einem alten Beispiel. In *Festschrift für Johannes Hubschmid zum 65. Geburtstag: Beiträge zur allgemeinen, indogermanischen und romanischen Sprachwissenschaft*, Otto Winkelmann and Mario Braisch (eds.), 53–66. Bern: Francke.
Helbig, Gerhard, and Joachim Buscha
  1989   *Deutsche Grammatik*. Leipzig: Enzyklopädie.
Heltoft, Lars
  1986   The V-2 analysis: a reply from Diderichsen-tradition. In *Scandinavian Syntax*, Östen Dahl and Anders Holmberg (eds.), 50–67. Stockholm: Institute of Linguistics.

534   Rosanna Sornicola

Hemon, Roparz
    1984   *A Historical Morphology and Syntax of Breton*. Dublin: The Dublin Institute for Advanced Studies.

Hendrik, Randall (ed.)
    1990   *The Syntax of the Modern Celtic Languages*. (Syntax and Semantics 23). San Diego, CA: Academic Press.

Henriksen, Carol
    1986   Historical Danish syntax, within the framework of Paul Diderichsen's syntactic theory. In *Scandinavian Syntax*, Östen Dahl and Anders Holmberg (eds.), 75–87. Stockholm: Institute of Linguistics.

Hetzron, Robert
    1971   Presentative function and presentative movement. *Studies in African Linguistics* (Supplement 2): 79–105.
    1975   The presentative movement, or why the ideal word-order is V S O P. In *Word-order and Word-order Change*, Charles Li (ed.), 345–388. Austin, TX: University of Texas Press.

Hinds, John
    1986   *Japanese*. London: Routledge.

Höhle, Tilman N.
    1982   Explikationen für "normale Betonung" und "normale Wortstellung". In *Satzglieder im Deutschen*, Werner Abraham (ed.), 75–153. Tübingen: Narr.

Holden, Kyril, and Maria Krupp
    1987   Word-order in Russian transitive sentences. *Folia Slavica* 8 (2–3): 254–271.

Holes, Clive
    1995   *Modern Arabic. Structures, Functions and Varieties*. London: Longman.

Holland, Gary B.
    1980   Problems of word-order change in selected Indo-European languages. Unpublished doctoral dissertation, University of California at Berkeley.

Holloway King, Tracy
    1995   *Configuring Topic and Focus in Russian*. Stanford, CA: Center for the Study of Language and Information.

Holmberg, Anders
    1986   *Word-order and Syntactic Features in the Scandinavian Languages and English*. Stockholm: Department of General Linguistics.
    1989   Verb movement in Finnish. In *Papers from the Fifth Nordic Conference on Bilingualism*, Vol. 1, (Studies in Language 14), Jussi Niemi (ed.), 112–130. Joensuu: Joensuun yliopisto, Humanistinen tiedekunta.

Holmes, Philip, and Ian Hinchliffe
    1994   *Swedish. A Comprehensive Grammar*. London: Routledge.

Horvath, Julia
    1986   *Focus in the Theory of Grammar and the Syntax of Hungarian*. Dordrecht: Foris.

Huang, C.-T. James
    1984   On the distribution and reference of empty pronouns. *Linguistic Inquiry* 15: 531–574.

Ingham, B.
    1991   Sentence structure in Khuzistani Arabic. In *Semitic Studies in Honour of W. Leslau on the Occasion of his 85th Birthday*, Vol. 1, Alan S. Kaye (ed.), 714–728. Wiesbaden: Harrassowitz.

Jacob, Daniel
    1990   *Markierung von Aktantenfunktionen und "Prädetermination" im Französischen*. Tübingen: Niemeyer.

Jacobs, Joachim
  1988   Fokus-Hintergrund-Gliederung und Grammatik. In *Intonationsforschungen*, Hans Alt-
              mann (ed.), 98–134. Tübingen: Niemeyer.
Jakobson, Roman
  1963   Implications of language universals for linguistics. In *Universals of Language*, Joseph
              H. Greenberg (ed.), 208–219. Cambridge, MA: MIT Press.
Jespersen, Otto
  1969   *Language. Its Nature, Development and Origin*. London: Allen & Unwin. Original edi-
              tion 1922.
Jo, Mi-Jeung
  1995   The theory of syntactic focalization on a subcategorization feature of verbs. In *Dis-
              course Configurational Languages*, Katalin É. Kiss (ed.), 335–374. Oxford: Oxford
              University Press.
Jongeling, K.
  1991   On the V S O character of Hebrew. In *Studies in Hebrew and Aramaic Syntax Presented
              to Professor J. Hoftijzer*, K. Jongeling, H. L. Murre-Van den Berg, and L. Van Rompay
              (eds.), 103–111. Leiden: Brill.
Kalitsunakis, Johannes
  1963   *Grammatik der neugriechischen Volkssprache*. Berlin: Walter de Gruyter.
Kálmán, László, Gábor Prószéky, Ádám Nádasdy, and György Kálmán C.
  1986   Hocus, focus and verb types in Hungarian infinitive constructions. In *Topic, Focus and
              Configurationality*, Werner Abraham and Sjaak de Meij (eds.), 129–142. Amsterdam
              and Philadelphia: John Benjamins.
Karlsson, Fred
  1987   *Finnish Grammar*. Porvoo-Helsinki-Juva: Werner Söderström Osakeyhtiö.
Karttunen, Lauri, and Martin Kay
  1985   Parsing in a free word-order language. In *Natural Language Parsing*, D. Dowty, L.
              Karttunen, and A. Zwicky (eds.), 279–306. Cambridge: Cambridge University Press.
Keenan, Edward
  1976   Towards a universal definition of subject. In *Subject and Topic*, Charles Li (ed.), 303–
              333. New York: Academic Press.
Kenesei, István
  1986   On the logic of word-order in Hungarian. In *Topic, Focus and Configurationality*,
              Werner Abraham and Sjaak de Meij (eds.), 143–160. Amsterdam and Philadelphia: John
              Benjamins.
Kim, Alan Hyun-Oak
  1988   Pre-verbal focussing and type XXIII language. In *Studies in Syntactic Typology*,
              Michael Hammond, Edith Moravcsik, and Jessica R. Wirth (eds.), 147–169. Amster-
              dam: John Benjamins.
Kippasto, Ann
  1989   Word-order in Votian. In *Proceedings of the Fifth International Finno-Ugrist Stu-
              dents' Conference*, Riho Grünthal, Sirpa Penttinen, and Tapani Salminen (eds.), 84–89.
              Helsinki: Finno-Ugrian Society.
Kirsner, Robert S.
  1979   *The Problem of Presentative Sentences in Modern Dutch*. Amsterdam: North-Holland.
Kiss, Katalin É.
  1987   *Configurationality in Hungarian*. Dordrecht: Reidel.
  1995   Introduction. In *Discourse Configurational Languages*, Katalin É. Kiss (ed.), 3–27. Ox-
              ford: Oxford University Press.
Kiss, Katalin É. (ed.)
  1995   *Discourse Configurational Languages*. Oxford: Oxford University Press.

536   Rosanna Sornicola

Kohonen, Viljo
    1982   A note on factors affecting the position of accusative objects and complements in Æl-
                fric's Catholic Homilies I. In *Approaches to Word-order*: *Reports on Text Linguistics*,
                Nils Erik Enkvist and Viljo Kohonen (eds.), 175–196. Åbo: Åbo Akademi.
Komlósy, András
    1986   Focussing on focus in Hungarian. In *Topic, Focus and Configurationality*, Werner Abra-
                ham and Sjaak de Meij (eds.), 215–226. Amsterdam and Philadelphia: John Benjamins.
Kovtunova, I. I.
    1976   *Sovremennyj russkij jazyk. Porjadok slov i aktual'noe členenie predloženija*. Moscow:
                Prosveščenie.
Kress, Bruno
    1982   *Isländische Grammatik*. Leipzig: Enzyklopädie.
Krylova, O., and S. Khavronina
    1976   *Word-order in Russian Sentences*. Moscow: Russian Language Publishers.
Ladd, Robert D.
    1996   *Intonational Phonology*. Cambridge: Cambridge University Press.
Lafitte, Pierre
    1978   *Grammaire Basque (Navarro-Labourdin Littéraire)*. Donostia: Elkar. Original edition
                1944.
Lambrecht, Knud
    1981   *Topic, Antitopic and Verb-agreement in Non-standard French*. Amsterdam: John Ben-
                jamins.
    1994   *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.
Lampe, G. W. H. (ed.)
    1969   *The Cambridge History of the Bible, Vol. 2: The West from the Fathers to the Reforma-
                tion*. Cambridge: Cambridge University Press.
Lehmann, Winfred P.
    1973   A structural principle of language and its implications. *Language* 49: 47–66.
    1974   *Proto-Indo-European Syntax*. Austin, TX: University of Texas Press.
Lerch, Eugen
    1934   *Historische Französische Syntax, Bd. III: Modalität*. Leipzig: Reisland.
Levelt, William, J. M.
    1989   *Speaking. From Intention to Articulation*. Cambridge, MA: MIT Press.
Lewis, Henry
    1942   *The Sentence in Welsh*. London: British Academy.
Lewis, Henry, and Holger Pedersen
    1937   *A Concise Comparative Celtic Grammar*. Göttingen: Vandenhoeck & Ruprecht.
Lewy, E.
    1922   *Tscheremissische Grammatik*. Leipzig: Haessel Verlag.
Li, Charles (ed.)
    1976   *Subject and Topic*. New York: Academic Press.
Linde, P.
    1923   Die Stellung des Verbs in der lateinischen Prosa. *Glotta* 12: 153–178.
Lyons, John
    1977   *Semantics*. 2 Vols. Cambridge: Cambridge University Press.
MacCana, Proinsias
    1973   On Celtic word-order and the Welsh abnormal sentence. *Ériu* 24: 90–120.
    1991   Further notes on constituent order in Welsh. In *Studies in Brythonic Word-order*, James
                Fife and Erich Poppe (eds.), 45–80. Amsterdam: John Benjamins.

MacCoisdealbha, Pádraig
  1998   *The Syntax of the Sentence in Old Irish. Selected Studies from a Descriptive, Historical and Comparative Point of View*. Tübingen: Niemeyer.

Marantz, Alec P.
  1984   *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.

Marchello-Nizia, Christiane
  1995   *L'évolution du français. Ordre des mots, démonstratifs, accent tonique*. Paris: Colin.

Marouzeau, Jules
  1938   *L'ordre des mots dans la phrase latine, Vol. 2: Le verbe*. Paris: Les Belles Lettres.

Masica, Colin P.
  1991   *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.

Maslova, Elena
  1995   V S / S V opposition in Russian and the thetic/categorical distinction. *Sprachtypologie und Universalienforschung* 48 (1/2): 106–124.

Mathesius, Vilém
  1939   O tak zvaném aktuálním členění větném. *Slovo a slovesnost* 5: 171–174.
  1941–1942   Ze srovnávacích studií slovosledných. *Časopis pro moderní filologii* 28: 181–190; 302–307.

Matthews, Peter H.
  1991   *Morphology*. Cambridge: Cambridge University Press.
  1994   *Grammatical Theory in the United States from Bloomfield to Chomsky*. Cambridge: Cambridge University Press.

Matthews, W. K.
  1960   *Russian Historical Grammar*. London: The Athlone Press.

McCloskey, James
  1983   A VP in a VSO language? In *Order, Concord and Constituency*, Gerald Gazdar, Ewan Klein, and Geoffrey K. Pullum (eds.), 9–55. Dordrecht: Foris.

Meillet, Antoine
  1906   L'état actuel des études de linguistique générale, Leçon d'ouverture du cours de Grammaire comparée au Collège de France, lue le mardi 13 février 1906. [Also in 1965 *Linguistique historique et linguistique générale*, Antoine Meillet (ed.), 1–18. Paris: Champion.]
  1912   L'évolution des formes grammaticales. *Scientia* 12 (26): 6. [Also in 1965 *Linguistique historique et linguistique générale*, Antoine Meillet (ed.), 130–148. Paris: Champion.]

Meillet, A., and H. De Willman-Grabowska
  1921   *Grammaire de la langue polonaise*. Paris: Champion.

Meillet, A., and A. Vaillant
  1924   *Grammaire de la langue serbo-croate*. Paris: Champion.

Milano, Emma
  1994   Sul parlato: problemi di ordine delle parole. Unpublished M.A. thesis in sociolinguistics, Università di Napoli "Federico II".

Miller, Jim
  1995   VS clauses in Russian: contrast, new episodes and theticity. *Sprachtypologie und Universalienforschung* 48 (1/2): 125–141.

Miller, Jim, and Regina Weinert
  1998   *Spontaneous Spoken Language*. Oxford: Clarendon Press.

Mitchell, Bruce
  1985   *Old English Syntax*. 2 Vols. Oxford: Clarendon Press.

538   Rosanna Sornicola

Mohanan, Karuvannur Puthanveettil, and Tara Mohanan
    1994   Issues in word-order. In *Theoretical Perspectives on Word-order in South Asian Languages*, Miriam Butt, Tracy Holloway King, and Gillian Ramchand (eds.), 153–216. Stanford, CA: Center for the Study of Language and Information.
Muysken, Pieter
    1995   Focus in Quechua. In *Discourse Configurational Languages*, Katalin É. Kiss (ed.), 375–393. Oxford: Oxford University Press.
Nichols, Johanna
    1985   The grammatical marking of theme in literary Russian. In *Issues in Russian Morphosyntax*, Michael S. Flier and Richard D. Brecht (eds.), 170–208. Columbus, OH: Slavica Publishers.
Nuyts, Jan, and George de Schutter (eds.)
    1987   *Getting One's Words into Line. On Word Order and Functional Grammar*. Dordrecht: Foris.
Nykrog, Per
    1957   L'influence latine savante sur la syntaxe du français. In *The Classical Pattern of Modern Western Civilization*: *Language*, (Acta Congressus Madvigiani, Travaux du Cercle Linguistique de Copenhagen 11), Franz Blatt (ed.), 89–114. Copenhagen: E. Munksgaard.
Ortiz de Urbina, John
    1995   Residual verb-second and verb-first in Basque. In *Discourse Configurational Languages*, Katalin É. Kiss (ed.), 99–121. Oxford: Oxford University Press.
Ó Siadhail, Mícheál
    1989   *Modern Irish. Grammatical Structure and Dialectal Variation*. Cambridge: Cambridge University Press.
Payne, Dorothy (ed.)
    1992   *Pragmatics of Word-order Flexibility*. Amsterdam and Philadelphia: John Benjamins.
Pedersen, Holger
    1976   *Vergleichende Grammatik der keltischen Sprachen*. 2 Vols. Göttingen: Vandenhoeck & Ruprecht. Original edition 1909–1913.
Perlmutter, David
    1978   Impersonal passives and the unaccusative hypothesis. *Berkeley Linguistic Society* 4: 157–189.
PET
    1995–1997   *Piccola Enciclopedia Treccani*, 12 Vols. Rome: Istituto dell'Enciclopedia Italiana.
Pétursson, Magnus
    1978   *Isländisch*. Hamburg: Helmut Buske Verlag.
Plank, Frans
    1984   *Ces obscurs objets du désir*. Introduction. In *Objects. Towards a Theory of Grammatical Relations*, Frans Plank (ed.), 1–8. London: Academic Press.
Plank, Frans (ed.)
    1984   *Objects. Towards a Theory of Grammatical Relations*. London: Academic Press.
Platzack, Christer
    1986   The position of the finite verb in Swedish. In *Verb-second Phenomena in Germanic Languages*, Hubert Haider and Martin Prinzhorn (eds.), 27–48. Dordrecht: Foris.
Poppe, Erich
    1989   Constituent ordering in *Breudwyt Maxen Wledic*. *Bulletin of the Board of Celtic Studies* 36: 43–63.
    1990   Word-order patterns in *Breudwyt Ronabwy*. In *Celtic Linguistics: Ieithyddiaeth Geltaidd. Readings in the Brythonic Languages. Festschrift for T.A. Watkins*, Martin J. Ball, James Fife, Erich Poppe, and Jenny Rowland (eds.), 445–460. Amsterdam: John Benjamins.

1991a	Word-order in *Cyfranc Lludd a Llefelys*: notes on the pragmatics of constituent-ordering in MW narrative prose. In *Studies in Brythonic Word-order*, James Fife and Erich Poppe (eds.), 155–204. Amsterdam: John Benjamins.
1991b	*Untersuchungen zur Worstellung im Mittelkymrischen*. Hamburg: Helmut Buske.
2000	Constituent order in Middle Welsh: the stability of the pragmatic principle. In *Stability, Variation and Change of Word Order Patterns over Time*, Rosanna Sornicola, Erich Poppe, and Ariel Shisha-Halevy (eds.), 41–51. Amsterdam: John Benjamins.

Poppe, Nicholas
1954	*Grammar of Written Mongolian*. Wiesbaden: Harrassowitz.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik
1985	*A Comprehensive Grammar of the English Language*. London: Longman.

Rebuschi, Georges
1984	*Structure de l'énoncé en basque*. Paris: Société d'études linguistiques et anthropologiques de France.

Renzi, Lorenzo
1984	La tipologia dell'ordine delle parole e le lingue romanze. *Linguistica* 24: 27–59.

Rouveret, Alain
1990	X-bar theory, minimality and barrierhood in Welsh. In *The Syntax of the Modern Celtic Languages*, (Syntax and Semantics 23), Randall Hendrik (ed.), 27–79. San Diego, CA: Academic Press.

Rowlands, Eurys I.
1980	Y frawddeg gymysg a'r frawddeg dro. *Bulletin of the Board of Celtic Studies* 28: 218–222.

Russ, Charles V. J. (ed.)
1989	*The Dialects of Modern German. A Linguistic Survey*. Stanford, CA: Stanford University Press.

Rydén, Mats, and Sverker Brorström
1987	*The Be/Have Variation with Intransitives in English*. Stockholm: Almqvist & Wiksell.

Sáenz-Badillos, Angel
1993	*A History of the Hebrew Language*. Cambridge: Cambridge University Press.

Salvi, Giampaolo
1993	Ordine delle parole e struttura della frase nelle lingue romanze antiche. In *Grammatikographie der romanischen Sprachen*, (Akten der gleichnamigen Sektion des Bamberger Romanistentages [23–29.9.1991]), Christian Schmitt (ed.), 455–477. Bonn: Romanistischer Verlag.
1997	La posizione tipologica dell'italiano tra le lingue romanze. *Italienische Studien* 18: 25–38.

Sasse, Hans-Jürgen
1987	The thetic/categorical distinction revisited. *Linguistics* 25: 511–580.
1995a	*Theticity* and V S order: a case study. *Sprachtypologie und Universalienforschung* 48 (1/2): 3–31.
1995b	V S clauses in Modern Greek and Hungarian. *Sprachtypologie und Universalienforschung* 48 (1/2): 142–188.

Sauvageot, Aurélien
1973	*L'élaboration de la langue finnoise*. Paris: Klincksieck.

Scaglione, Aldo
1972	*The Classical Theory of Composition from its Origins to the Present. A Historical Survey*. Chapel Hill, NC: University of North Carolina Press.
1981	*The Theory of German Word-order from the Renaissance to the Present*. Minneapolis, MN: University of Minnesota Press.

540   Rosanna Sornicola

Scherpenisse, Wim
    1986   *The Connection between Base Structure and Linearization Restrictions in German and Dutch*. Frankfurt a. M.: Peter Lang.

Schroeder, Christoph
    1995   Post-predicate argument in Turkish. *Sprachtypologie und Universalienforschung* 48 (1/2): 204–218.

Schwyzer, Eduard
    1950   *Griechische Grammatik, Band 2: Syntax und syntaktische Stilistik*. Munich: Beck'sche Verlagsbuchhandlung.

Sgall, Peter, Eva Hajičová, and Jarmila Panevová
    1986   *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Dordrecht: Reidel.

Shisha-Halevy, Ariel
    1995   Structural sketches of Middle Welsh syntax. I: The converter systems. *Studia Celtica* 29: 127–223.
    2000   Stability in clausal/phrasal pattern constituent sequencing: 4000 years of Egyptian (with some theoretical reflections, also on Celtic). In *Stability, Variation and Change of Word Order Patterns over Time*, Rosanna Sornicola, Erich Poppe, and Ariel Shisha-Halevy (eds.), 71–100. Amsterdam: John Benjamins.

Siewierska, Anna
    1988   *Word-Order Rules*. Beckenham: Croom Helm.

Siewierska, Anna (ed.)
    1998   *Constituent Order in the Languages of Europe*. Berlin and New York: Mouton de Gruyter.

Sigurðsson, Halldór Á.
    1986   Verb post-second in a V2 language. In *Scandinavian Syntax*, Östen Dahl and Anders Holmberg (eds.), 138–149. Stockholm: Institute of Linguistics.
    1989   *Verbal Syntax and Case in Icelandic in a Comparative GB Approach*. Lund: Department of Scandinavian Languages.

Sinor, Denis
    1990   Introduction. In *The Cambridge History of Early Inner Asia*, Denis Sinor (ed.), 1–18. Cambridge: Cambridge University Press.

Sohn, Ho-min
    1994   *Korean*. London: Routledge.

Sørensen, Knud
    1957   Latin influence on English syntax. In *The Classical Pattern of Modern Western Civilization*: *Language*, (Acta Congressus Madvigiani, Travaux du Cercle Linguistique de Copenhagen 11), Franz Blatt (ed.), 131–155. Copenhagen: E. Munksgaard.

Sornicola, Rosanna
    1981   *Sul parlato*. Bologna: Il Mulino.
    1986   Costituenza, dipendenza e la struttura tema-rema in Italiano. In *Tema-rema in Italiano*, Harro Stammerjohann (ed.), 121–139. Tübingen: Narr.
    1988   *It*-clefts and *wh*-clefts: two awkward sentence types. *Journal of Linguistics* 24: 343–379.
    1990   Soggetti prototipici e soggetti non prototipici: l'italiano a confronto con le altre lingue europee. *Lingua e Stile* 25 (1): 3–24.
    1993a   Topic, focus and word-order. In *Encyclopedia of Language and Linguistics, Vol. 9: Syntax*, R. E. Asher and J. M. Y. Simpson (eds.), 4633b–4640b. Oxford and New York: Pergamon Press.
    1993b   The many routes of functionalism. *Rivista di Linguistica* 5 (1): 157–176.
    1994   On word-order variability: a study from a corpus of Italian. *Lingua e Stile* 29 (1): 25–57.

1995a   Mutamenti di prospettiva culturale nelle lingue europee moderne: l'influenza del latino sulla sintassi. In *Kulturwandel im Spiegel des Sprachwandels*, Karl-Egon Lönne (ed.), 41–58. Tübingen and Basel: Francke.

1995b   Theticity, V S order and the interplay of syntax, semantics and pragmatics. *Sprachtypologie und Universalienforschung* 48 (1/2): 72–83.

1998   The interpretation of historical sources as a problem for diachronic typology: word-order in English rhetorical and grammatical treatises of XVIth and early XVIIth centuries. In *The Virtues of Language. History in Language, Linguistics and Text in Memory of Thomas Frank*, Dieter Stein and Rosanna Sornicola (eds.), 81–128. Amsterdam: John Benjamins.

1999   Labile orders and grammatical functions. The functional representation of one-argument structures. In *Boundaries of Morphology and Syntax*, Lunella Mereu (ed.), 291–305. Amsterdam: John Benjamins.

2000   Stability, variation and change in word order: some evidence from the Romance languages. In *Stability, Variation and Change of Word Order Patterns over Time*, Rosanna Sornicola, Erich Poppe, and Ariel Shisha-Halevy (eds.), 101–115. Amsterdam: John Benjamins.

Sornicola, Rosanna, and Aleš Svoboda (eds.)

1992   *Il campo di tensione. La sintassi della Scuola di Praga*. Naples: Liguori.

Sornicola, Rosanna, Marcello Barbato, Maria Cesarano, and Daniela Puolato

1994   Word order in old Romance languages. Paper presented at the EUROTYP project meeting of the research group on Pragmatic Organization of Discourse, Naples, 11–13 November 1994.

Sridhar, S. N.

1990   *Kannada*. London: Routledge.

Srpová, Milena

1990   Quelques remarques sur la dynamique énonciative en tchèque et en français. *Revue d'Études Slaves* 62 (1/2): 403–416.

Stanchev, Svillen

1987   Some observations on the order of constituents in Bulgarian. In *Getting One's Words into Line. On Word Order and Functional Grammar*, Jan Nuyts and George de Schutter (eds.), 33–43. Dordrecht: Foris.

Steele, Susan

1978   Word-order variation: a typological study. In *Universals of Human Languages, Vol. 4: Syntax*, Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik (eds.), 585–623. Stanford, CA: Stanford University Press.

Stender-Petersen, Adolf, and Knud Jordal

1957   Das griechisch-byzantinische Erbe im Russischen. In *The Classical Pattern of Modern Western Civilization: Language*, (Acta Congressus Madvigiani, Travaux du Cercle Linguistique de Copenhagen 11), Franz Blatt (ed.), 163–218. Copenhagen: E. Munksgaard.

Stenson, Nancy

1981   *Studies in Irish Syntax*. Tübingen: Narr.

Svoboda, Aleš

1984   České slovosledné pozice z pohledu aktuálního členění [Czech word-order from the viewpoint of functional sentence perspective]. *Slovo a slovesnost* 45 (1): 22–34; 45 (2): 88–103.

Svolacchia, Marco, Lunella Mereu, and Annarita Puglielli

1995   Aspects of discourse configurationality in Somali. In *Discourse Configurational Languages*, Katalin É. Kiss (ed.), 65–98. Oxford: Oxford University Press.

542   Rosanna Sornicola

Swarajya Lakshmi, V., and Aditi Mukherjee (eds.)
  1996   *Word-order in Indian Languages*. Hyderabad: Centre of Advanced Study in Linguistics
              Osmania University & Brooklins Corporation.
Tajsner, Przemyslaw
  1990   Scrambling and the Polish word-order. An alternative hypothesis. *Papers and Studies in
              Contrastive Linguistics* 25: 43–57.
Tallerman, Maggie
  1991   The directionality of the head subcategorization in Welsh. In *Studies in Brythonic Word-
              order*, James Fife and Erich Poppe (eds.), 311–328. Amsterdam: John Benjamins.
Tarvainen, Kalevi
  1985   *Kontrastive Syntax Deutsch-Finnisch*. Heidelberg: Julius Groos Verlag.
Tauli, Valter
  1983   *Standard Estonian Grammar, Part 2: Syntax*. Uppsala: Borgströms Tryckeri.
Tesnière, Lucien
  1959   *Éléments de syntaxe structurale*. Paris: Klincksieck.
Thomason, Sarah, and Terrence Kaufman
  1988   *Language Contact, Creolization and Genetic Linguistics*. Berkeley and Los Angeles:
              California University Press.
Thompson, Sandra
  1978   Modern English from a typological point of view: some implications on the function of
              word-order. *Linguistische Berichte* 54: 19–35.
Thráinsson, Höskuldeer
  1986   V1, V2, V3 in Icelandic. In *Verb-second Phenomena in Germanic Languages*, Huber
              Haider and Martin Prinzhorn (eds.), 169–194. Dordrecht: Foris.
Thumb, Albert
  1912   *Handbook of the Modern Greek Vernacular. Grammar, Texts, Glossary*. Edinburgh:
              Clark.
Tiffou, Étienne, and Jürgen Pesot
  1989   *Contes du Yasin. Introduction au bourouchaski du Yasin avec grammaire et dictionnaire
              analytique*. Paris: Peeters.
Timm, Lenora
  1989   Word-order in 20th century Breton. *Natural Languages and Linguistic Theory* 7 (3):
              361–378.
  1991   Discourse pragmatics of NP-initial sentences in Breton. In *Studies in Brythonic Word-
              order*, James Fife and Erich Poppe (eds.), 275–310. Amsterdam: John Benjamins.
Tomlin, Russell S.
  1986   *Basic Word-order. Functional Principles*. London: Croom Helm.
Tompa, József
  1985   *Kleine ungarische Grammatik*. Budapest: Akadémiai Kiado.
Tsimpli, Ianthi Maria
  1995   Focusing in Modern Greek. In *Discourse Configurational Languages*, Katalin É. Kiss
              (ed.), 176–206. Oxford: Oxford University Press.
Uhlířová, Ludmila
  1969   Vztah syntaktické funkce větného členu a jeho místa ve větě [The relationship between
              syntactic function and linear position of sentence elements. Methods and results of sta-
              tistical research]. *Slovo a slovesnost* 30: 358–370 [Italian translation in Rosanna Sorni-
              cola and Aleš Svoboda (eds.)].
Ulrich, Moriţa
  1985   *Thetisch und Kategorisch*. Tübingen: Narr.
Urošević, Zoran M., C. Carello, M. Saviic, G. Lukatela, and M. T. Turvey
  1986   Some word-order effects in Serbo-Croat. *Language and Speech* 29 (2): 177–195.

Vähämäki, Börje
    1984   *Existence and Identity. A Study of the Semantics and Syntax of Existential Sentences in Finnish*. Åbo: Åbo Akademi.

van Oosten, Jeanne
    1985   *On the Nature of Topics, Agents, and Subjects: A Cognitive Explanation*. Bloomington, IN: Indiana University Linguistic Club.

van Steenbergen, Marlies
    1989   Finnish: configurational or not? In *Configurationality. The Typology of Asymmetries*, László Marácz and Pieter Muysken (eds.), 143–158. Dordrecht: Foris.

Vennemann, Theo
    1972   Analogy in generative grammar: the origin of word-order. In *Proceedings of the XIIth Congress of Linguists*, Vol. 2, Luigi Heilmann (ed.), 79–83. Bologna: Il Mulino.

Verpoorten, J.-M.
    1977   *L'ordre des mots dans l'Aitareya-Brāhmaṇa*. Paris: Les Belles Lettres.

Vilkuna, Maria
    1989   *Free Word-order in Finnish. Its Syntax and Discourse Functions*. Helsinki: Soumalaisen Kirjallisuunden Seura.

Vogt, Hans
    1971   *Grammaire de la langue géorgienne*. Oslo: Universitets-forlaget.
    1974   L'ordre des mots en géorgien moderne. *Bedi Kartlisa* 32: 48–56.

von Stechow, Armin, and Susanne Uhmann
    1986   Some remarks on focus projection. In *Topic, Focus and Configurationality*, Werner Abraham and Sjaak de Meij (eds.), 295–320. Amsterdam and Philadelphia: John Benjamins.

Wackernagel, Jacob
    1892   Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen* 1: 333–436. [Also in Wackernagel, Jacob (1953). *Kleine Schriften 1*. 1–104. Göttingen: Vandenhoeck & Ruprecht].

Wagner, Heinrich
    1959   *Das Verbum in den Sprachen der britischen Inseln. Ein Beitrag zur geographischen Typologie des Verbums*. Tübingen: Buchreihe der Zeitschrift für Celtische Philologie 1.
    1967   Zur unregelmässigen Wortstellung in der altirischen Allitterationsdichtung. *Beiträge zur Indogermanistik und Keltologie*: 289–314.

Wartburg, Walter von
    1950   *Évolution et structure de la langue française*. Bern: Francke.

Watkins, Calvert
    1963   Preliminaries to a historical and comparative analysis of the syntax of the Old Irish verb. *Celtica* 6: 1–49.
    1976   Towards Proto-Indo-European syntax: problems and pseudo-problems. In *Papers from the Parasession on Diachronic Syntax (April 22, 1976)*, B. Steever, Carol A. Walker, and Salikoko S. Mufwene (eds.), 305–326. Chicago: Chicago Linguistic Society.

Watkins, T. Arwyn
    1991   The function of the cleft and non-cleft constituents orders in Modern Welsh. In *Studies in Brythonic Word-order*, James Fife and Erich Poppe (eds.), 329–351. Amsterdam: John Benjamins.

Watters, John Robert
    1979   Focus in Aghem: a study of its formal correlates and typology. In *Aghem Grammatical Structure*, (Southern California Occasional Papers in Linguistics 7), Larry M. Hyman (ed.), 137–197. Los Angeles: University of Southern California.

544   Rosanna Sornicola

Weber, David J.
    1989   *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley and London: University of California Press.
Weil, Henri
    1879   *De l'ordre des mots dans les langues anciennes comparées aux langues modernes*. Paris: Vieweg.
Weinert, Regina, and Jim Miller
    1996   Cleft constructions in spoken language. *Journal of Pragmatics* 25: 173–206.
Weinrich, Harald
    1993   *Textgrammatik der deutschen Sprache*. Mannheim and Leipzig: Duden Verlag.
Wells, Christopher J.
    1987   *German: A Linguistic History to 1945*. Oxford: Oxford University Press.
Williams, Stephen J.
    1980   *A Welsh Grammar*. Cardiff: University of Wales Press.
Willis, David W. E.
    1998   *Syntactic Change in Welsh. A Study of the Loss of Verb-second*. Oxford: Clarendon Press.
Yokoyama, Olga
    1986   *Discourse and Word-order*. Amsterdam and Philadelphia: John Benjamins.
Zifonun, Gisela, Ludger Hoffmann, and Bruno Strecker
    1997   *Grammatik der deutschen Sprache*. 3 Vols. Berlin: Walter de Gruyter.