

THE ENCYCLOPEDIA OF  
LANGUAGE  
AND  
LINGUISTICS

Volume 9

Editor-in-Chief  
**R. E. ASHER**  
*University of Edinburgh, UK*

Coordinating Editor  
**J. M. Y. SIMPSON**  
*University of Glasgow, UK*



PERGAMON PRESS  
OXFORD • NEW YORK • SEOUL • TOKYO

## Topic, Focus, and Word Order

### 1. Word Order: Its Controversial Nature

Throughout the history of linguistics, word order has basically been conceived as a phenomenon of a dual nature, first being related to grammar, second to style. As a phenomenon pertaining to grammar, word order is a device that codifies grammatical relations, the result being the 'basic' or fundamental word order patterns of a language. English as well as Italian, French, etc., are languages with a basic SVO order, because the functions of subject and object are, in each of these languages, carried by, respectively, preverbal and postverbal position (*John loves Mary*, where *John* is the subject, compared with *Mary loves John*, where *John* is the object). As a pragmatic phenomenon, word order encompasses those deviations from basic patterns that are due to pragmatic factors, such as, primarily, marked focus placement. (Other influential factors of a pragmatic/semantic nature are the contextual dependency of sentence constituents, which is responsible for the so-called 'theme-rheme distribution,' and also the referentiality and animacy hierarchies, according to which the most referential (or the most animate) element tends to occur on the left of the sentence.)

On the whole, what is implied in this polarized view of word order as a grammatical versus a pragmatic phenomenon is that the conditions determining the word order patterns of a given natural language are formal, i.e., highly abstract and mechanical, and that functional principles, such as focus placement, contextual dependency, the referentiality and animacy hierarchies, etc., can only at a later stage affect them. This view has been supported in late twentieth-century linguistic research mainly by generative grammarians (see *Generative Grammar*). It raises more general problems of linguistic theory, concerning the level of representation of syntactic structure on which 'order' should be placed. The idea that order is a mere realization device of structure can be traced back to Meillet and further back to Condillac and has been variously maintained by Tesnière (see *Tesnière, Lucien Valérius*), Halliday, and many other nongenerative linguists. The idea that structural configurations of generative grammar inherently have a linear dimension has from time to time been criticized by those who favor nonlinear models, such as Šaumjan and

Soboleva, Sgall, and, in the early 1980s, in different theoretical frameworks, Bresnan and the supporters of relational grammar (see *Relational Grammar*).

The alternative view has also been defended, according to which in every language word order is determined by the interplay of both formal (i.e., strictly grammatical) and functional factors. This view can be traced back to Mathesius (see *Mathesius, Vilém*) and to the Prague School tradition of syntactic functionalism (see *Prague School Syntax and Semantics*). In this approach every language has a certain degree of sensitivity to functional factors: word order patterns peculiar to different languages are thus the result of different dynamic interactions of such factors with grammatical ones. Grammatical factors, however, are considered as mechanical tendencies at work in every language, which are determined by the habit of always setting the same sentence constituents in the same place.

In this light, the concepts of basic versus 'nonbasic' word order, as well as the concepts of 'fixed' versus 'nonfixed' word order turn out to be idealizations of a high degree, being constructed on purely abstract sentence models rather than being arrived at by observations of real utterances. In dealing with pragmatic functions and word order, however, empirical studies of language in use should be taken into due account.

The two views are both challenged in the following discussion of topic, focus, and word order. However, any overall treatment of this subject cannot escape the fundamental difficulty that so far no unitary account has been proposed in both formal and functional terms. On the one hand, the notions of 'topic' and 'focus' have been introduced in the theoretical framework of generative grammar as terms devoid of their semantic/pragmatic import. On the other hand, the literature on word order often lacks a broader view of the set of possible syntactic structures in natural languages which are associated with functions such as topic and focus. What is more, it lacks a full understanding of formal regularities exhibited by these structures.

The following sections deal first with the crucial notions of topic and focus and second with a set of structures related to these two pragmatic functions which seem to be widely spread across natural languages: topicalization, left dislocation, right dislocation, clefting (in its various subtypes, such as *it*-clefting and *wh*-clefting), raising, passivization, extraposition. Finally, the relation between topic, focus, and syntactic structure is discussed.

## 2. Topic and Focus

There are no all-encompassing definitions of topic and focus in the literature. Both terms cover phenomena belonging to the whole spectrum of syntax, semantics, and pragmatics, with an extension to the phonological level.

Perhaps the most controversial concepts lie at the pragmatic level. Three (at least) definitions of topic and focus can be found with different terminology, but variants of the same thing:

- (a) Topic refers to information already present in linguistic or situational context (i.e., it is the contextually bound unit or configuration of units); focus refers to information nonpresent (or partially present) in linguistic or situational context (i.e., it is the contextually unbound (or partially unbound)

unit or configuration of units). With different terminology these two notions have been opposed to each other as 'given' (topic) versus 'new' (focus).

- (b) Topic (or theme) is the part of the sentence conveying the lowest degree of communicative dynamism; focus (or rheme) is the part with the highest degree of communicative dynamism.
- (c) Topic is presupposed information, focus is nonpresupposed information.

It is a fact that has been widely recognized, but for which up till now poor theoretical explanations have been given, that across natural languages there is a high tendency in unmarked sentences to map the contextually bound stretch of the sentence on to the subject and the contextually unbound stretch on to the predicate. This tendency as well as the fact that the overwhelming majority of the world's languages are either SVO or SOV raises a problem concerning linearity. There is in fact no *a priori* reason why the contextually bound (given/less dynamic/presupposed) part of the sentence should come first.

The same correlation shows itself in the semantic definition associated with the two terms (note, however, that at this level the term 'comment' is often found instead of focus). Here topic is 'what is being spoken about,' focus (or comment) is 'what is being said on what is being spoken about.'

When it comes to a purely syntactic definition of topic and focus, apparently it is the linear dimension of the sentence which is essentially involved. Different models have variously assessed the property of being a topic as the occurrence of a constituent *x* in the first position of the sentence. This generalization, however, can be questioned. The property of being an argument of the verb is no less important as a syntactic criterion than purely linear considerations. By a large consensus in the literature only those constituents that convey grammatical functions are considered as candidates for the topic function. Thus languages may have topics that do not occur in the first position of the sentence; to take but one example, the so-called circumstantial elements that express the temporal or spatial setting (in many languages, time or place adverbs) may be placed in the first position, as in the sentence *Yesterday Mary was in a very bad mood*. Here the topic is not *yesterday*, but the constituent with the subject function. Thus the idea seems well-founded that in general a position  $P_1$  should be differentiated from topic position  $P(\text{top})$ , although there will be many cases in which  $P_1 = P(\text{top})$ . Things are further complicated by the fact that the topic often coincides with a phrase with multiple constituents. An alternative and more satisfactory definition would not specify for topic just a unique position  $P_1$ , but a whole range of positions  $P_1 \dots P_k$  (where  $j \dots k$  space over a set of integers ranging from 1 to  $k$ ); a further condition should specify that,  $P_1 \dots P_n$  being the series of positions of syntactic structure,  $P_k$  never coincides with  $P_n$ .

Similar problems are faced in the attempt to obtain a syntactic definition of focus. Here again what is crucial is not merely the position inside the sentence, but also the categorial/functional nature of the constituent involved. Following a purely linear criterion, in fact, one could be led to assume that, at least in unmarked (i.e., nonemphatic) sentences, the focus position is the final one, since the linear

dimension of the sentence can be conceived as a serial process of adding information quanta, each quantum conveying a higher information value than its antecedent. This assumption is wrong for two reasons: first, as in the case of topic, focus often does not coincide with a single constituent, but with a configuration of constituents. To resume the preceding example, in the sentence *Yesterday Mary was in a bad mood* the whole string *was in a bad mood* is the focus (or, more precisely, the 'broader focus'), although inside this domain, some constituents are more focal than others (inside the VP domain the focus proper is the NP *a bad mood*; inside the NP domain, in unmarked sentences, the general consensus would be that the focus proper is the modifier *bad* (this is what is called 'narrow focus')). Second, the final position could be occupied by a circumstantial element, which is a typical nonargument of the verb: the previous assignment of 'foci' holds true even in *Mary was in a bad mood yesterday* as well as in *Mary was in a bad mood during her stay in Rome*, when these sentences are uttered with normal intonation contours (i.e., *yesterday* and *during her stay in Rome* have either no nucleus or a secondary one). Similar considerations hold true for other constituents occurring in the final position that do not have a strong dependency relation with the predicate frame and thus are extrasentential (e.g., *he said* in *John was upset by the War, he said*) or appositional (e.g., *Peter* in *I have just met my brother, Peter*; note that in *I have just met my brother Peter*, with no pause between *brother* and *Peter* and the nucleus on *Peter*, the latter constituent is the focus proper of the sentence).

One is thus entitled to think that the strength of dependency relations is a fundamental parameter in determining focus. What is suggested here is that focus is correlated to dependency structure. This formulation could be refined in terms of the relation between heads (i.e., governing constituents) and their governed constituents. In unmarked sentences, in fact, natural languages seem to show a correlation between highly focal constituents and the property of being governed. Thus, for example, in sentences with the configuration  $[s[NP \dots] + [VP V + [NPART + ADJ + N_{NP}]VP]S]$  the focus would cover the NP (in many cases ADJ would be the focus proper); in sentences with the configuration  $[s[NP \dots] + [VP V + [PPPREP + [NP \dots]PP]VP]S]$  the focus will cover the PP (with NP as the focus proper—more precisely, the governed constituent, if any, in NP).

Of course, languages vary according to the degree of conformity between the structural and the linear principle of focus assignment. This has to do with typological characteristics affecting grammar to a different extent. Consider, for example, the case of Japanese or other Altaic languages, where the predominant SOV order is systematically reflected in the general operator–operand order. In sentences such as Japanese (1–2) the focus proper (*Rōma* and *akaku*, respectively) does not occur in sentence final position. Languages of this type are thus said to have an 'unsusceptibility to functional sentence perspective' (FSP) (i.e., the linear criterion). In principle, the highest degree of conformity to FSP is exhibited by SVO (or operand–operator) languages. However, even these latter may deviate from it in some part of their grammar. Consider the case of the adjectival prenominal position in English, which results in

focus not occurring in phrase-final or sentence-final position:

Kinoo (watashi wa) Rōma e itta (1)  
yesterday I TOPIC MARKER Rome to have been  
'Yesterday I went to Rome'

Ano onna noko wa kami o akaku shite iru (2)  
that girl TOPIC hair OBJECT red has  
MARKER MARKER  
'That girl has red hair'

Finally, to return to the problem of assessing a relationship between linearity (order) and focus, the conclusion cannot be avoided that it is deeply affected by typological characteristics of languages. If focus tends to coincide with predication, and this, in turn, is realized at the configurational level by VP, there is no universally valid focus position. Here indeed crucial and difficult problems of syntactic theory are implied, since for languages with VSO order the major constituent VP is split. Furthermore, in SOV languages, where the governed constituents tend to occur on the left of the verb, focus seems systematically to escape the 'VP last' criterion. This in fact holds particularly for SVO languages.

However, it should be noted that a linear criterion may hold under like conditions of government. If a verb governs both an NP and a PP, the last governed major constituent in the VP is the focus (with the last governed lexical category as the focus proper).

A third criterion may impose itself. Semantic criteria may interact with both dependency and linearity. An interesting example is the so-called double-object construction of English. In *I gave a book to Mary* versus *I gave Mary a book*, a role is played by the animacy hierarchy of constituents, according to the principle of 'the most animate first.' Again, the linear criterion seems to determine the focus as the last argument of the verb.

The previous discussion has been concerned with 'unmarked' focus distribution. At this point, however, it must be noted that linear properties of 'marked' focus distribution, that is, the distribution in emphatic sentences, strongly deviate from what has been called the *ordo naturalis* principle, arranging the information increase along a left to right direction. This different distribution conforms rather to an inverse flow of information according to a decrease from left to right in the sentence. Thus, for example, in *To Rome I've never been* or in *These things we like*, focus is placed in  $P_1$ .

So far an attempt has been made at defining focus in terms of linear, structural, and semantic properties. None of these, however, is in itself sufficient. The same is true of suprasegmental correlates of focus, such as variations in pitch, length, and loudness. All these properties can be considered as realization devices of a more abstract notion of a pragmatic nature, that is, 'prominence.' A similar conclusion can be reached for the notion of topic, whose configurational and prosodic properties are envisaged in late twentieth-century literature as devices that codify the pragmatic notion of 'center of attention.'

Another possibility is to think of topic and focus as pragmatic primitives of grammar, with which particular structures can be associated in different languages (this idea has been maintained by Lakoff (1971) and by Dik (1978, 1989)). This could be especially useful in the treatment of configurational regularities concerning order,

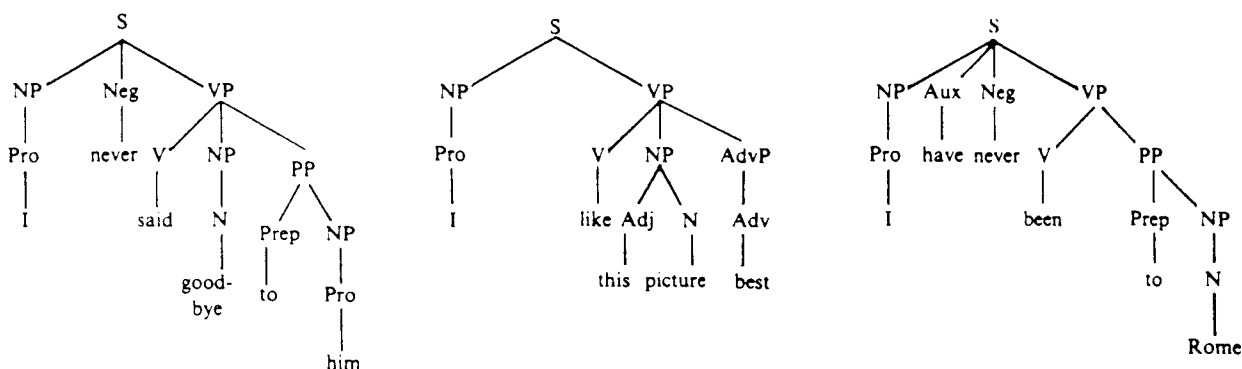


Figure 1a.

which are involved in such phenomena as topicalization, left dislocation, right dislocation, and clefting. However, the approach in terms of pragmatic primitives seems more satisfactory in the framework of a formal theory than for research on empirical properties of natural languages. On the other hand, if topic and focus are defined in terms of configurations, the definition of the associated phenomena becomes rather circular.

The next section combines pragmatic and syntactic approaches in an examination of syntactic structures related to topic and focus.

### 3. Topic-related Structures

#### 3.1 A General Definition

When the syntax of topic is discussed in the literature it is generally both the linear and the argumental criteria which override the pragmatic one. Thus in this section the topic will be identified as the item in one of the initial positions of the sentence which conveys a grammatical function.

Perhaps the most general syntactic property of the class of topic-related structures might be determined as the occurrence in the sentence P (top) position of a constituent with a different grammatical function from that of the subject. It can be argued, in fact, that a hierarchy of accessibility to topic position can be established in natural languages, with subject being the prime candidate for P(top):

Subject > Indirect object > Direct object  
> Locative complement > Manner complement (3)

The above definition can account for sentences such as (4-6):

- To him I never said good-bye (4)  
This picture I like best (5)

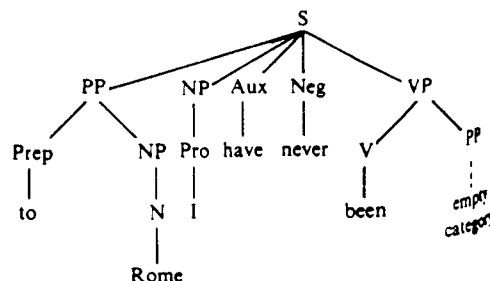
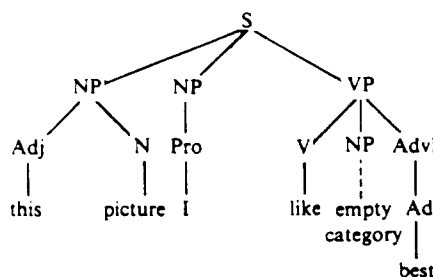
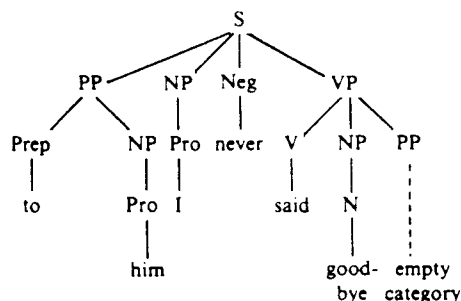


Figure 1b.

To Rome I've never been (6)

whose structures can be thought of as derived respectively from (4'-6'):

I never said good-bye to him (4')

I like this picture best (5')

I've never been to Rome (6')

By adopting a phrase-structure tree of the kind used in generative grammar to represent the base structures (4'-6'), shown in Fig. 1a, it can easily be understood that sentences (4-6) are derived by a movement rule raising the nodes PP, NP, PP to the top node as shown in Fig. 1b.

#### 3.2 Hanging Topic

The definition in example (3), however, needs to be further refined in order to be really workable for the large variety of topic-related phenomena occurring in natural languages. Consider for example the following Mandarin Chinese (7-8), Japanese (9), and Lahu (10-11) sentences (from Li and Thompson 1976: 462):

Nèi-chang huǒ xíngkuì xiāofāng-duì  
that-classifier fire fortunate fire-brigade  
lái de kuài  
come ADV particle quick (7)

'That fire, fortunately the fire-brigade came quickly' (8)

Nèi-xie shùmu shù-shēn dà  
those tree tree-trunk big  
'Those trees, the trunks are big' (9)

Gakkoo-wa boku-ga isogasi-kat-ta  
School-topic marker I-subject marker busy-past tense  
'School, I was busy' (10)

Hè chí tē pē2 5 dà2 jà  
Field this one classifier rice very good  
'This field, the rice is very good' (11)

Hə	ʒ	na-ghə	yī ve	yò	(11)
elephant	TOPIC	nose	LONG	DECLARATIVE	
	MARKER		PARTICLE	MARKER	
'Elephants, noses are long'					

In these sentences the elements occupying position P(top) are ungoverned by the verb (a more technical formulation of this would be that they do not have any selectional relation to the main verb). Thus the basic representation of (7–11) would be as shown in Fig. 2. It seems reasonable

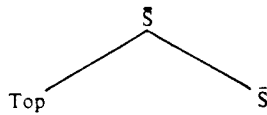


Figure 2.

then to differentiate two fundamental classes of topic-related phenomena, the first having to do with extra-sentential constituents in P(top) which constitute a kind of hanging topic, the second with sentential arguments moved or extraposed to P(top).

Languages may vary as to the permissible range of topicalization. According to a well-known hypothesis by Li and Thompson two main language types can be recognized around the world, topic-prominent and subject-prominent languages. Topicalizations of the kind shown in sentences (7–11) would be permissible in topic-prominent languages, but not in subject-prominent ones. However, the typological distinction per se is rather questionable, since the range of possible topicalization allowed by a given language seems to be related to sociolinguistic parameters such as the speaker's capability to plan the discourse (planned/unplanned discourse), different strategies in spoken versus written communication, the speaker's level of education, and so on. Spoken registers of what would be believed on the evidence of written data to be subject-prominent languages do exhibit structures like sentences (7–11). Examples (12–13) are from a corpus of spoken Italian:

*La scuola, mi sono messo in congedo* (12)  
 The school I have taken a leave  
 'As to the school, I have taken a leave'

*La radio, hanno trasmesso un programma interessante* (13)  
 The radio they have transmitted a program interesting  
 'On the radio they have transmitted an interesting program'

### 3.3 Left Dislocation and Topicalization

The second class (moved sentence arguments) can be further divided into two subclasses, according to the presence (e.g., 14) or absence (e.g., 15) of anaphoric relations between the constituent moved to the top node and a coreferential pronoun filling the position which was previously occupied by it (note, however, that Italian and French sentences such as (15) and (16) would require further observations to justify the position occupied by the coreferential pronoun):

*John Smith, I haven't met him for a long time* (14)

*Cet élève, je l'aime bien* (15)  
 'That pupil, I like her very much'

*Patrizia, l'ho vista ieri* (16)  
 'Patrizia, I saw her yesterday'

*These problems they can't deal with* (17)

*Il decano ho incontrato ieri* (18)  
 The dean I met yesterday  
 'It is the dean I met yesterday'

Structures such as (14–16) are instances of 'left dislocation,' structures (17–18) of 'topicalization' (or 'fronting'). There seems to be, at least in many European languages, a difference in the pragmatic value of structures (14–16) and (17–18): while in the first structures the topic may not convey emphasis (unless specific suprasegmental features are associated with it), in the latter ones it seems the case that the constituent in P(top) has a contrastive value. In other words, it could be said that in sentences such as (17–18) the topic function coincides with the focus function, a peculiar property from the pragmatic point of view, related to marked word order.

From the syntactic point of view, however, rules generating both structures (14–16) and (17–18) have for a long time been recognized as being of a 'movement' nature.

An interesting problem concerns the range of constituents that may undergo movement rules of this kind. Not surprisingly, languages vary as to this range, although a strong tendency can be recognized to put in P(top) nonverbal constituents. In European languages, for example, ADJP, PP, ADVP, in addition to NP, are constituents that are allowed to move to P(top), although there is a difference from language to language as to the pragmatic value associated with the syntactic operation. For example, in both Italian and English, adjective or adverb movement to the top node results in structures conveying emphasis, as can be seen comparing the translations (19–22):

*... e rosso era* (19)

*... and red it was* (20)

*volentieri venne* (21)

*willingly he came* (22)

### 3.4 Raising and Passivization

Left dislocation shows the same property of constituent moving leftward to P(top), as structures which have undergone raising processes. These can be exemplified by English constructions where a NP is 'raised' out of subordinate clause subject position into main clause subject position:

*Mary seems to me to be happy* (23)

Another interesting similarity is shown in passivization processes, which also have the property of moving a NP constituent leftward. It has been recognized for a long time in the literature that passive sentences can be represented as structures derived from the corresponding active ones by means of movement rules (and morphological alterations of the verb, which will not be considered here). From the pragmatic point of view, passivization has convincingly been argued to be but another case of bringing into P(top) (i.e., into the 'center of attention') the NP with the grammatical function object at some underlying level of representation.

### 3.5 Clefting

Clefting encompasses another class of syntactic processes which are related to topic-focus distribution. The term

covers various subtypes of processes, some of them more properly similar to the topic-related ones. These will be referred to as subtype A. Other clefting processes result in structures of the equative kind. These will be referred to as subtype B, to be dealt with in Sect. 4.

Languages tend to differ as to minor details in the structure generated by A-clefting, as will be clear from the following examples from typologically and/or genetically different languages (with examples (24–27) being translations each of the other):

It is me who said that (24)

*C'est moi qui l'ai dit* (French) (25)

*Soy yo quien lo he dicho* (Spanish) (26)

*Sono io che l'ho detto* (Italian) (27)

*Shì wǒ lái zhèr* (Mandarin Chinese) (28)

Be I come here  
'It is me who comes here'

*Kuttiyaṇṇa aanaye nulliyata* (Malayalam) (29)

Child-is elephant pinched-it  
'It was the child who pinched the elephant'

In all these examples the verb 'be' acts as a device assigning the focus function to the NP that immediately follows (cf. 24–28) or precedes (cf. 29). As the occurrence of a dummy subject before the verb 'be' is a feature irrelevant here—it has to do with more idiosyncratic language tendencies—it can be said that the more general A-clefting pattern is represented by (30):

'Be' + NP +  $\bar{S}$  (30)

where NP carries the focus function. Here again, as in the case of topicalization, a single constituent has both the functions of topic ('be' is no candidate for topichood) and focus.

The properties of A-clefting are similar to those of topic-related processes also from a strictly syntactic point of view. The range of constituents that may occur in postcopular position constitutes in fact a class identified by the feature [–verbal]. Note, however, that here again languages seem to vary as to the width of the [–verbal] class, i.e., as to the range of constituents that are allowed to occupy the postcopular position. For example, in languages such as English or Italian only NPs, PPs, and time and place adverbs may occur in that position (English *\*It is yellow that it is* and the Italian translation *\*E' giallo che è*; English *\*It is well that I have found him* and the Italian translation *\*E' bene che l'ho trovato*), whereas in Welsh adverbs and adjectives may also do so:

*Bit chvero y talhaur* (Middle Welsh) (31)  
(It) will be bitterly that (it) will be paid for

More problematic is to decide whether A-clefting should be considered as a movement process (like left-dislocation and topicalization) or not. Examples such as (24–28), where a NP with the grammatical function subject occurs in postcopular position, are not generated by movement rules. Their structures could rather be base-generated with NP in P(top). This is not the case with (32):

It is you that I am looking after (32)

whose structure is obviously related to (33):

I am looking after you

(33)

and could be derived from it by a movement rule.

#### 4. Syntactic Processes that keep the Focus in Unmarked Position

Syntactic processes whose pragmatic import is the focalization of a constituent are less homogeneous than those related to the topic. As a matter of fact, no general definition of focalization processes can be formulated in terms of syntactic configurations, nor in terms of grammatical relations. What could be said would be rather tautological: in unmarked conditions syntactic processes related to focus either keep or move a constituent rightward in the sentence, i.e., in the domain of the unmarked focus (of course this definition would not include structures with marked focus, such as the ones in Sect. 3.5).

The normal distribution of focus as described in Sect. 2 is kept in structures such as *wh*-clefts (34–35):

The one who lies is he (34)

What is in question is his reputation as a scientist (35)

These are sentences with an equative-identifying value, which are generated by type B-clefting. The more general B-clefting pattern can be represented by:

$X + be + Y$  (36)

where both X and Y can be any of the categories in the set (NP,  $\bar{S}$ ). Other equative sentences conforming to pattern (36) are of a kind that is rather frequent in spoken language:

The thing is (that) they are in trouble (37)

The point is that they have never understood the situation (38)

It is worth mentioning that structure (36) has two fundamental properties: (a) reversibility, i.e., its reverse structure (36'):

$Y + be + X$  (36')

is also well-formed (sentences like (37–38) clearly deviate from this regularity); and (b) the marking of X by the feature [–definite] (more precisely, if  $X = NP$ , the head of NP is [–definite]; if  $X = \bar{S}$ , the head of the NP with the subject function is [–definite]).

Sentences conforming to pattern (36) have the focus function stretching over the postcopular constituent. Note that a conflict between the linear and the semantic criteria arises in reverse sentences such as (39):

John is the one who went to Edinburgh (39)

Here in fact the feature [–definite] in the postcopular constituent is incompatible with focality (or less compatible than the feature [+definite]). It can be seen, however, how powerful the linear criterion is, as it overrides the unfavorable semantic feature, assigning the focus function to the postcopular constituent.

An interesting and peculiar effect is obtained in structures with an anaphoric pronoun anticipating either a direct object or an indirect object (or both), as in the following examples from Romance languages:

*La he visto a tu mujer* (Spanish) (40)  
Her I have seen to your wife  
'I did see your wife'

*Gliele raccontai la notizia a Maria* (Italian) (41)  
 To her + it I told the piece of news to Mary  
 'I did tell the piece of news to Mary'

*Je l' ai donné le livre* (French) (42)  
 I it have given the book  
 'I did give the book as a gift'

Here the focus is not on the last sentential argument (*a tu* in (40); *a Maria* in (41); *le livre* in (42)), as in the corresponding sentences without pronominal copies before the verb (40'–42'):

*He visto a tu mujer* (40')

*Raccontai la notizia a Maria* (41')

*J'ai donné le livre* (42')

Rather, it is shifted leftward on the verb.

#### 4.1 Right Dislocation and Extraposition

Structures with right dislocation of the subject, such as (43):

*Hanno considerato il caso molti esperti* (Italian) (43)  
 They have considered the case many experts  
 'Many experts have considered the case'

are instances of syntactic processes moving in focus what normally is a nonfocal constituent. Another case in point is extraposition, as in (44):

A critical review has just appeared of his latest book (44)

where the PP of *his latest book* is detached from NP with the subject function and moved rightward in focus.

### 5. Topic and Focus as Pragmatic Primitives and Syntactic Structure

In this concluding section a few general remarks will be attempted on the relationship between topic, focus, and word order as well as on the nature of topic and focus themselves.

In Sect. 2 it was pointed out that topic and focus are to be considered as pragmatic functions which should be defined neither in terms of syntactic structure nor in terms of prosodic structure. Rather, they should be conceived as independent functions with structural correlates (of a syntactic and prosodic kind). In the light of what has been observed in the last paragraphs two important correlations can be pointed out between topic, focus, and syntactic structure. The first deals precisely with word order. Topic is related to syntactic processes that keep a unit in the leftmost position, out of the sentence pattern proper, or to syntactic processes that result in moving a specific constituent leftward in the sentence. On the other hand, focus is related to syntactic processes that keep or move a constituent rightward in the sentence. This of course conforms to the linear distribution of topic and focus, as has already been pointed out in Sect. 2.

The left-to-right dimension observed, however, is to be connected to two pragmatic notions such as 'center of attention' and 'prominence,' whose import is of a cognitive nature: constituents occurring leftward (leftmost) in the sentence belong to that part of the utterance which is the center of attention for the speaker/listener (finer considerations would be highly desirable as to possible differences

between speakers and listeners in establishing centers of attention; in fact much more experimental work is needed here). On the other hand, constituents occurring rightward (rightmost) in a connected sentence will probably set up prominence peaks. As a matter of fact, it seems a fairly general property of human communication under nonemphatic conditions, to organize the information flow in the utterance according to a strategy of centering attention on specific information units first, and then giving prominence to others. This might well be a universal tendency across natural languages, which accounts for semantic or structural configurations such as topic-comment, subject-predicate, NP-VP, etc.

The second correlation between topic, focus, and syntactic structure concerns: (a) the nonverbal nature of constituents that across natural languages are more frequently allowed to occupy topic position, i.e., the center of attention in the information flow; and (b) the lack of categorial restrictions on constituents in focus. The first property could be formulated in terms of the topic position ruling out elements that bear predication and requiring referential ones. The relationship between topic and referentiality might lead to a cross-linguistic generalization of a pragmatic nature, that is, only constituents with a referential value can function as centers of attention. As to the second property, it should be noted that the lack of categorial restrictions only concerns lexical categories and not major categories: under nonemphatic conditions, VP is in fact the normal domain of the focus function in subject-initial languages. Thus the conclusion can be reached that prominence in principle is to be correlated to predication. Note that this is the case not only when focus stretches over VP, but also in marked structures such as those in Sect. 3.5, where the topic and focus functions collapse together. Predication carried out in this latter case seems to be of a special kind, which may be described as 'identifying.' Sentences with emphatic topicalization or clefting, in fact, can be analyzed as having a double predication pattern, the first associated to the leftmost NP, the second to the following stretch of utterance.

What has been said seems to have some consequences for the study of topic and focus. One of the problems in any attempt to combine pragmatic and syntactic analysis in the study of topic, focus, and syntactic structure has been the fuzziness of notions such as topic and focus themselves. It has already been pointed out in Sect. 2 that if their definition is to be kept separated from syntactic structures, no option remains but to consider them as primitives. This is far from being a satisfactory conclusion however, at least for a full understanding of how phenomena of natural languages work. On the other hand, if topic and focus are given the empirical content of center of attention and prominence, respectively, the problem of further differentiating these two notions is left open. Although an answer to this can be expected to come mainly from psycholinguistic work, the case of topic and focus collapsing together can give some provisional hints. If prominence is defined as an identifying predication, the affinity of the two pragmatic notions will show itself as a consequence. Identifying, after all, is related to referentiality. Thus, center of attention and prominence can be considered as two notions that can be differentiated in degree, being of the same



nature: prominence, in fact, could be conceived as the high degree of attention centering. In this light, syntactic processes such as left-dislocation, topicalization, clefting, and the like are nothing else than effects of pragmatic and cognitive properties of the human mind and human communication.

## Bibliography

- Bolinger D L 1952 Linear modification. *Publications of the Modern Language Association of America* 67: 1117-44
- Bresnan J (ed.) 1982 *The Mental Representation of Grammatical Relations*. MIT Press, Cambridge, MA
- Carlson L 1983 *Dialogue Games*. Reidel, Dordrecht
- Cinque G 1977 The movement nature of left dislocation. *Linguistic Inquiry* 8: 397-412
- Chafe W L 1970 *Meaning and the Structure of Language*. University of Chicago Press, Chicago, IL
- Chafe W L 1976 Givenness, contrastiveness, definiteness, subjects, topics and point of view. In: Li Ch N (ed.) *Subject and Topic*. Academic Press, New York
- Chomsky N A 1965 *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA
- Chomsky N A 1981 *Lectures on Government and Binding*. Foris, Dordrecht
- Daneš F 1967 Order of elements and sentence intonation. In: Bolinger D L (ed.) *Intonation*. Penguin, Harmondsworth
- Dik S C 1978 *Functional Grammar*. North Holland, Amsterdam
- Dik S C 1989 *The Theory of Functional Grammar*. Foris, Dordrecht
- Firbas J 1964 From comparative word-order studies. *Brno Studies in English* 4: 111-28
- Firbas J 1987 On the operation of communicative dynamism in functional sentence perspective. *Leuvense Bijdragen* 76: 289-304
- Gazdar G, Klein E, Pullum G-K (eds.) 1983 *Order, Concord and Constituency*. Foris, Dordrecht
- Givón T 1984 *Syntax: A Functional Typological Introduction*. Benjamins, Amsterdam
- Halliday M A K 1961 Categories of the theory of grammar. *Word* 17(3): 1-7
- Halliday M A K 1967-68 Notes on transitivity and theme in English. *Journal of Linguistics* 3(1): 37-81; 3(2): 199-244; 4(2): 179-215
- Jackendoff R S 1972 *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA
- Lakoff G 1971 On generative semantics. In: Steinberg D D, Jakobovits L A (eds.) *Semantics*. Cambridge University Press, Cambridge
- Li Ch N, Thompson S A 1976 Subject and topic: A new typology of language. In Li Ch N (ed.) *Subject and Topic*. Academic Press, New York
- Marantz A 1984 *On the Nature of Grammatical Relations*. MIT Press, Cambridge, MA
- Mathesius V 1939 O tak zvaném aktuálním členění větém. *Slovo a slovesnost* 5: 171-74
- Mathesius V 1941-42 Ze srovnávacích studií slovosledných. *Časopis pro moderní filologii* 28: 181-90; 302-07
- Meillet A P J 1965 *Linguistique historique et linguistique générale*. Champion, Paris
- Šaumjan S K, Soboleva P A 1963 *Applikativnaja poroždajščaja model' i iščislenie transformacij v russkom jazyke*. Akademik Nauk, Moscow
- Sgall P 1972 Topic, focus and the ordering of elements in semantic representation. *Philologica Pragensia* 15: 1-14
- Sgall P, et al. 1986 *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Reidel, Dordrecht

Tesnière L V 1958 *Elements de Syntaxe Structurale*. Klincksieck, Paris

R. Sornicola

## Totonac

Totonacan is a Middle American language family comprising the daughter languages Totonac and Tepehua, spoken in the central Mexican states of southern Hidalgo, northern Puebla, and northwestern Veracruz along the Gulf of Mexico. The only fact known about the pre-Conquest speakers of Totonac is that they entered this homeland around 900 AD; their whereabouts beforehand and their activities in the region subsequently are unknown, although they are surmised to have been the builders of the important archeological site, Teotihuacan. During the Spanish conquest, they allied themselves to Cortez against their old enemies, the Aztecs.

### 1. Genetic Relationship

In the early 1990s, comprehensive comparative analyses of the Totonacan languages remain to be completed; however, a reconstructed phonemic inventory of the protolanguage does exist, listing /p t ts tʃ tl k q s ʃ ɬ x m n w y l; i u a i; u; aɪ/. A remote genetic affiliation with Mayan and Mixe-Zoque has been proposed as 'Macro-Mayan'; some researchers also include Huave in their classifications as 'Mexican-Penutian' (a subcategory of the controversial 'Penutian'). These aggregations are not accepted by most scholars. At this stage in the research, it seems safest to conclude that Totonac has not been demonstrated to have any outside genetic relationship (see *Language Classification*).

### 2. Linguistic Structure

Totonac boasts a highly productive synthetic morphology (see *Morphology, Polysynthetic*). A single verb frequently functions as an entire clause: for example, Northern Totonac *ϕ-nik-putu-ϕ-n* 'he wanted to hit you' (3SG-to hit someone-DESIDERATIVE-ASPECT [and implicitly past tense]-2SG OBJ); note that the pronominal subject and object are included in the verb complex. Verbs are composed of roots and numerous affixes which convey tense, aspect, various adverbial notions, person and number of both subject and object, and degrees of transitivity (see *Transitivity*). Nouns, too, are frequently compounds, for instance of two simplex nouns: for example, /laka-stapu/ 'pupil of eye,' literally 'eye-bean' (Tlachichilco Tepehua). Complex phonological processes of assimilation then affect the derived verbs and nouns. Word order is largely governed by pragmatic considerations and is rather free, but tends to be subject-verb-object. The phonemic inventory lacks voiced stops and affricates; it does not differ significantly from that reconstructed for the protolanguage. Sound symbolism to signify degrees of intensity occurs between such consonant sets as /ɬ, s, ʃ/, with the first item being most intense; for example, compare /hunpu:ɬukuk/ 'there is a big hole' and /hunpu:sukuk/ 'there is a small