

## CLIPP

**Christiani Lehmanni inedita, publicanda, publicata**

titulus	Word order change by grammaticalization
huius textus situs retis mundialis	<a href="http://www.uni-erfurt.de/sprachwissenschaft/personal/lehmann/CL_Publ/word_order.pdf">http://www.uni-erfurt.de/sprachwissenschaft/personal/lehmann/CL_Publ/word_order.pdf</a>
dies manuscripti postremum modificati	28.02.2009
ocasio orationis habitae	9 <sup>th</sup> International Conference on Historical Linguistics, Rutgers University, NJ, 12.-17.8.1989
volumen publicationem continens	Gerritsen, Marinel & Stein, Dieter (eds.), <i>Internal and external factors in syntactic change</i> . The Hague & Berlin: Mouton-de Gruyter
annus publicationis	1992
paginae	395-416

# Word order change by grammaticalization

Christian Lehmann  
University of Bielefeld

## Abstract

Traditionally, word order has been regarded as one of the devices for the expression of grammatical meaning, alongside with grammatical words, affixation, internal modification, and suprasegmental means. This view has blocked the way towards an adequate understanding of the role of word order in language. Every linguistic sign has a certain degree of autonomy, realized in its manipulability on the paradigmatic and syntagmatic axes. Viewed from this angle, greater or lesser order freedom of a sign is one of the aspects of its autonomy.

Autonomy in this sense is reduced by grammaticalization. Grammaticalization therefore leads to fixation of word order. While old constructions are grammaticalized and fixed at lower grammatical levels, new ones are introduced at higher levels. Thus, while inversion or laxation of the sequential order in a given construction is impossible (this would be degrammaticalization), the construction and its order can be renewed by creating a functionally similar construction at a higher level.

These principles are illustrated from word order change in a variety of languages. Verb-initial order in colloquial German is shown to be a consequence of the extreme grammaticalization of the expletive *es/da* construction. The same order in colloquial French is a consequence of the grammaticalization of the suspension of the theme. The conclusion is that just as word order is only an aspect of the autonomy of the language sign, word order change may be just an aspect of grammaticalization.

## 1. Introduction

In traditional historical-comparative linguistics, syntactic change played a relatively minor role. What is now seen to form the center of syntax by many linguists, viz. order of words or constituents, figures but marginally in the works of K. Brugmann, B. Delbrück, H. Paul and also of most of the modern Indo-Europeanists. This has various reasons. One of them lies in the prevailing interest of received Indo-European linguistics in the reconstruction of Proto-Indo-European. For reasons that have been discussed repeatedly in the literature (cf. Vincent 1980), reconstruction in the area of syntax is less certain than in the other areas of the linguistic system. Moreover, most of the ancient Indo-European languages have free word order. The most simple and probably correct hypothesis is that this was also true of Proto-Indo-European. This, however, entails that no revolutions in the rules of word order took place in the span bridged by reconstruction, namely the span between Proto-Indo-European and the specific Indo-European languages, which is of greatest interest to researchers. Crucial changes took place only at stages with which historical-comparative linguistics is but marginally concerned, namely on the way from the ancient to the modern

Indo-European languages. These historical changes are relatively well-known individually, but not yet sufficiently integrated into a general-comparative theory of language change.

A completely new approach to the investigation of word order change was taken in the sixties in general-comparative linguistics. In his epochal work of 1963, J. Greenberg showed that the sequential position of linguistic signs, in particular main constituent order, is subject to implicational laws. In Greenberg 1969, he demonstrated how these laws can be dynamicized, i.e. how they are to be applied not only in synchrony but also in diachrony. These so-called diachronic universals appeared to many linguists to be an appropriate basis for a principled description of syntactic change, esp. word order change. Diachronic basic word order typology was promoted especially in various articles by Th. Vennemann (1974) and W.P. Lehmann (1973). In diametrical contrast to traditional historical-comparative linguistics, these authors saw the syntax, especially word order, in the center of the functioning of the language system. They tried to base language types on order types and to reduce grammatical change to order change.

These attempts may be considered failed today; cf., e.g., Comrie 1981, ch. 10.3 and Ch. Lehmann 1982 for a thorough criticism. One of their methodological mistakes was over-simplification paired with immunization against falsification. For the present purpose, a theoretical mistake is of greater interest. It lies in the isolation and hypostatization of one grammatical phenomenon, namely word order. There was practically no reflection on the role of word order in the linguistic system, on its conceptual and systematic status and, further, on its connection with other grammatical phenomena.

## **2. Word order in the syntactic system**

In traditional interpretation, word order is one of the devices for the expression of grammatical meanings. I will assume that grammatical meanings are meanings and need not be called functions or anything else. H. Paul (1920:123) gives the following list of devices for the "linguistic expression of the connection of ideas", which I render in slightly modernized terminology.

1. Juxtaposition of words,
2. sequential order of words,
3. sentence stress,
4. intonation,
5. tempo,
6. grammatical words,
7. inflection.

E. Sapir (1921:61) sets up the following six main types of grammatical processes:

1. word order,
2. composition,
3. affixation,
4. internal modification,
5. reduplication,
6. stress and pitch.

The differences between the two lists derive partly from the fact that Paul refers only to the means for the linguistic expression of the connection of ideas, whereas Sapir deals, more generally, with grammatical processes. It should be noted that Paul distinguishes, inside sequencing, between mere juxtaposition and order. What is important at the moment, however, is merely that both authors class sequential order as one of the grammatical expression devices beside others like prosody and morphological modification (inflection). Even today it is customary to say of particular syntactic relations such as the relation of the direct object that they can be expressed either by word order (for example, in English) or by inflectional morphology (for example, in Latin).

If we consider the syntagmatic position of linguistic signs in a broader context, this conception becomes questionable (cf. Lehmann 1985, 1). Commonly a distinction is made between fixed and free word order. By fixed word order is meant an order that is determined by purely grammatical rules, i.e. by rules that refer exclusively to the categories of the linguistic signs involved and their mutual syntactic relation. By free word order is meant an order which is not subject to grammatical rules, but determined by more semantic considerations or by considerations of functional sentence perspective (called 'pragmatics' by some). For example, the position of the main verb in declarative sentences is fixed in German, as it is completely determined by rules which refer to the categories of the sentence (the sentence type) and of the verb. In Latin, on the contrary, it is free, because it is not determined by such rules and instead depends essentially on the theme-rheme-structure and the focus-background-structure.

However, it should be clear that the distinction between fixed and free word order is not an absolute one. On the one hand, there are possibilities in German of varying the verb position even within one and the same sentence type - I will come back to this in 6. On the other hand, even in Latin there is a tendency, which is independent of functional sentence perspective, towards final position of the verb. The basic idea of this division must therefore be that a word order pattern is the more fixed the more it is subject to rules of grammar.

Expression of meaning is bound up with the possibility of choice. If I am forced by the rules of the system to use a certain expression, then I cannot convey a meaning by it because it does not contrast with any other expression. Hence it becomes doubtful in which sense it shall be said of a fixed word order that it serves the expression of some meaning. In this framework, this appears to be possible only with free word order. However, a moment ago we saw that free word order does not, by definition, express grammatical relations, but instead semantically specific relations or relations of functional sentence perspective. In other words, it is inappropriate to say of fixed word order, which is bound up with grammatical relations, that it serves the expression of meanings; and at the same time free word order, although capable of expressing meanings, is not at the service of grammar. Hence the status of word order, be it as a device for the linguistic expression of the connection of representations (Paul), be it as a grammatical process paired with grammatical concepts (Sapir), becomes questionable.

We can now briefly come back to the two aspects of word order which were distinguished by H. Paul. It is easy to see how limited is the possibility of expressing anything by mere juxtaposition of signs (Paul's no. 1). The signs cannot but follow each other in the chain, whether or not they bear a direct grammatical relation to each other. Therefore, contiguity in the chain cannot by itself signify grammatical relatedness. Instead, the categories of the words in question and their relational potential have to be taken into account. For instance, a noun and an adjective carry their categories qua lexical entries and bring them into play when they are actualized. An adjective has, moreover, the relational potential to combine, as an attribute, with a noun. Only after all this is presupposed can we interpret a sequence of an adjective and a noun, for instance in English, as a nominal containing an attribute. Contrariwise, we will never interpret a sequence of an adjective and a conjunction as a phrase, because the grammatical presuppositions are not met.

In other words, the juxtaposition of two signs *a* and *b* cannot tell more than this: the grammatical relation to a sign of category *B* which is set up in sign *a* is actualized in case that a sign of category *B* is positioned immediately beside *a*, otherwise not. Positioning of *b* beside *a* therefore means only the presence or availability of *b* for the grammatical relation set up in *a*; it cannot express the relation itself. This is true both if the relative order of *a* and *b* is invertible and, a fortiori, if it is fixed.

We have thus arrived at Paul's second aspect, relative order. If the order of *a* and *b* is invertible, two cases have to be distinguished. Either the grammatical relation between *a* and *b* remains constant under inversion. This means, of course, that the grammatical relation is independent of order and, consequently, is not expressed by order. The order of adjective and noun in Italian would be an example. Or else the grammatical relation becomes a different one under inversion. Only in this case does the order express, or rather help to express, the grammatical relation. I am aware of only one case of this kind, viz. the position of an NP vis-à-vis the verb in languages like English. Here the NP is subject if it precedes the verb, and it is object if it follows. But even in this case it is not word order alone which signals the grammatical relation; with most verbs, selection restrictions bear the main burden in the assignment of actant functions.

The conclusion of this argument is that word order does not have a positive expressive function, but much more a negative, oppositive function. It does not, like a grammatical affix, signify a particular grammatical relation; instead it functions rather like a phoneme by admitting or excluding a given grammatical relation.

### **3. Word order and inflectional morphology**

Since one has started comparing ancient and modern Indo-European languages, in particular Latin and the Romance languages, the common-place has perpetuated itself that a rich inflectional morphology allows free word order and that lack of inflectional morphology requires fixed word order and is, in regard to expressive possibilities, compensated for by the latter. It is said that the relative weight of the two expressive devices was turned round in the development from Latin to the Romance languages. Since a long time, this view has been generalized in linguistic typology. Thus

Mallinson & Blake (1981) repeatedly speak of a "trade-off" between inflectional morphology and word order in the expression of syntactic relations.

Given what we have seen in the preceding section, this conception cannot be true as it stands. On its way into the Romance languages, word order became indeed more fixed in many relations. Limitation of alternatives, however, means reduction of conveyable information. Far from word order taking on the tasks of the lacking morphology, its own expressive capacities - if indeed it has any - were confined, too.

At the typological level, there is another disconcerting observation to be made. The following can be established as an implicational tendency: if a language has free word order, then it has rich inflectional morphology. This is empirically confirmed in most cases (see Mallinson & Blake 1981, ch. 3.4 for possible exceptions). The inversion of this law, however, is invalid: it is not the case that if a language has rich inflectional morphology, it has free word order. This would be falsified by numerous languages with dominantly agglutinative morphology, e.g. by Turkish and Yucatec Maya. It thus becomes evident that there can be no question about a mutual compensation between word order and inflectional morphology. The unilateral implicational relationship between the two phenomena indicates that they are not hierarchically equal, but instead in a dependency relationship.

On the diachronic axis, the problem presents itself as follows: Assuming that loss of inflectional morphology goes hand in hand with fixation of word order, and acquisition of inflectional morphology allows the introduction of freer word order; what then are the mechanisms by which this comes about? If this problem is solved, answers to more specific questions like the following can be expected: If free word order is lost, do the functions formerly fulfilled by it get lost, too? If free word order is acquired, are formerly rigid phrases broken up and become variable? Or how else is the acquisition of free word order to be conceived?

#### 4. Grammaticalization

There is one mechanism chiefly responsible for the acquisition and loss of inflectional morphology, grammaticalization. This is clearly shown by the well-known stock examples of grammaticalization. The Romance languages have replaced the inherited future by a new one and introduced a completely new verbal category, the conditional, both of which are based on the infinitive and whose inflection morphemes are grammaticalized from Latin *habere* 'have' ((1)). These are therefore cases of the evolution of inflection by grammaticalization.

- (1) a. Vulgar Latin *cantare habet*      'he has to sing'  
       > Italian *canterà*                'he will sing'  
       b. Vulgar Latin *cantare habuit*    'he had to sing'  
       > Italian *canterebbe*            'he would sing'

Moreover, the Romance languages have, in concomitance with the grammaticalization of prepositions, reduced the inherited case paradigm and finally lost it; cf. (2). French and, to a lesser extent, other Romance languages (cf. Haiman, in this vol.) have, in concomitance with the grammaticalization of free personal pronouns, reduced the paradigm of the verbal personal endings; cf. (3). These cases therefore illustrate the loss of inflection by grammaticalization.

(2) Latin *flos/floris/flori/florem/flore* > Italian *flore*

(3) Latin > French  
 canto /ʃatə/  
 cantas /ʃatə/  
 cantat /ʃatə/  
 cantamus /ʃatõ/  
 cantatis /ʃate/  
 cantant /ʃatə/

In several agglutinative languages such as Yucatec Maya, it can be shown historically or by internal reconstruction that inflectional morphology has developed by grammaticalization and coalescence of formerly free words or morphemes. (4) exemplifies the grammaticalization of terminative aspect.

(4) *ts'ook* *u* *k'aay* > *ts'u* *k'aay*  
 YUC finish(CMPL.3.SG) SBJ.3sing TERM:SBJ.3 sing  
 'his singing finished' > 'he has sung'

On the other hand, the advanced loss of inflectional morphology such as observable in English, but also in some Slavic languages, is based on a continuation upto the zero point of the reduction by grammaticalization.

Grammaticalization is a phenomenon that has received intensive study especially in the last years. To the extent that morphological change can be taken as a case of grammaticalization, chances are favorable that its mechanisms will be understood. Our initial question, however, was how change of word order is related to morphological change. If grammaticalization is the transformation of a lexical morpheme into a grammatical one and the further reduction of the grammatical morpheme to the point of loss, what then does this have to do with word order? A narrow conception of grammaticalization which is limited to the fate of morphemes obviously does not help us in the attempt to pin down the relation between word order and inflectional morphology.

And indeed, grammaticalization is not something that just affects a morpheme in isolation. Grammaticalization concerns linguistic signs as wholes, i.e. in their paradigmatic and syntagmatic relations. A linguistic sign may be relatively autonomous, i.e. it may have relatively few or loose paradigmatic and syntagmatic links; or it may be more subject to grammatical rules, i.e. it may be confined in its paradigmatic and syntagmatic relations. Grammaticalization is, in fact, loss of autonomy in this sense.

This conception does not neglect the palpable characteristics of grammaticalization, viz. loss of concrete semantic content and of phonological substance. Limitation of the paradigmatic relations of a sign actually means that it contracts such relations to ever fewer other signs. This entails that its distinctive features, both on the semantic and on the phonological sides, decrease.

Hence, the conception of grammaticalization as the reduction of a lexical to a grammatical morpheme is only part of the truth.

(5) a. *cantare habeo* 'I have to sing'  
 ROM b. *(je) chanterai* '(I) will sing'

If Latin *cantare habeo* is grammaticalized to French *chanterai* ((5); cf. also (1).a), then this does not boil down to a reduction of *habeo* ‘have to’ to *-ai* FUT.1.SG. Instead, the former *habeo* becomes, viewed paradigmatically, an element of the paradigm of tense elements; and viewed syntagmatically, the construction is grammaticalized as a whole, i.e. a new synthetic verb form arises.

One of the aspects of the autonomy of a linguistic sign in its syntagmatic relations is its syntagmatic variability, i.e. its order freedom. From what was said above it is evident that this gets lost through grammaticalization. In Latin, one could say either *habeo cantare* or *cantare habeo*. In French, there is only *(je) chanterai*, not *(j’)ai chanter*. The same can be shown for the grammaticalization of the Latin demonstrative *ille* to the definite article and to the clitic personal pronoun in Romance languages. The demonstrative adjective could precede or follow the determined nominal; the definite article of the Romance languages is bound to one side of the nominal. The demonstrative pronoun of Latin had the entire order freedom of any NP; the clitic personal pronoun of the Romance languages is bound to the verb, and even the position either before or after the verb is subject to grammatical rules. To the extent that a morpheme is grammaticalized, its order freedom is reduced. By the time it has become an inflectional affix, its position is completely fixed within the morphological slots of its host.

## 5. Fixation of word order

However, when we speak of grammatical order, we are generally not referring to the sequencing of grammatical formatives. Instead, we mostly refer to the sequencing of full words or even of complex phrases. Thus, it remains to be seen whether this, too, has any relation to grammaticalization. Here, again, the broader context has to be taken into account. Let us consider the grammaticalization of adpositions as an example. As I lack sufficient historical evidence from one language, I will have to put the argument together from various languages.

An adposition such as German *wegen* ‘because of’ has a primary syntagmatic relation to the NP governed by it and a secondary relation to the controlling phrase - in the most simple case a VP - that is modified by the adpositional phrase. (6) is an example.

(6) *Wegen seines Argwohns / seines Argwohns wegen glückt dem Alten nicht viel.*

GERM ‘Because of his mistrust, the old man does not succeed in much.’

In its primary relation, this adposition has a certain syntagmatic variability insofar as it can be used alternatively as a pre- or as a postposition. It shares this property with some other German adpositions such as *nach* ‘after’ and *entlang* ‘along’. These have a relatively low degree of grammaticalization in common. All of the more strongly grammaticalized adpositions of German, such as *von* ‘from, of’, *zu* ‘to’, *an* ‘at’, *in* ‘in’ etc., are exclusively prepositions.

There are, of course, other structural correlates to this differential degree of grammaticalization of the German adpositions. First, the strongly grammaticalized adpositions coalesce with the definite article to such forms as in (7), while the weakly grammaticalized adpositions such as *wegen*, *nach*, *entlang* do not do this. This is another clue to the greater syntagmatic autonomy of the latter.

- (7) *von dem* > *vom* 'of the'  
 GERM *zu dem* > *zum* 'to the'  
*an dem* > *am* 'at the'  
*in dem* > *im* 'in the'

Second, *von*, *zu*, *an*, *in* serve for the expression of purely grammatical relations (such as the genitive, the infinitive complement, the (dialectal) progressive with *am* etc.), while *wegen*, *nach*, *entlang* express semantically specific relations. Third, the reduced and relatively homogeneous phonological form of the grammatical prepositions as opposed to the fuller and more multiform shape of the semantically more concrete prepositions should be noted.

Adpositions can grammaticalize to case affixes. At this stage at the latest, their position in their primary relation is entirely fixed. Some linguists have already regarded French *à* 'to' and *de* 'of' as case prefixes. This is, however, debatable, and in general case prefixes are extremely rare among the languages of the world. The grammaticalization of postpositions to case suffixes, on the other hand, is a common process which is observable at historical times in languages such as Turkish and Tamil. Thus we have here another example of the total loss of order freedom by grammaticalization.

Interested as we are in the relation between full words, we now have to consider the secondary relation of the adposition. The adpositional phrase is relatively free in its position vis-à-vis the controlling VP. (6) allows, among others, the permutations in (6)'. All of them are independent of the pre- vs. postpositional use of the adposition.

- (6)'. a. Dem Alten glückt wegen seines Argwohns nicht viel.  
 b. Dem Alten glückt nicht viel wegen seines Argwohns.

That is, apart from permutations of the other constituents of (6), the adpositional phrase can occupy all of the syntactic positions except, of course, for the position reserved for the finite verb. Hence the secondary relation of the adposition is syntagmatically freer than the primary one. However, with the grammaticalization of the adposition, freedom in the secondary relation decreases just as freedom in the primary relation. It is true that this does not, in German, manifest itself in syntagmatic variability, since all adpositional phrases enjoy the same order freedom. It can, however, be perceived in the case government of verbs. Just as verbs can govern certain cases, there are verbs which govern grammatical prepositions, e.g. *abhängen von* 'depend on', *zählen zu* 'belong to', *liegen an* 'be due to', *irren in* 'err in'. There is, however, no verb which governs a preposition such as *wegen*, *nach*, *entlang*. The difference between free adpositional phrases and adpositional phrases functioning as a complement manifests itself in other languages in differential order freedom. This can be seen in (8) and (9).

- (8) a. We danced on the boat.  
 b. On the boat, we danced.  
 (9) a. We decided on the boat.  
 b. \*On the boat, we decided.

The adjunct in (8) is permutable, the complement in (9) is not. Given that only grammatical prepositions occur in more grammatical, e.g. complement, relations, such

a situation can also be adduced as an argument in favor of the reduction of syntagmatic variability in the secondary relation of a grammaticalized adposition.

Once the adposition has become a case affix, its secondary relation passes over entirely to the host NP. Consider (10) and (11).

(10) a. *Caesar scripsit epistulam ad Hannibalem.*

LATIN 'Caesar wrote a letter to Hannibal.'

b. *Caesar scripsit epistulam Hannibali.*

'Caesar wrote Hannibal a letter.'

(11) *Cesare scrisse una lettera ad Annibale.* (idem, ITALIAN)

While the Latin version (10).a features a free adpositional phrase, version b has a dative NP which is a complement of the controlling verb. The Italian version (11) is situated, in regard to grammaticalization, between the two Latin examples, since the preposition *a* can already express the pure dative, but is not (yet) an affix. We can, therefore, set up an -achronic - continuum leading from (10).a via (11) to (10).b. What is important for the present purpose is the following: At the end of this continuum, we no longer have a grammatical element (here, the dative suffix) which by itself contracts a primary and a secondary grammatical relation. Instead, its primary relation has become a morphological one, namely the relation of the nominal case affix to the noun, while its secondary relation has passed over to its host: *Hannibali* has the relation of indirect object to *scripsit*. At this stage, we are therefore dealing with an immediate relation between two full words or potentially complex phrases.

As mentioned before, further grammaticalization of the case element now leads to its loss. At the same time, the relation between the host of this element and the controlling term becomes ever tighter. In English, for instance, the relation of the erstwhile (Germanic) indirect object acquires properties of a direct object and becomes available for passivization, as in (12).

(12) a. Cesar wrote Hannibal a letter.

b. Hannibal was written a letter by Cesar.

At the same time, order freedom is severely restricted here. If compared with (8) and (9), (12).a is more like the latter because we cannot front *Hannibal* in (12).a. Hence, in the course of the grammaticalization of the case element to zero not only its own order freedom, but also the order freedom of its host has decreased.

The result of this case study of the grammaticalization of adpositions is the following: The grammaticalization of a sign is bound up inseparably with the reduction of its syntagmatic variability. This means that grammaticalization does not merely seize a word or a morpheme - namely the one which it reduces to a grammatical formative and finally to zero -, but instead the whole construction formed by the syntagmatic relations of the element in question. To the extent that the external relations of this construction are contracted by the grammaticalized formative, they are also seized by the grammaticalization process. Consequently, with the grammaticalization of a bound morpheme the syntagmatic variability of its host shrinks, too. Thus, the fixation of any word order can be a consequence of grammaticalization.

## 6. Acquisition of word order freedom

There are languages such as Amharic that formerly had prepositions, but have passed, by syntactic change, to using postpositions. For an early stage of Proto-Indo-European, postpositions may be posited (or else the case suffixes could hardly have evolved); but many Indo-European languages have prepositions. How is such a word order change to be conceived? It would be logically possible that one day freedom in the positioning of the erstwhile postpositions set in so that they could be used alternatively as prepositions. The syntagmatic variation thus developed would, in the further course of things, have been reduced to the prepositional variant. This would mean, however, that syntagmatic variability increased in the first phase of this development. This would run counter to the sense of grammaticalization, where syntagmatic variability decreases; it would be a case of so-called degrammaticalization. As a matter of fact, such cases are not known.<sup>1</sup> Rigid order restrictions cannot be directly loosened. They can only be substituted by a freer order when the construction in question is renewed. Hence, it is impossible that the same adpositions that formerly were postpositions become prepositions one day. What is possible, however, is that the old postpositions are reduced by grammaticalization and simultaneously prepositions evolve from a new source, which by and by replace the old postpositions. Something of this sort must have happened in the prehistorical development from Proto-Indo-European to Latin.

An example from historical times, and from a different grammatical domain, is offered by main constituent order in modern French as illustrated in (13).

- (13) a. *Jean est venu.* 'John has come.'  
 FRENCH b. *Il est venu, Jean.* 'He has come, John.'  
 c. *Il est venu Jean.* 'John has come.'

Positioning of subject and verb in this order at the start of the clause, as in (13).a, is obligatory in written French. There is, however, the possibility of forming a complex sentence by right-dislocation of the topic, as in (13).b. This formerly complex construction is reinterpreted as a simple clause in spoken French ((13).c). That is, the pragmatic markedness - suspension of the theme - falls away, and so does the intonation break as signalled by the comma. The personal pronoun loses its referential function and becomes an agreement marker. What was formerly a right-dislocated topic now occupies the subject position. All of this means that a change in main constituent order takes place from (13).a to c, from initial position of the subject to its position at the end of the clause (cf. Bailard 1982).

This example shows very clearly that the order regularities of a particular construction are not simply dissolved or inverted. Instead, a new source is tapped for the formation of a construction that is functionally similar, if, for the time being, less grammaticalized, and all the while the existing order regularities are observed. The new construction with its variant word order can then oust the traditional one.

In German, a similar case is observable. Written German has a rule for main constituent order in independent declarative clauses, which says that the finite verb occupies the second position. This is the strictest rule at this level; in other words, overall order at the sentence level is relatively free. Nonetheless this rule confines the speaker. One cannot start the sentence with the verb, although one might want to do

so in presentative sentences. Nor can one defer the verb to a later position, although one might want to do so for the attainment of rhematic effects.

However, this rule can be circumvented by fulfilling it to the letter. As to the second sentence position, the rule merely demands its filling with a finite verb form. If one wants to defer the verbal lexeme to a later position, one fills the second position with an auxiliary, as in (14).b.

- (14) a. *Das Auto schleuderte ihn über den Zaun.*  
 GERM 'The car hurled him over the fence.'  
 b. *Das Auto hat ihn über den Zaun geschleudert.*  
 'The car has hurled him over the fence.'

For the time being, however, the development potential of this alternative is limited by the fact that all the relevant auxiliaries form specific tense/aspect/mood/voice categories. Consequently, if one chooses this functional sentence perspective, one chooses, at the same time, the verbal categories which are expressed by auxiliaries.

As to filling the first sentence position, the rule only demands a non-verbal constituent. If one wants to put all major constituents after the verb, one occupies the first position with a semantically empty dummy element, as in (15) and (16).

- (15) *Es weiß eben niemand so recht, was das heißen soll.*  
 GERM 'It is the case that nobody knows quite clearly what this is supposed to mean.'  
 (16) *Da soll mich doch gleich der Schläg rühren.*  
 GERM Lit.: 'The stroke shall hit me right away.'; i.e. 'I am perplexed.'

German grammar provides two such dummy pronominal elements, *es* 'it' ((15)) and *da* 'there' ((16)). Although they are interchangeable in some contexts (both of the examples would be possible with the other element, but not so idiomatic), the general regularity is that *es* is cataphoric for an NP, while *da* vaguely refers to some spatio-temporal or preceding textual context.

This strategy is exploited and further developed in the colloquial language. Given that the pronominal forms at the beginning of the sentence are nothing more than dummies, they are redundant and can fall away. Sentences such as those in (17) are very common at the colloquial level.

- (17) a. *Könnte ja jeder kommen.*  
 GERMAN 'Then anybody could come.'  
 b. *Hastu [hast du] eben Pech gehabt.*  
 'You just have had tough luck.'  
 c. *Mußtu [mußt du] halt noch mal hingehen.*  
 'You just have to go there once again.'  
 d. *Müßte man mal drüber nachdenken.*  
 'One should think about this.'

However, the position in front of the finite verb is not simply left unfilled in such sentences - that would be a violation of a grammatical rule. Instead, for the time being all such sentences are colloquially elliptical versions of standard sentences which have the dummy *da* in first position. That this is the correct interpretation is confirmed by the fact that the dropping of initial thematic elements is extended to other pronominal elements, as illustrated in (18).

- (18) a. *Is[t] ja noch mal gút gegangen.*

- GERMAN 'It has turned out well once again.'
- b. *Wird schon nicht so schlimm werden.*  
'It won't get that bad.'
  - c. *Kann dir doch egal sein.*  
'That should be of no concern to you.'
  - d. *Weiß ich nicht.*  
'I don't know.'
  - e. *Kann man nie wissen.*  
'You never know.'
  - f. *Ham [haben] wir ja noch nie gesehen.*  
'That we have never seen before.'
  - g. *Müßte man mal näher untersuchen.*  
'That should be investigated more closely.'

These sentences lack an initial *das* 'that' which would function as the subject in (18).a - c, but as the direct object in d - g. This shows that the motivation for this tendency is essentially non-syntactic. Cf. also the parallelism between (17).d and (18).g. On the other hand, there appear to be some syntactic constraints. For instance, *das* in other syntactic functions cannot be omitted, as shown in (19).

- (19) a. *\*(Dem) müßte man mal nachgehen.*  
GERMAN 'That should be looked into.'
- b. *\*(Dessen) müßte man sich mal annehmen.*  
'That should be taken care of.'

Anyway, the situation opens the possibility for a further step in which such constructions are reanalyzed as simple declarative sentences with initial verb. As a result, German would acquire an alternative main constituent order.

This is not mere speculation. Sentences such as the above are already fairly common in the colloquial language. Moreover, they exhibit a common structural tendency: the initial verb is a modal or auxiliary in the far majority of cases (*können* 'can' in (17).a, *müssen* 'must' in c, *sein* 'be' in (18).a, *haben* 'have' in (17).b and (18).c, *werden* 'will' in (18).b). This is even true for (16) (*sollen* 'shall'). This means that the construction is preferred in those cases in which everything including the verb belongs to the rheme of the sentence. In those Indo-European languages which have passed through the analogous change since a long time, namely the Celtic languages, it is also preferably the auxiliary which starts the sentence.<sup>2</sup> And while this change or at least this tendency comes up in the colloquial language of German, something similar is happening in French and in Modern Greek. Possibly we are faced with a genetic-typologically conditioned drift.

Hence, an order pattern is not directly twisted into a different one. Instead, its alternative first has to be introduced by the grammaticalization of an originally more complex construction. This complex construction completely obeys the rules of established grammar, but undermines them by the artifice of providing mere grammatical elements for the crucial syntactic positions, which lose their word status with further grammaticalization of the construction or are even reduced to zero. As a consequence, they no longer occupy a syntactic position, and the pattern of the sequence of the syntactic constituents is thus changed. This course of things is

noticeable both in the French and in the German examples. The last step in this development, which is yet far from observable in both languages, consists in the complete ousting of the traditional pattern by the innovated one.

### 7. Word order and grammatical level

There is yet an open question: what happens to the functions fulfilled by free word order if this becomes fixed by grammaticalization? Consider the following series of Latin sentences (a-versions) with their French equivalents (b-versions).

- (20) a. *Ioannem non puto venisse.*  
 b. *Jean, je ne crois pas qu'il soit venu.*  
 'As for John, I don't believe that he has come.'
- (21) a. *Porcus est Ioannes.*  
 b. *C'est un cochon, Jean.*  
 'He's a pig, John.'
- (22) a. *Vinum bibit Ioannes.*  
 b. *C'est du vin que Jean boit.*  
 'It's wine that John drinks.'

What (20) - (22) exemplify is, in turn: topicalization of a nominal constituent, suspension of the theme ("afterthought") and focussation of a nominal constituent. Latin fulfills all these functions within the clause frame by different word order variants. French takes recourse, in all cases, to complex constructions which transcend the clause boundary, viz.: left-dislocation in (20), right-dislocation in (21) and sentence clefting in (22). Such devices are available in Latin either to a very limited degree or not at all (cf. Havers 1926 for the so-called *nominativus pendens* [left-dislocation] and Löfstedt 1966 for the cleft-sentence).

Two conclusions may be drawn from this. First, the renewed fulfillment of functions for which the old means are no longer suited goes the same way as the renewal or innovation of syntactic patterns which we saw in the preceding section: available syntactic means are used to form complex constructions. Once these are grammaticalized, they replace the older simple constructions. Second, in order to get greater order freedom, one has to step up to a higher grammatical level. If one is not content with the freedom that is available within the French clause, one has to form a complex sentence, e.g. by right- or left-dislocation or by sentence-clefting. In other words: given a hierarchy of grammatical levels from the sentence via clause and phrase down to the word, then there is principally greater order freedom at the higher levels than at the lower ones; cf. Ross 1973.

By the example of the adpositions that are grammaticalized to case affixes, we saw that word order turns into the positioning of affixes in morphological slots of their hosts. At this level, order rules are strictest. Thus one can say that a language with rich inflectional morphology has shifted most of its order rules to the word level. This has been seen already by E. Sapir (1921:109f), who writes:

Every language has its special method or methods of binding words into a larger unity. The importance of these methods is apt to vary with the complexity of the individual word. The more synthetic the language, in other words, the more clearly the status of each word in the sentence is indicated by its own resources,

the less need is there for looking beyond the word to the sentence as a whole. ... And yet to say that a sufficiently elaborate word-structure compensates for external syntactic methods is perilously close to begging the question. The elements of the word are related to each other in a specific way and follow each other in a rigorously determined sequence. This is tantamount to saying that a word which consists of more than a radical element is a crystallization of a sentence or some portion of a sentence, that a form like *agit* is roughly the psychological equivalent of a form like *age is "act he"*.

From here we can go on and say: The parallel gradation of grammatical levels and of degrees of syntagmatic variability is universal. That is, in every language there is less order freedom at the lower levels than at the higher ones. What is not universal is the grammatical levels themselves; i.e., there is no universal set of grammatical levels that would have to recur in every language. Let me clarify this by (20) - (22). If we want to correlate the grammatical levels of Latin with those of French, we have two possibilities. Either we start from structural criteria and thus distinguish in both languages the level of the - potentially complex - sentence from the level of the clause. Then we find out that a complex sentence in French is characterized by the fulfillment of different functions than a complex sentence in Latin. Or else we start from functional criteria and differentiate between a level at which the functions of sentence perspective ("pragmatic functions") are fulfilled and a lower level at which predication is achieved. Then we find out that the functions of sentence perspective are fulfilled at the structural level of the complex sentence in French, but at the structural level of the clause in Latin. Consequently, if grammatical levels are of a semiotic nature, which means they are characterized by the association of certain functions with certain structures, then they are not universal.

Thus, the hierarchies of grammatical levels of different languages may be staggered in regard to the association of functions and structures. A clear case of this is the typological classification that has been bestowed upon Classical Chinese. This is a language without inflectional morphology and with strictly regulated word order. It has been considered an isolating language since the beginning of language typology. This terminology implied that words do not contain grammatical signs by which they relate to each other. Again, V. Skalička (1966) classifies Classical Chinese and all the languages formerly referred to as isolating as polysynthetic. By this he means that the languages form large complexes of morphemes which are rigidly determined in their mutual position. Obviously the grammatical levels are here conceived differently. For those typologists who call the languages in question isolating, the monomorphemic signs of Chinese are words and, consequently, the constructions formed by them are phrases and clauses. For Skalička, instead, they are just morphemes, and the constructions formed by them are therefore words. What is important in the present connection is not a quarrel about terms, but the fact that languages such as Classical Chinese apparently possess a grammatical level that lies between our levels of the word and the phrase with regard to syntagmatic variability of its constituents. A typological comparison of the functions associated with these levels remains to be carried out.

## 8. Grammaticalization of word order

Let us summarize:

1. Grammaticalization is reduction of the autonomy of a linguistic sign. It comprises, on the one hand, the transformation of lexical elements into grammatical ones and their further reduction to zero, and on the other, the reduction of the paradigmatic and syntagmatic variability of the sign.
2. Reduction of syntagmatic variability includes fixation of word order. This is why grammaticalization goes hand in hand with the loss of word order freedom.
3. Grammaticalization also includes renewal of constructions at the highest grammatical levels. Both new grammatical elements and order patterns are attained by the development of hitherto unexploited lexical and syntactic sources. Accordingly, word order freedom or a new order pattern at a given grammatical level is gained by grammaticalization of a pattern that belongs to the next higher level.
4. A grammatical level is defined by the association of functions and structures. The structural aspects comprise the allowable degree of syntagmatic variability. It decreases at the lower grammatical levels.
5. The overall potential of syntagmatic variability of signs does not necessarily differ among languages. It is, however, distributed among different levels in different languages. In a language with rich inflectional morphology, order restrictions are amassed at the word level. In languages with so-called fixed word order, restrictions are concentrated at a higher level, for instance at the phrase or even the clause level.
6. Word order is not an expression device on a par with inflectional morphology. Instead it is an instance of syntagmatic variability and, thus, a structural aspect of the autonomy of the language sign. Its freedom or fixation depends on which grammatical levels are particularly strongly developed in the language.
7. The functional and dynamic conception of language advocated here provides the interface between internal and external factors in grammatical change. Grammaticalization is a necessary consequence of the desire to lend forceful expression to one's thought. A corollary of this is the suppression of whatever is not expressive. Insofar, grammaticalization is a strictly internal factor or, rather, method of grammatical change; it is at work in all languages at all times. Speech communities, however, differ as to which social groups will take greatest freedom in the realization of this desire, and under which circumstances its pursuit is more or less sanctioned by the society. Thus, in one and the same society, journalistic and bureaucratic style may indulge in the formation of fancy complex prepositions (cf. Lehmann 1990), while the colloquial language may allow for certain kinds of ellipsis (cf. 6 above). Thus, the distribution of specific grammaticalization processes within a language may well be controlled by social factors.

### Notes

<sup>1</sup> Alleged counterexamples are mostly reconstructed ones. Here I would appeal to R. Jakobson's (1958) principle that reconstructions should keep within the limits generalizable from historically documented cases.

<sup>2</sup> There is another common structural trait in all of my German examples: the verb is in each case followed by a clitic modal particle (which is in some cases preceded by a clitic personal pronoun).

## References

- Bailard, Joëlle 1982, "Le français de demain: VSO ou VOS". *ICHL* 5, 1982:20-28.
- Comrie, Bernard 1981, *Language universals and linguistic typology*. Syntax and morphology. Oxford: Blackwell.
- Greenberg, Joseph H. 1963, "Some universals of grammar with particular reference to the order of meaningful elements". Greenberg, J.H. (ed.), *Universals of language*. Report of a conference held at Dobbs Ferry, New York, April 13-15, 1961. Cambridge, Mass.: MIT Press; 58-90. 2. ed. 1966:73-113.
- Greenberg, Joseph H. 1969, "Some methods of dynamic comparison in linguistics". Puhvel, Jan (ed.), *Substance and structure of language*. Lectures delivered before the Linguistic Institute of the Linguistic Society of America, University of California, Los Angeles, June 17 - August 12, 1966. Berkeley & Los Angeles, Cal.: Univ. of California Press; 147-202.
- Havers, Wilhelm 1926, "Der sog. *nominativus pendens*". *IF* 43:207-256.
- Jakobson, Roman 1958, "Typological studies and their contribution to historical comparative linguistics". *PICL* 8:17-25.
- Lehmann, Christian 1982, "Zur Universaliendiskussion in der italienischen Linguistik. (La nozione di 'tipo' e le sue articolazioni nelle discipline del linguaggio, *Lingua e Stile* XV/3, 1980)". *FoL* 15, 1981:443-458.
- Lehmann, Christian 1985, "On grammatical relationality". *FoL* 19:67-109.
- Lehmann, Christian 1990, "Grammaticalization and related changes in contemporary German". Heine, Bernd & Elizabeth C. Traugott (eds.), *Approaches to grammaticalization*. Vol. II: Focus on types of grammatical markers. Amsterdam & Philadelphia: J. Benjamins (Typological Studies in Language, 19:2); 2:493-535.
- Lehmann, Winfred P. 1973, "A structural principle of language and its implications". *Language* 49:47-66.
- Löfstedt, Bengt 1966, "Die Konstruktion *c'est lui qui l'a fait* im Lateinischen". *IF* 71: 253-277.
- Mallinson, Graham & Barry J. Blake 1981, *Language typology*. Cross-linguistic studies in syntax. Amsterdam etc.: North-Holland (North-Holland Linguistic Series, 46).
- Paul, Hermann 1920, *Prinzipien der Sprachgeschichte*. Tübingen: Niemeyer. 5. ed.
- Ross, John R. 1973, "The penthouse principle and the order of constituents". Corum, Claudia et al. (eds.), *You take the high node and I'll take the low node*. Papers from the comparative syntax festival. The difference between main and subordinate clauses. 12 April 1973. A paravolume to Papers from the Ninth Regional Meeting. Chicago: CLS; 397-422.
- Sapir, Edward 1921, *Language*. An introduction to the study of speech. New York: Harcourt, Brace & World.
- Skalička, Vladimír 1966, "Ein 'typologisches Konstrukt'". *TLP* 2:157-163.
- Vennemann, Theo 1974, "Topics, subjects, and word order: from SXV to SVX via TVX". *ICHL* 1:339-376.
- Vincent, Nigel 1980, "Iconic and symbolic aspects of syntax: prospects for reconstruction". Ramat, Paolo et al. (eds.), *Linguistic reconstruction and Indo-European syntax*. Proceedings of the Colloquium of the 'Indogermanische Gesellschaft', University of Pavia, 6-7 September 1979. Amsterdam: J. Benjamins (AStHoLS; CILT, 19); 47-68.