

WARNING CONCERNING COPYRIGHT RESTRICTIONS: The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproduction of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If electronic transmission of reserve material is used for purposes in excess of what constitutes "fair use", that user may be liable for copyright infringement.

LONGMAN LINGUISTICS LIBRARY

General editors

R.H. ROBINS, *University of London*

GEOFFREY HORROCKS, *University of Cambridge*

DAVID DENISON, *University of Manchester*

For a complete list of books in the series, see pages v and vi.

Linguistic Typology:

Morphology and Syntax

Jae Jung Song



An imprint of **Pearson Education**

Harlow, England · London · New York · Reading, Massachusetts · San Francisco · Toronto · Don Mills, Ontario · Sydney
Tokyo · Singapore · Hong Kong · Seoul · Taipei · Cape Town · Madrid · Mexico City · Amsterdam · Munich · Paris · Milan

Abbreviations

ABS	absolutive	FUT	future
ACC	accusative	GEN	genitive
A(GT)	agent	IMP	imperative
APASS	antipassive	INC	inclusive
APPL	applicative	INCOMP	incomplete
ASP	aspect	IND	indicative
AUX	auxiliary	INST	instrumental
BEN	benefactive	INTR	intransitive
C	causative affix	INV	inverse
CL	classifier	IO	indirect object
CLT	clitic	IPFV	imperfective
CMP	completive	LOC	locative
COM	comitative	M	masculine
COMP	complementizer	MVT	movement affix
CONJ	conjunction	NFUT	nonfuture
CP	causative prefix	NOM	nominative
CREL	correlative	NP	noun phrase
CS	causative suffix	NR	nominalizer
D	deictic	OBJ	object
DAT	dative	OBL	oblique
DECL	declarative	OBV	obviative
DEF	definite	OPT	optative
DEM	demonstrative	P	patient
DEP	dependency	PART	participle
DET	determiner	PASS	passive
DO	direct object	PFV	perfective
DR	direct	PL	plural
ERG	ergative	POSS	possessive
F	feminine	PROG	progressive
FIN	finite	PROX	proximate

1

Introducing linguistic typology

1.1 What do linguistic typologists study?

There are generally estimated to be about 4,000 to 6,000 languages in the world. Depending on where the distinction between languages and dialects is drawn, the total number of the languages of the world may easily shoot up to 7,000.¹ This sheer number alone gives one at least a rough idea of the immense diversity of the languages of the world. A moderate example of this diversity comes from basic word order – which will be discussed in detail in Chapter 2. Consider the following six sentences, each exhibiting one of the six basic word orders.

- (1) Korean (SOV)
kiho-ka saca-lil cha-ass-ta
Keeho-NOM lion-ACC kick-PST-IND
'Keeho kicked the/a lion.'
- (2) Thai (SVO)
khon nū kàt mǎa tua nán
man this bite dog CL that
'This man bit that dog.'
- (3) Welsh (VSO)
Lladdodd draig ddyn
killed dragon man
'A dragon killed a man.'
- (4) Malagasy (VOS)
manasa ny lamba ny vehivavy
wash the clothes the woman
'The woman is washing the clothes.'

- (5) Panare (OVS)
 pi? kokampö unki?
 child washes woman
 'The woman washes the child.'
- (6) Nadëb (OSV)
 samüüy yi qa-wùh
 howler-monkey people eat
 'People eat howler-monkeys.'

What is most intriguing about these examples is the way the three main expressions – denoting the entity which initiates an action (S), the entity at which that action is directed (O) and the action itself (V) – occur in all six logically possible permutations, i.e. SOV, SVO, VSO, VOS, OVS and OSV. The diversity evident in basic word order will further increase with the order of constituents at other levels also taken into consideration, e.g. the word order within noun phrases, prepositional or postpositional phrases, and the like. With the six logically possible basic word orders all realized in the languages of the world, however, it may not immediately be clear whether or not there is anything significant about basic word order which can shed light on the nature of human language. (See 2.1 for discussion of the concept of basic word order, and also of flexible or free word order languages, as opposed to languages with basic word order.)

Despite the differences among them there must be certain properties whereby the languages of the world are all recognized as falling into the category of human languages – perhaps an obvious point to make. There must, therefore, be an underlying unity to human languages. There are linguists who are concerned directly with discovering this unity by studying the rich structural variation found in the languages of the world. These linguists are known as *linguistic typologists*, or *typologists* for short. Their investigation of cross-linguistic variation is referred to as *linguistic typology*, or *typology* for short. For instance, having observed the existence of the six logically possible basic word orders in the languages of the world, linguistic typologists will ask questions such as: what is the actual distribution of the six basic word orders in the languages of the world?, is each of the basic word orders more or less equally distributed in the languages of the world?, if not, what is the frequency of each basic word order?, and which basic word order(s) is (are) the most or least common in the languages of the world? As cross-linguistic research has so far shown, there is, in fact, a preponderance of the two basic word orders exemplified in (1) and (2), that is SOV, and SVO. This suggests strongly that the distribution of the basic word orders cannot possibly be random or arbitrary because otherwise the six basic word orders should be more or less evenly distributed among the languages of the world, that is about 16.6 per cent for each basic word order. More importantly, the

distinct preference for SOV and SVO points the way towards the existence of (a) certain motivating factor(s) for the skewed distribution of basic word order. Linguistic typologists will thus also raise the question as to why the distribution of the basic word orders is the way it is.

Furthermore, linguistic typologists compare two or more structural properties with a view to ascertaining whether there exists a (statistically significant) correlation between them and, if so, how strong that correlation is. For instance, basic word order has been compared with the presence (or absence) of prepositions, or postpositions. Verb-initial languages (or languages with the verb appearing first in the sentence, i.e. VSO and VOS) are always found to be equipped with prepositions, not with postpositions.² This means that verb-initial word order does not co-occur with postpositions. Welsh is a typical language exhibiting this correlation, as in (7) and (8):

- (7) Welsh
 Gwelodd y bachgen ddyn ddoe
 saw the boy man yesterday
 'The boy saw a man yesterday.'
- (8) Welsh
 a. trwy Gaerdydd
 through Cardiff
 'through Cardiff'
 b. yn y côr
 in the choir
 'in the choir'

This finding constitutes one important property of human language in that it does represent a constraint on possible variation within human language: no verb-initial languages are predicted to have postpositions instead of prepositions. There is no reason why the two independent properties, basic word order, and the presence of prepositions or postpositions, should correlate with each other to the effect that the presence of verb-initial word order implies that of prepositions. Logically speaking, there should also be verb-initial languages with postpositions, which is not the case. The question that directly confronts linguistic typologists is, why should such a correlation exist at all? In more general terms, the primary task for linguistic typologists is to identify and explain the properties that make human language what it is. The question can thus be rephrased: what is a possible, as opposed to impossible, human language? But, as has already been demonstrated with respect to the distribution of the basic word orders (that is, with all the six logical possibilities of word order attested to different degrees in the languages of the world), linguistic typologists may also address an attenuated form of the question: what is a more probable, as opposed to less probable, human language?

1.2 Typological analysis

In order to reach the stage of asking and, hopefully, answering the question posed at the end of the previous section, linguistic typologists will need to go through the following four stages in typological analysis: (i) identification of a phenomenon to be studied; (ii) typological classification of the phenomenon being investigated; (iii) the formulation of (a) generalization(s) over the classification; and finally (iv) the explanation of the generalization(s).

First, linguistic typologists must determine what they would like to investigate. There are, of course, no theoretical restrictions on what structural properties or grammatical phenomena should or should not be studied. Nor are there any restrictions on how many properties should simultaneously be studied at a given time. Some linguistic typologists may choose one feature of language as an object of inquiry, whereas others may at once probe into more than one. But what one must exercise circumspection about is which of the properties selected for typological analysis is actually worthwhile investigating, with some properties proving to be more interesting or revealing than others. Said differently, some are more likely than others to lead to linguistically significant typological generalizations. For instance, compare the typological property of basic word order with the presence of question particles. As has already been hinted at, the selection of basic word order as a typological property has led to a number of empirically or theoretically interesting questions and issues. But what about use of question particles? The languages of the world will be typologized into two groups: those with question particles and those without. But what is there to be understood from this bland typological classification? There does not seem to be much more to be done or learned about it. It is difficult to imagine that this typological classification can be put to much use at all in understanding the nature of human language – unless perhaps it is studied in conjunction with some other structural properties. In a way, therefore, the first stage of typological analysis may depend crucially on the investigator's insight or intuition to a great extent just as in any kind of scientific endeavour. Furthermore, the first and second stages of typological analysis may have to be carried out concurrently to a certain degree. This is because, unfortunately, one does not know in advance whether or not the chosen property is going to be a typologically significant one.

Once a property (or properties) has (or have) been chosen for typological analysis, structural types pertaining to that property (or those properties) will be identified or defined so that the languages of the world can eventually be classified into those types. In the case of basic word order, for instance, six (logically possible) types are identified, whereby languages are typologized according to the basic word order type that they exhibit. Some languages will be grouped as SOV, others as VSO, and so forth. The identification of the six basic word order types, and the classification into those types of the lan-

guages of the world will then constitute the linguistic typology of basic word order. The skewed distribution of the six basic word orders emerging from this typological classification is such that there is concluded to be a distinct tendency towards SOV, and SVO in the languages of the world. This can then be taken to be a significant generalization over the data classified – representing stage (iii) above. It will also ultimately lead to the question as to why there is this strong tendency, i.e. stage (iv). At this final stage, linguistic typologists will make every attempt to explain the structural tendency in question.

Similar comments can be made of the correlation between verb-initial word order, and prepositions, which was alluded to in 1.1. First, the languages of the world are surveyed in terms of basic word order on the one hand, and the presence (or absence) of prepositions or postpositions on the other. There are four different logical combinations of these two properties (2×2): (i) verb-initial languages with prepositions; (ii) verb-initial languages with postpositions (i.e. non-prepositions); (iii) non-verb-initial (i.e. verb-medial or verb-final) languages with prepositions; and finally (iv) non-verb-initial languages with postpositions. The languages of the world will then be classified into these four types. As it turns out, all the types except for that in (ii) are well represented in the languages of the world as illustrated by the following examples:

- (9) Tzotzil (VOS & prepositions)
- a. ʔi-s-pet lok'el ʔantz ti t'ul-e
 CMP-3-carry away woman the rabbit-CLT
 'The rabbit carried away the woman.'
 - b. xchiʔuk s-malal
 with 3-husband
 'with her husband'
- (10) Yoruba (SVO & prepositions)
- a. bàbá ra bàtà
 father bought shoes
 'Father bought shoes.'
 - b. ní ojà
 at market
 'at the market'
- (11) Canela-Krahô (SOV & postpositions)
- a. hümre te rop cakwĩn
 man PST dog beat
 'The man beat the dog.'
 - b. pur kam
 field in
 'in the field'

This classification will eventually give rise to the generalization: the presence of verb-initial word order implies that of prepositions. This will in turn call for an explanation as to why this implicational relationship exists between verb-initial word order and prepositions.

1.3 Typology of language universals

Properties such as the preponderance of SOV and SVO, and the correlation between verb-initial word order and prepositions are often referred to as language universals in linguistic typology. Strictly speaking, however, language universals must be true of all languages. Under this strict definition of the term, neither the preference for SOV and SVO, nor the correlation between verb-initial word order and prepositions may qualify as a language universal since the former is merely a structural tendency in human language, albeit a strong one, and since the latter makes explicit reference only to verb-initial languages with prepositions. In other words, only properties which all human languages have in common may be taken to be language universals. In linguistic typology, however, whatever statistically significant patterns or tendencies that are found in the languages of the world are also referred to as language universals.³ The correlation between verb-initial word order, and prepositions *is* a language universal in the sense that it embodies one significant statement about the nature of human language. (So is the preponderance of SOV, and SVO.) Thus, while not making explicit reference to non-verb-initial languages, the correlation between verb-initial word order and prepositions does capture a fundamental structural property of human language by excluding the logically possible co-occurrence of verb-initial word order and postpositions from the domain of human language as will presently be explained in more detail. In the present book, then, it is this broad interpretation of language universals that will be adopted (but cf. Ramat 1987).

Interestingly enough, language universals can themselves be typologized into four different types, by using two parameters: (i) absolute vs. non-absolute; and (ii) implicational vs. non-implicational. Absolute universals are exceptionless by definition. An example of this type of universal is: all languages have ways to turn affirmative sentences into negative ones (e.g. *James kicked the dog* → *James did not kick the dog*). Non-absolute universals – also known as universal tendencies, or statistical universals – are not without exceptions but the empirical validity of this type of universal does far outweigh the number of exceptions that may exist. The preponderance of SOV and SVO in the languages of the world is a non-absolute universal. More often than not, various statistical methods are employed in order to determine whether or not a given tendency is statistically significant (cf. 1.5.3).

Implicational universals take the form of ‘if *p*, then *q*’. Thus the presence of one property (i.e. the *implicans*) implies that of another (i.e. the *implicatum*). A good example of this type of universal has already been provided: verb-initial languages are always found to be equipped with prepositions. This can be rewritten as follows: if a language is verb-initial, then it is also prepositional. By design implicational universals will be based on interaction of more than one property. Thus there may also be implicational universals, involving more than two properties. One such example is Greenberg’s (1963b) Universal 5: if a language has dominant SOV order and the genitive follows the governing noun, then the adjective likewise follows the noun. In this implicational universal two properties are needed to predict a third. It is also possible that the implicatum can be more than one property. Again, Greenberg (1963b) offers an example of this kind: if some or all adverbs follow the adjective they modify, then the language is one in which the qualifying adjective follows the noun and the verb precedes its nominal object as the dominant order (Universal 21). It is not difficult to see that, other things being equal, implicational universals that predict the presence of multiple properties on the basis of a single property are more highly valued than those that predict the presence of a single property on the basis of multiple properties. To put it differently, it is preferable to predict as much as possible on the basis of as little as possible (Moravcsik 1997: 107). By this criterion of economy alone Greenberg’s Universal 21 is of more value than his Universal 5.

The predicting power of implicational universals is not confined solely to the properties to which they make explicit reference. Thus given the implicational universal ‘if a language is verb-initial, then it is also prepositional’, there are two other situations that fall out from that universal (in addition to the impossibility of verb-initial languages with postpositions). By making no claims about them, it, in effect, has the advantage of saying something about non-verb-initial languages with prepositions or with postpositions, thereby recognizing these languages also as possible human languages. In other words, the implicational universal in question rules out only verb-initial languages with postpositions as impossible human languages – that is, *p* & $\neg q$ (or not *q*), contradicting the original statement of ‘if *p*, then *q*’. What is referred to as a *tetrachoric* table is often used to indicate clearly which of the logically possible combinations of two (or more) properties is allowed or disallowed.

(12)		<i>prepositions</i>	<i>postpositions</i>
	verb-initial	Yes	No
	non-verb-initial	–	–

The tetrachoric table in (12) shows that the combination of verb-initial word order, and postpositions is an impossibility in human language. Thus the

implicational universal effectively serves as a constraint on possible variation within human language.

Non-implicational universals, on the other hand, do not involve the predicting of property X on the basis of property Y. They involve only a single typological property. They do not appear in the form of 'if p , then q '. The preponderance of SOV, and SVO is such a universal. Note that this particular universal is not only non-implicational but also non-absolute, thereby illustrating that universals may cut across the distinction between the absolute/non-absolute, and implicational/non-implicational parameters. Thus, in addition to non-absolute non-implicational universals, there may also be (i) absolute implicational universals, (ii) absolute non-implicational universals, and (iii) non-absolute implicational universals. An example of (i) has already been provided: if a language is verb-initial, it is also prepositional; an example of (ii) comes from the fact that all languages have ways to convert affirmative sentences into negative ones, and finally (iii) is exemplified by Greenberg's Universal 21 (see above).

1.4 Language universals and linguistic typology

Language universals are properties which must at least be true of the majority of the human languages. They also impose constraints or limits on possible variation within human language. Linguistic typology, on the other hand, is concerned with classification of languages into different structural types (i.e. individual structural properties, or correlations between them).⁴ Therefore, 'it may seem to the uninitiated something of a contradiction in terms to handle these apparently quite distinct areas of investigation together' (Mallinson and Blake 1981: 7). But, as may be gleaned from previous sections, the contradiction is more apparent than real. Language universals research, in fact, thrives on linguistic typology. This is because in order to discover language universals linguistic typologists first need typological classification on which to work. Thus linguistic typology 'provides material for establishing language universals' (Mallinson and Blake 1981: 7). With languages classified into different types, linguistic typologists may be able to discern patterns or regularities in the distribution of the types, for example, with some types being significantly more common than others, or with one (or more) of the logically possible types completely unattested or only marginally attested in the languages of the world.

This close relationship between language universals and linguistic typology can be most clearly demonstrated by one of the language universals that have been discussed in brief in previous sections: the strong tendency towards SOV and SVO. If the languages of the world had not been surveyed in terms of all possible six basic word orders, this structural tendency would never have been brought to light in the first place. To put it differently, the

typological classification in terms of basic word order of the languages of the world is a prerequisite for the significant statement to be made about human language. Imagine the prospect of discovering the tendency in question by examining only one language, or even a handful of languages! This may be too extreme an example but the point could not be made more strongly. Language universals underlying structural tendencies are statistical in nature. Thus it comes as no surprise that they demand statistical data.

Further demonstration of the symbiotic interaction between language universals research and linguistic typology comes from implicational universals. Recall that the presence of verb-initial word order implies that of prepositions (or p & q). This universal also makes indirect reference to the two other logical possibilities ($-p$ & q , and $-p$ & $-q$; cf. (12)). While making no negative claims about non-verb-initial languages, it does not only sanction verb-initial languages with prepositions as possible human languages but it also rules out verb-initial languages with postpositions as impossible human languages. In order to arrive at the actual formulation of this implicational universal, however, it first needs to be ascertained which of the four logical possibilities is attested or unattested in the languages of the world. That can be achieved only on the basis of initial typological classification of the languages of the world in terms of basic word order, and also in terms of the distribution of prepositions and postpositions.

The interaction between language universals and linguistic typology also highlights one of the virtues of formulating language universals on the basis of typological classification. Typological classification naturally calls for data from a wide range of languages (see 1.5.3 on how languages are selected or sampled for this purpose). Only by working with such a wide range of data is one able to minimize the risk of misinterpreting some of the least common structural properties as being (part of) language universals. This risk is more real than some linguists may be willing to admit because, when deciding to work with a small number of familiar or well-known languages (for whatever reasons), one is likely to deal with structural properties which may not in any real sense be representative of the languages of the world. For instance, use of relative pronouns in relative clauses is very common in European languages but it has been pointed out that it is a cross-linguistically infrequent type (Comrie 1989: 149). Therefore, universal claims about, or universal theories of, relative clauses which are put forth on the basis of these European languages alone should immediately be suspect.

1.5 Assumptions and problems in typological analysis

There are at least two theoretical assumptions which must be made in typological analysis: *cross-linguistic comparability* and *uniformitarianism*. In addition there are at least two practical problems that linguistic typologists must

address and, if possible, make explicit statements as to how they are going to deal or have dealt with in their typological investigation. These are problems of *language sampling* and *data collection*. It seems at least that, while the two assumptions are deemed reasonably uncontroversial, the problems of language sampling and data collection are far from resolved although, as will be seen in 1.5.3 and 1.5.4, the problem of language sampling has over the years attracted a good deal of attention from linguistic typologists. These issues and problems will in turn be discussed with a view to laying the groundwork for the chapters that follow.

1.5.1 Cross-linguistic comparability

Linguistic typologists study cross-linguistic variation in order to understand the nature of human language. The best way to gain access to the cross-linguistic variation of a grammatical phenomenon is to study as wide a range of languages as possible. Because they study a large number of languages all at once, linguistic typologists must therefore ensure that what they are comparing across languages be the same grammatical phenomenon, not different grammatical phenomena. It goes without saying that languages should be studied against one another in terms of the same property. Otherwise one will never be able to achieve what one sets out to: the description of cross-linguistic variation of the same grammatical phenomenon. If one wants to construct a typology of, for example, comparative constructions, how does one then actually make sure that one is comparing the comparative construction in language X with the same construction, and not something else, in language Y? To ask it differently, how does one identify the same grammatical phenomenon across languages? This is what Stassen (1985: 14) refers to aptly as *the problem of cross-linguistic identification*.

There are basically two ways of dealing with the problem of cross-linguistic identification (cf. Stassen 1985: 14–15; Croft 1995: 88–9). First, one may choose to carry out cross-linguistic identification on the basis of purely formal or structural criteria. A set of formal properties, e.g. verbal marking, adpositions (i.e. prepositions and postpositions), etc., may first be put together in order to identify a given grammatical phenomenon. Alternatively, one can opt for functional – i.e. semantic, pragmatic and/or cognitive – definitions of the grammatical phenomenon to be studied.

Which of the two types of definition – formal or functional – will meet the needs of typological analysis better? Croft (1995: 88) gives two reasons as to why formal definitions do not work for cross-linguistic comparison. First, structural variation across languages is so great that it cannot serve as the basis of cross-linguistic identification. As an example, Croft takes note of the fact that the subject relation in English may be expressed by means of two different grammatical relations in languages such as Quiché, Lakshota and Spanish. Second, formal definitions are internal to the structural system

of a single language, thereby again failing to constitute the basis of a language-independent definition. In a similar vein Stassen (1985: 14) also points out that language-dependent formal definitions do not tie in with linguistic typology, one of the primary aims of which is to characterize structural variation across languages. In addition, there are two fundamental reasons why formal definitions are not deemed appropriate for the resolving of the problem of cross-linguistic identification. Structural variation itself is what in the first place linguistic typologists want to identify for cross-linguistic comparison. In other words, one cannot make use of the structural variation which has not yet been established in order to identify that structural variation. It will be tantamount to using a (non-existent) description of X in order to describe X. Moreover, there can hardly be any *purely* formal definitions. Formal definitions of grammatical phenomenon X can only be identified and thus understood in the context of the function that X carries out. One cannot simply examine a given grammatical property and predict what function that grammatical property is used to carry out. This would be possible only if functions were inherent in, and thus deducible from, grammatical phenomena themselves. Rather, functions do arise out of what linguistic expressions are utilized for. For example, if one wants to study comparative constructions across languages, one cannot infer the function of comparison only from the linguistic expression in which that function is encoded (e.g. the use of adpositions, grammatical marking, etc.). One will not know what grammatical properties to look for in the first place, thereby being unable to recognize a comparative construction when one sees it.

In view of the foregoing objections to formal definitions, linguistic typologists opt for functional definitions for purposes of cross-linguistic identification. However, functional definitions may not be without problems, either. Far more frequently than not, functional definitions themselves tend to be based on pre-theoretic concepts, or ill-defined notions. This is not to say, of course, that the problem is unique to this type of definition. The definition of a given concept is always dependent on the understanding of other concepts which make up that definition – unless these other concepts are undefined theoretical primitives. For example, consider the semantic definition of comparison utilized by Stassen (1985: 15):

a construction counts as a comparative construction (and will therefore be taken into account in the typology), if that construction has the semantic function of assigning a graded (i.e. non-identical) position on a predicative scale to two (possibly complex) objects.

In order to understand this definition fully one needs to have an understanding of what a predicative scale, a graded position, etc. are. Also note that the definition has nothing to say about what form or shape the construction in question will take. Thus functional definitions are more of heuristics for

cross-linguistic identification than of definitions in the strict sense of the word. For this reason it may not always be entirely clear how wide a range of grammatical phenomena may be 'permitted' to fall under a given functional definition. As an example, take the semantically based definition of relative clauses adopted in Keenan and Comrie (1977: 63), which runs as follows:

Our solution to [the problem of cross-linguistic identification] is to use an essentially semantically based definition of RC [relative clause]. We consider any syntactic object to be an RC if it specifies a set of objects (perhaps a one-member set) in two steps: a larger set is specified, called the *domain* of relativization, and then restricted to some subset of which a certain sentence, the *restricting* sentence, is true. The domain of relativization is expressed in surface structure by the *head NP*, and the restricting sentence by the *restricting clause*, which may look more or less like a surface sentence depending on the language.

As Mallinson and Blake (1981: 266) correctly point out, it is not the case that Keenan and Comrie's definition of the RC 'sets a lower limit on the degree to which the RC can resemble a simple sentence or full clause and still be an RC.' Whatever structure is seen to perform the relative clause function as described above will be taken to be an RC, no matter how little resemblance it may bear to the relative clause in well-known languages, e.g. English. Note that the definition contains no distinct structural properties by which to identify RCs, other than the mention of the restricting clause and the head NP. Thus one may not always be certain whether or not a given grammatical structure in language X is a relative clause. It may well be nothing more or less than a 'general' structure which happens to be taken pragmatically or contextually as having a relative clause interpretation. Consider the following example of a so-called adjoined clause from Warlpiri (Hale 1976), which is susceptible to both relative clause and temporal interpretations as the English translation of (13) indicates (also Mallinson and Blake (1981: 266-8)). (The same structure can also have a conditional interpretation.)

- (13) Warlpiri
 ŋatjulu-lu o-ŋa yankiri pantu-nu,
 I-ERG AUX-I emu spear-PST
 kutja-lpa ŋapa ŋa-nu
 COMP-AUX water drink-PST
 'I speared the emu which was/while it was drinking water.'

There is evidence that calls into question the grammatical status as a genuine relative clause of (13). First, the adjoined clause as a whole can be positioned before the main clause as in (14).

- (14) Warlpiri
 yankiri-li kutja-lpa ŋapa ŋa-nu,
 ŋatjulu-lu o-ŋa pantu-nu
 'The emu which was drinking water, I speared it.' *or*
 'While the emu was drinking the water, I speared it.'

Moreover, the adjoined clause need not have an NP co-referential with an NP in the main clause (in which case a relative clause interpretation is not possible) as in (15).

- (15) Warlpiri
 ŋatjulu-lu lpa-na kali tjaŋtu-nu,
 I-ERG AUX-I boomerang trim-PST
 kutja-o-ŋpa ya-nu-nu ŋjuntu
 COMP-AUX walk-PST-hither you
 'I was trimming a boomerang when you came up.'

Also note that the syntactic linkage of the adjoined clause with respect to the main clause is, as Hale (1976: 78) puts it, 'marginal [or very loose] rather than embedded'. In fact, how the adjoined clause is interpreted 'is not determined by the grammar, but rather by a subset of the system of maxims, which are presumably observed in the construction of felicitous discourse, involving such notions as "relevance", "informativeness", and the like' (Hale 1976: 88). Given these pieces of evidence, the question does immediately arise as to whether or not the adjoined clause in (13) should really be regarded as a relative clause although under Keenan and Comrie's semantically based definition of RCs it may still qualify as an RC.

This is exactly the same question that Comrie and Horie (1995) raise as to the status as the relative clause of grammatical structures like the one in Warlpiri. They (1995) observe that in Japanese relative clauses do not behave as they do in languages such as English. In English relative clauses behave distinctly from other types of complement clause, whether with verbal or nominal heads.⁵ In Japanese, on the other hand, relative clauses are akin to complement clauses with nominal heads, distinct from complement clauses with verbal heads. In other words, there do not seem to be clear grammatical differences between relative clauses and complement clauses with nominal heads, with sentences potentially interpreted either as relative clauses or as complement clauses, 'depending on such factors as the semantics of the head noun (e.g. only certain head nouns allow the complement clause interpretation), and the plausibility of interpreting the head noun semantically as a constituent of the subordinate clause' (Comrie and Horie 1995: 69). They (1995: 73) also point out that in Khmer the grammatical marker used in relative clauses is 'not specifically a relative clause marker, but rather a general marker for associating subordinate clauses with head

nouns'. They (1995: 74) thus draw the conclusion from these observations that, there being no clear distinction between relative clauses and complement clauses of head nouns, the basic notion of relative clauses may not be of universal validity if it is meant by that notion that relative clauses are a distinct syntactic construction correlating highly with relative clause interpretations. In other words, they suggest that languages which lack relative clauses, such as Japanese and Khmer, make use of a general syntactic construction for relating subordinate clauses to head nouns, which is in turn subject to a wide range of pragmatic, not semantic, interpretations including that of relative clauses.

When confronted with such a problem as this, individual investigators may ultimately have to make up their own mind as to whether or not the structures in Warlpiri, Japanese and Khmer should be taken to be relative clauses. However, such a decision should not be taken in an arbitrary or random manner. One must, in fact, take into account at least two criteria, one language-internal and the other cross-linguistic, when making that kind of decision: (i) *functional-structural consistency*, and (ii) *measure of recurrence* or, more accurately, *measure of cross-linguistic recurrence*. Without supporting evidence from these two, it may hardly be justifiable to interpret the functional definition of relative clauses too broadly, that is to accept the adjoined clause in Warlpiri, or the 'relative clause' in Japanese or Khmer as a genuine relative clause.

First, one must determine whether or not relative interpretations are mapped consistently onto the adjoined clause in languages like Warlpiri. Thus, if the adjoined clause is the option or strategy used consistently for the expressing of relative clause function, it must be regarded as a genuine example of relative clauses. If, on the other hand, the adjoined clause is associated only on an *ad hoc* basis with relative clause interpretations, its status as a relative clause will be very doubtful. Being one of the two structures employed consistently for the expressing of relative clause function in Warlpiri (Mary Laughren, personal communication), the adjoined clause must be taken to be none other than the relative clause construction *par excellence* in that language.⁶

Even if the criterion of functional-structural consistency has been met, one cannot be too cautious about the status as the relative clause of the adjoined clause in Warlpiri, for instance. One should also be circumspect enough to take the structure to be an exemplar of the relative clause if and only if it recurs with relative clause function in language after language. This is the criterion of measure of recurrence. Of course, it cannot categorically be said in how many languages the structure in question should appear in order to be classified as a relative clause construction. But what can be said is this: the more languages make use of the structure for the expressing of relative clause function, the stronger one's confidence grows in accepting that structure as constituting one of the types of relative clause construction.

The measure of recurrence may sound to some ears too 'commonsense' to be legitimate in serious scientific investigation. This kind of measure of recurrence, however, is also adopted in other types of scientific investigation, albeit in much more rigorous form. For example, water is predicted to boil at 100 degrees Celsius at one atmosphere pressure (i.e. 760 torr, or about 14.7 lb/sq in) and, in fact, we know that it does so precisely because of its *recurrent* physical behaviour of reaching the boiling point at that temperature. Similarly, if a given structure is used recurrently, and recurrently enough across languages to express relative clause function, it must be regarded as exemplifying one of the types of relative clause construction available in human language.

1.5.2 The Principle of Uniformitarianism

Linguistic typologists often study not only currently spoken languages but also extinct ones, provided that they have been documented. This may perhaps strike one as odd, if not surprising, because one may expect typological classification to be concerned only with the currently spoken languages of the world. One may be inclined to think that language universals represent constraints or limits on structural variation within human language as it *is*, not as it *was* (or for that matter as it *will be*). But then why do linguistic typologists also include extinct languages in their typological investigation? The assumption underlying this inclusion is what is generally known as the *Principle of Uniformitarianism* in linguistics (see Lass (1980: 53–7, 1997: 24–32) for discussion thereof in the context of historical linguistics). Basically, what it means is that human languages of the past – or of the future for that matter – are not essentially different in qualitative terms from those of the present. It claims, therefore, that the fundamental properties of human language have remained invariable over time. There are believed to be no differences in evolutionary terms between languages of the past – as far back as one can go and claim the existence of human languages – and those spoken today. In other words, human language of today is at the same level of evolution as that of, say, 60,000 years ago.

This assumption is, of course, something that has never been subjected to empirical verification, and cannot be put to the test for obvious reasons; one simply cannot go back in time and examine languages spoken 60,000 years ago to see whether or not they were qualitatively different from those of today. Nor is there any logical reason why the principle should be correct. Nonetheless it plays an important role in typology (and equally in historical linguistics). The primary aim of typology is to discover universal properties of human language. Language universals should by definition be true of all human languages. If human languages were spoken 60,000 years ago, then these languages must also be included in any typological study, which is

utterly impossible. In the absence of the Principle of Uniformitarianism, then, no typological analysis will be possible or, more accurately, complete simply because it is impossible to 'recover' all unrecorded extinct languages from oblivion. With the Principle of Uniformitarianism in place, however, linguistic typologists can examine languages spoken today and, if and where possible, attested extinct languages as well – since they are available for inclusion in typological study, anyway – and can still make claims about the nature of human language. Similar comments can also be made of languages of the future. Since it is expected that they will also be human languages, any typological study must in principle include them as well, which is out of the question. But the Principle of Uniformitarianism also works in the opposite direction of time from the present, thereby allowing linguistic typologists to extend to languages of the future such universal properties as they may have discovered on the basis of currently available data. After all, under the Principle of Uniformitarianism the nature of human language is assumed not to change over time.

There are also rather practical reasons why the Principle of Uniformitarianism is adhered to in linguistic typology. Without this principle, languages must be seen to evolve constantly as time passes by. But if languages were evolving through time, and were conceived of as being at different stages of linguistic evolution, grammatical descriptions that linguistic typologists employ for their research would be completely useless for typological research because they invariably – and inevitably – record languages at different points in time, or at different stages of evolution, with some grammars being descriptions of languages of more than a few hundred years ago, and others being far more recent ones.

The absence of the principle will also lead to the view – which incidentally is generally not accepted in linguistics – that some languages should be at a more advanced stage of evolution than others because one would not be able to claim that all human languages have evolved to the same level. If languages were at different stages of linguistic evolution, it would be impossible to engage in any typological research since one would (arbitrarily) have to target at one particular stage of evolution which all human languages have reached at one time or another, and to study all grammatical descriptions of the languages of the world at *that* stage of evolution (assuming, of course, that it is possible to select such a stage, and also to have access to all grammatical descriptions at once).⁸

The Principle of Uniformitarianism, then, provides a kind of frame of reference within which fruitful typological research can be carried out productively without being hindered unduly by the intractable methodological issue, which does not necessarily have to be resolved at the current stage of development of linguistic typology as an empirical approach to the study of language (see Croft (1995: 86–7) for typology as a linguistic theory).⁹

1.5.3 Approaches to language sampling

Intuitively speaking, the best way to discover language universals is perhaps to examine all languages of the world.¹⁰ For obvious reasons, however, it is very easy to see why it is out of the question to do so. As was pointed out at the beginning of this chapter, there are about 4,000–7,000 languages currently spoken in the world. Individual linguistic typologists (or even a team of linguistic typologists for that matter) are unable to compare such a large number of languages or even a small fraction thereof in a reasonable span of time. In fact, economic considerations alone will rule the large-scale survey out as wholly unfeasible. What makes it even more unrealistic is the fact that not all languages of the world have been studied and described and are thus available for inclusion in typological research. It is correct to say that, as things stand, there are far more languages which await linguistic documentation than those which have been described. It is also true that many languages are so inadequately or poorly documented that linguistic typologists may not be able to find anything about the grammatical phenomena in which they are interested. In fact, it is plainly impossible to study all languages of the world because many languages have already died out, with some of them leaving little or no record, although their existence may be known to us (e.g. Arin, Assan, Kassitic, Illyrian, etc.). There may also be many other languages which are not even known to have existed. Furthermore, with dialects developing in separate languages over time, there will also be 'new' languages coming into being. If one's aim is to study all languages of the world, there must certainly also be room for these new languages in one's investigation. But, needless to say, there is no taking these 'future' languages into account before their birth, or emergence (cf. 1.5.2).

In view of these practical limitations of time, money, existence and descriptions (Perkins 1989: 297), linguistic typologists often work with a practically manageable set of languages or what is referred to commonly as a language sample. Naturally, questions arise as to how many languages should be included in a given sample, and how languages should be selected for that sample. A few researchers have over the years addressed these and other related questions but it seems that some of them 'unfortunately remain unresolved' (Croft 1995: 89).

Bell (1978) is the first to raise the issue of language sampling for typological research. In his programmatic yet most influential paper he explains the role of stratification in language sampling (i.e. the process of placing languages into different strata, e.g. genetic affiliation, geographic location, word order types, etc.) and discusses genetic, areal and bibliographic biases to be avoided in language sampling.

First, his discussion points the way towards the major methodological advantage of using stratified language samples. As Perkins (1989: 300) correctly points out, language sampling has two contradicting requirements to

meet. A language sample must have as many languages as possible so that any inferences or generalizations drawn from that sample can be extended in an empirically sound manner to language in general. At the same time it must have as few languages as possible because that is the whole idea of using the sample in the first place. Pertinent to this second requirement is the risk of including more languages from each stratum than required, thereby introducing 'variables that are not independent of [genetic] affiliation and location'. In stratified samples the number of cases to be studied can be more substantially reduced than the number of cases required in other types of sample. This suggests strongly that one can at least in principle learn as much about the nature of human language from a sample with a few languages as from a sample with many languages (Perkins 1989: 298). Equally important is the fact that stratification in language sampling ensures at least in principle that languages to be chosen be independent, and 'not [be] identical cases that should more properly be considered different instances of the same case' (Perkins 1980: 60–61; 1992: 124). This is known as the issue of the independence of cases, which, as will presently be seen, has perhaps been the most intractable problem that linguistic typologists have to deal with in language sampling. Bell (1978: 127–9, 137–40) also highlights the importance of probability or random sampling. Use of probability or random samples does not only enable researchers to run various tests in order to ascertain whether or not their findings are statistically significant (or whether or not they are due to chance). But it also makes it possible 'to estimate the extent of the error from a single sample' (Bell 1978: 157). 'Evaluating risks [of sampling error] involved in inferences from a sample to the population of all languages is extremely desirable in ascertaining the precision of the results obtained in a particular test of a theory' (Perkins 1980: 59–60; 1992: 127). Indeed Tomlin (1986: 25–6), for instance, is able to choose the 'best' sample from a number of potential ones precisely because he employs this kind of stratified probability sampling technique.

Second, Bell (1978: 145–9) calls for language samples to be accurately representative of the actual distribution of the languages of the world. In order to achieve such a representative sample one must make every effort to overcome certain sampling biases. He refers specifically to genetic and areal biases that should be eliminated from language samples. Languages of a single language family have structural properties or features in common by virtue of having derived from the same source. They are bound to have inherited structural properties from their parent or ancestor language. Moreover, languages of different genetic lineages are also known to share structural properties or feature when they have long been in contact with one another. In other words, structural features may come to diffuse across genetic boundaries to the effect that they may be shared by various languages which are unrelated to one another or at least are from different subgroups within a family. Such an area of linguistic convergence is technically called a linguistic

area or *Sprachbund*. For example, Meso-America is reported to be a fine example of *Sprachbund* (Campbell, Kaufman and Smith-Stark 1986). There are a number of features which the languages of Meso-America have in common but which are absent from languages outside the area: e.g. nominal possession; use of relational nouns; vigesimal numeral systems; non-verb-final basic word order; several widespread semantic calques; etc. Properties that languages share due to their common genetic heritage or contact are what may be called 'chance' or accidental structural properties of language families or linguistic areas, respectively (Comrie 1989: 10). They must be carefully distinguished from those properties which truly represent language universals, and should not be taken to be characteristic of human language. Special care must, therefore, be taken to ensure that particular language families or groups not be over-represented (or under-represented) in language samples but that languages be selected equitably from all known language families, or groups. Suppose one wants to study relative clause constructions on a cross-linguistic basis. If a given sample contains too many European languages at the expense of other groups or families, the use of relative pronouns will be given more weight or importance in one's investigation than it should be because, as was pointed out earlier, the use of relative pronouns is most frequently found in European languages and is, in fact, infrequent in other languages of the world. Any generalizations to be drawn from that sample will consequently suffer from over-representation of European languages, and the over-estimation of the relative pronoun type in universal theory of relative clause constructions. Conversely, if a language family is under-represented in a given sample, then it means that structural properties associated with that language family will in turn be afforded less significance than they should be, and also that other language families will end up being over-represented – even if it were only for the sake of making up for the predetermined size of the sample. Too many languages from a single *Sprachbund* may also be included in a sample, in which case there is a grave danger of misinterpreting as language universals structural properties which are characteristic of that *Sprachbund*. This danger looms larger when one is dealing with a linguistic area much greater than is conventionally known to exist. For instance, it was widely believed that there was a linguistic tendency for O(bject)-V(erb) languages to place modifying adjectives (or A) before nouns (or N). However, Dryer (1989: 274–5) finds that the putative correlation between OV and AN is due largely to the fact that it is the dominant pattern in Eurasia – which is not normally thought to constitute a *Sprachbund*. He points out that the tendency in the languages of the world is, in fact, the opposite pattern, i.e. OV and NA. Thus over-representation of languages of Eurasia must carefully be guarded against in the setting up of language samples.

Linguistic typologists often find themselves in an unenviable situation where they are forced to select languages for a sample, depending mainly on

whether or not grammatical descriptions or grammars are available. This is indeed a very unfortunate situation but sometimes cannot be avoided. For instance, Indo-European languages are very well documented in both breadth and depth, whereas the coverage of the languages of New Guinea and South America is very meagre. Even if linguistic typologists are willing to incorporate a representative number of languages from New Guinea or South America in their samples, they may thus be unable to have a reasonable amount of access to them simply because there are not enough (detailed) grammars of languages of these regions available in the first place. This is something that cannot easily be remedied, and will continue to create a certain amount of distortion or tension in linguistic typologists' samples 'even where the existence of the skewing and of its disadvantages are [*sic*] recognized' (Comrie 1989: 11).

Linguistic typologists may also work in a place where already published grammars are unfortunately not readily accessible to them; the libraries that they rely on for their research may hold mainly Indo-European or Oceanic languages and not much else, for instance. This is what Bell (1978: 145) refers to aptly as bibliographic bias. If this kind of bias is unavoidable and present in a sample, the least that the investigator can do is to state openly the existence of the problem for the benefit of other linguists.

For the foregoing pragmatic reasons linguistic typologists often opt for so-called *convenience* or *opportunity samples* (Bell 1978: 128). They may only select languages which they are familiar with or have ready access to through grammatical descriptions or language consultants. In fact, a good number of ground-breaking typological works are based on such convenience samples (e.g. Greenberg (1963b); Comrie (1976); Keenan and Comrie (1977); Nichols (1986) *inter alia*). The obvious shortcomings in their samples notwithstanding they did not only provide much insight into the nature of human language, which continues to play an important role in typological research. But, more often than not, they also gave impetus to subsequent large-scale research. Needless to say, any generalizations or inferences based on such convenience samples should only be taken as what they are – suggestions or preliminary findings concerning cross-linguistic patterns, or language universals – and they should naturally undergo further empirical verification, or revision on the basis of more languages, or more adequately constructed language samples.

How can languages then be selected equitably from all language families of the world? Bell (1978: 145–9) puts forth a specific proposal as to how this can be achieved. He sets an arbitrarily controlled time depth of genetic relatedness at 3,500 years so that all languages of the world can fall into a number of delimitable genetic groups, in fact a total of 478 groups. For example, the Indo-European family is regarded as consisting of twelve groups, whereas the Australian stock produces about twenty-seven groups. The use of the controlled time depth of 3,500 years, albeit hopelessly arbitrary,

is intended in no small measure to reflect genetic diversity within language families, or stocks (Bell 1978: 146). The more genetic groups there are within a family, or stock, the more genetic diversity that family, or stock exhibits. For instance, in Bell's calculation the Niger-Kordofanian, and Amerind stocks contain 900 languages each. But the genetic diversities of these two stocks are not comparable, with the former being less genetically complicated than the latter. The Niger-Kordofanian stock is allocated forty-four groups, while the Amerind stock is estimated to contain 150 groups. Bell's approach is based crucially on what may be termed *proportional representation*. Each language family contributes to a given sample in proportion to the number of the genetic groups in that family, or stock. Thus given the total number of 478 genetic groups in the languages of the world, each language family, or stock will take up a certain fraction of that total according to the number of genetic groups which that family, or stock encompasses. The Niger-Kordofanian stock will then be represented in a given sample at the ratio of 44/478 (or 9.2 per cent), whereas the Amerind stock will be given a much larger representation in the sample at the ratio of 150/478 (or 31.4 per cent). These same ratios of representation will apply regardless of the size of the sample. In a sample of 100 languages, for example, the Niger-Kordofanian stock will contribute about nine languages ($100 \times 44/478$), whereas the Amerind stock will be represented by about thirty-one languages ($100 \times 150/478$). Basically, this is a top-down approach in that the size of a sample is normally predetermined, and each language family or stock is proportionally represented in the final sample according to its share of the total number of genetic groups (Nichols 1992: 38).

Bell (1978) does not actually specify how areal bias is to be avoided in language sampling but it follows from his general discussion of sampling that it is possible to control for areal bias in much the same way as genetic bias has been dealt with. Thus no inordinate number of languages should be selected from the same geographical area for a given sample. To this end the whole world can be carved up into a number of geographical areas, for each of which only a representative number of languages can be selected. In this way the investigator can at least make a deliberate attempt to refrain from choosing too many languages from the same area since languages spoken in the same area are known to borrow from, or influence, one another, undergoing similar innovations.

Tomlin (1986: 24) follows Bell's (1978) sampling approach in that his primary aim in language sampling is also to produce a language sample which is representative of the languages of the world in both genetic and areal terms. But he makes a further attempt to hone the sampling technique by performing a statistical measurement known as the *Kolmogorov goodness-of-fit test* on his language sample. What this particular statistical test does is to evaluate the degree of disparity between the known theoretical genetic or areal distribution of the languages of the world (e.g. the overall genetic

classification of the languages of the world as presented in Voegelin and Voegelin (1977) or Ruhlen (1987)) and proposed language samples. Only when its deviance from the theoretical distribution is found to be statistically non-significant, will a given sample be adopted and implemented for typological research. By making use of the information on genetic affiliation of the languages of the world as provided in Voegelin and Voegelin (1977), Tomlin (1986: 27) creates a genetic frame which consists of a large number of 'cells', or genetic groups of approximately the level of Italic/Romance of the Indo-European family. In principle, these cells are supposed to be comparable to one another in terms of time depth, or level of genetic classification. This is, then, Tomlin's equivalent to Bell's controlled time depth of genetic relatedness, i.e. 3,500 years.

In addition Tomlin sets up an areal frame which consists of twenty-six 'cells', or linguistic areas. Basically, most of these areas are uncontroversially recognized linguistic or cultural areas, e.g. South Asia, Meso-America, New Guinea, etc. Where no such linguistic or cultural areas are evident, negative definition is employed. For example, after Southeast Asia, East Asia, Europe and Mid Central Asia having each been taken up as cells, one is left with a vast area, which is then identified as North and Northeastern Asia. Tomlin then derives a sample of 402 languages from his database of 1063 languages, and runs the Kolmogorov test of goodness-of-fit on the sample to measure the extent of its deviance from the 'characteristics' of the theoretical genetic and areal frames. Note that Tomlin does not start with a predetermined sample size but rather arrives at a sample of 402 languages in the process of 'approximat[ing] the genetic and areal distribution of the world's languages using his sampling technique' (Nichols 1992: 39), i.e. a bottom-up approach. Tomlin (1986) is an improvement on Bell's (1978) approach to controlling for genetic and areal biases in that his sampling technique demands of language samples a statistically acceptable level of genetic and areal representativeness.

There is at least one fundamental issue that must be addressed with respect to the proportionally representative language sample as outlined by Bell (1978) and further refined by Tomlin (1986): the independence of cases. The issue is related directly to the question – in fact, a fundamental question in all statistical procedures – as to whether or not languages selected for a given sample are indeed independent of one another or, in other words, are independent cases or units of analysis (Bell 1978: 146; Dryer 1989: 261–3; Perkins 1989: 299–301; Perkins 1992: 123–5). This is essentially what is known as *Galton's problem* or *Galton's objection* in cultural anthropology: anthropological traits which are related to cultures must be independent of one another since such traits often diffuse through migration and/or borrowing.

Perhaps the independence of cases is well illustrated by the often discussed weakness of proportionally representative language samples: the effect on representation of the actual sample size. That is, whether or not small

language families, or stocks can be taken into consideration hinges crucially on how many languages are going to be sampled. For instance, in Bell (1978) the Khoisan stock is thought to have five genetic groups. In order for this particular stock to contribute a single language to a sample the investigator must have no fewer than ninety-six languages in that sample. To put it differently, in samples of fewer than ninety-six languages there is no possibility of stocks like Khoisan being considered for typological research. This means that not all independent cases may have a chance of being included in a small sample. But the independence of cases certainly cannot be upheld when some independent cases are included in, but others are excluded from, the final sample because it will not be clear from that sample what the languages chosen are independent of. Even worse is the situation where small stocks that have been excluded from the sampling happen to possess exceptional or rare linguistic types because 'exceptional types test the rule' (Perkins 1988: 367). The weakness in question can actually be demonstrated by means of one of the samples discussed in Bell's (1978: 149) own paper. In his hypothetical sample of thirty languages, five out of the sixteen stocks (or 31.2 per cent) have nil representation!

Large language samples are generally not without problems, either. In a sample of, for example, 500 languages, there is no absolute guarantee that some of the languages chosen are not somehow remotely related to one another. There may be deeper genetic relatedness among the languages of the world than is generally believed (but cf. Nichols (1992: 40), who is not too much perturbed by this because she is of the view that there is much less deep relatedness among the languages of the world than is widely assumed). One can perhaps point out that there is not much one can do about deep genetic relatedness since there is at present no obvious way of detecting it. The independence of cases in language sampling is not just confined to genetic relatedness. Indeed it must also apply to linguistic areas since it proves very difficult not to include in large samples languages which may in one way or another have been in contact or may come from the same linguistic areas especially when these areas are not generally recognized in the literature. It is impossible, therefore, to extricate completely from large samples variables or factors that are not independent of genetic affiliation or geographical location. In fact, Dryer (1989: 263) goes so far as to suggest that it may not be possible to construct a sample of many more than ten languages if one decides to be really strict about the independence of cases in language sampling (but see Perkins (1989: 308)). Needless to say, a sample of ten languages is plainly unlikely to produce any significant generalizations about the nature of human language.

To sum up, in small samples where some language stocks are left out of consideration the independence of cases demanded of statistical sampling is not maintained because languages may not be independent of one another, with some being included in, and others being excluded from, the final

sample. In the case of large samples as well, languages chosen may not be independent, with some being remotely related to one another or coming from the same linguistic area(s). Therefore, both the validity, and reliability of conclusions or inferences drawn from samples will necessarily be compromised unless the independence of cases is upheld strictly in the constructing of those samples.

Related to the independence of cases in language sampling is the distinction between linguistic preferences and actual frequencies of different linguistic types in the languages of the world. It is the former, not the latter, that language universals should be equated with. Some linguistic types may be the most frequently found in the languages of the world simply because they are utilized in *large* language families. Such linguistic types are not necessarily related to language universals, or linguistic principles which make up human language. The distinction in question is most clearly demonstrated by Dryer (1989: 259–60). He puts forward a scenario in which there are 1,000 languages in the world. In this world there is one large language family of 900 languages, with the remaining 100 languages evenly distributed among ten small language families (i.e. ten languages in each of these small families). All the 900 languages in the large family are SVO, and all the languages of the other ten small families are SOV. As Dryer asks, do we conclude from this distribution that there is a linguistic preference for SVO over SOV? Of course not. We will instead come to the conclusion that SOV represents a linguistic preference in basic word order. The basis of the conclusion is very clear. The fact that 900 out of the 1,000 languages are SVO is merely a non-linguistic historical accident. These 900 languages are SVO because, for example, they must have inherited the same basic word order from their parent language. There are ten SOV families as opposed to only one SVO family. The distinction between linguistic preferences and actual frequencies of linguistic types is an important one, which is often not recognized and made in the typological literature. Since it will be discussed in detail in Chapter 2, suffice it to mention that Tomlin's (1986) large-scale study of basic word order reveals that there is no statistical difference between the frequency of SOV and that of SVO, both being the most frequent basic word order types. It is tempting to draw the conclusion from this that there is no linguistic preference for SOV to the exclusion of SVO. But Dryer (1989: 269–70) demonstrates that there is indeed a linguistic preference for SOV over SVO, and that the lack of a statistical difference between SOV and SVO in Tomlin's investigation is due to the distinction between linguistic preferences and actual frequencies of different linguistic types not being maintained. For example, in Tomlin's sample about 40 per cent of the SVO languages come from Niger-Congo, and there is also a large contingent of SVO languages from Austronesian (Dryer 1989: 260, 270).

Perkins (1980) deals with the independence of cases by adding a 'qualitative' dimension to language sampling. He derives a probability or random

sample of fifty languages/cultures from the population or universe of all cultures using as a frame the list of cultures proposed in Murdock's (1967) *Ethnographic Atlas*. He takes into consideration the languages which are spoken in the cultures listed by Murdock.¹¹ He further informs his selection of the fifty languages/cultures by ensuring on the basis of Voegelin and Voegelin (1966) and Kenny (1974) that they not be (substantially) close to each other in terms of genetic and cultural relatedness. He is acutely aware of the fact that in a large sample it is unavoidable to include closely related languages/cultures (see above). But this should not be permitted in all statistical procedures because, as has already been pointed out, these procedures 'presuppose the independence of cases' (Perkins 1992: 367). He thus establishes his sample of fifty languages in such a way that no two languages/cultures are from the same language family, or the same cultural area. Clearly, this is a significant improvement over the previous sampling attempts in that it is specifically designed to maximize the genetic (or cultural) distance between the languages of the sample. Therefore, Perkins's sampling strategy contrasts with Tomlin's, for instance, in that one and only one language will be selected for each language family regardless of the actual size of language families. For example, insofar as sampling is concerned, there is no 'qualitative' difference between a language family of ten languages, and a language family of 100 languages because both will contribute only one language each to the sample.

Perkins's deliberate attempt to maximize genetic distance within his sample is indeed a welcome step in the right direction of the discovering of linguistic preferences, rather than of the actual frequencies of different linguistic types, a point which was adequately illustrated by Dryer's example above. Furthermore, his sampling technique – at least in principle – makes a serious attempt to meet the requirement of the independence of cases in language sampling for purposes of statistical inferences. As Whaley (1997: 39) points out, a sample of fifty languages is also practically manageable in size for a single investigator; it may not be too difficult to collect fifty languages for typological research.

Dryer (1989: 261–3), however, casts doubt on whether or not the languages in Perkins's sample are truly independent of one another. In particular, he takes issue with the inclusion in the fifty language sample of three Nilo-Saharan languages (i.e. Ingassana, Maasai and Songhai) and six languages which are considered to be Mon-Khmer (Car, Semai, Khasi, Khmer, Palaung and Vietnamese).¹² Independence being a relative notion (Dryer 1989: 262), one may with little reservation accept the presence of the Nilo-Saharan languages in the same sample because Nilo-Saharan is a very remote grouping. But Dryer (1989: 262) is absolutely correct to raise the question as to how remotely related languages have to be in order to be admitted as independent for sampling purposes. Moreover, Dryer alludes to the fact that the notion of independence actually depends on what linguistic property is

being investigated. Word order, for example, changes quite easily, whereas morphological properties may be less susceptible to change. Thus 'a pair of languages that one might consider independent for the purposes of word order might not be properly treated as independent for the purposes of morphology' (Dryer 1989: 262).

A potential problem specific to Perkins's sample is its lack of control for areal bias. In a way his decision not to include two languages from the same cultural area can perhaps be regarded as an attempt to guard against areal bias (Dryer 1989: 263) because cultural relatedness may possibly be indicative of geographical proximity. This, in fact, seems to be Perkins's rationale for deciding not to control for areal bias when he (1992: 126) says that

[t]he sampling plan used insures that cultures included in the final sample are not closely related culturally. This is some insurance that languages that are similar *due to contact* and cultural borrowing are not included in the sample [emphasis added].

This assurance notwithstanding, Dryer (1989: 262) identifies in Perkins's sample a number of languages that come from well-defined linguistic areas, e.g. Southeast Asia, Pacific Northwest of North America, Meso-America, etc (cf. Tomlin (1986)), and he comes to the conclusion that the cultural frame that Perkins uses in lieu of an areal frame is far too fine for the purposes of controlling for areal bias. It is not clear at the moment whether or not Perkins's assumption is a sound or reasonably acceptable one to make in language sampling. But in this context it is worthwhile commenting, along with Dryer (1989: 283–4), that in general linguistic properties are not borrowed as easily as cultural ones (e.g. Driver and Chaney (1970) for discussion of the Yurok, Karok and Hupa tribes in California with their languages belonging to different stocks, yet their cultures being almost identical). Perhaps it may be judicious not to equate linguistic borrowing with cultural borrowing.

Perkins's sampling technique has taken a few innovative steps towards the basic requirements of statistical procedures. But the issue of the independence of cases does not seem to have been resolved completely in his fifty language sample. There is still room for suspecting that some of the fifty languages may at least be remotely related to one another in either genetic or areal terms. As has already been pointed out, Dryer (1989: 263) is of the opinion that no language sample with more than ten languages may be able to meet the requirement of the independence of cases. But for obvious reasons samples of ten or so languages are very unlikely to produce anything interesting or insightful about the nature of human language. This harks back to Perkins's observation that in language sampling a balance has to be struck between the two contradicting requirements: a language sample must have as many languages as possible so that any inferences or generalizations

drawn from that sample can be extended in an empirically sound manner to language in general, while for practical, conceptual and statistical reasons it must also have as few languages as possible (see Perkins 1992: 123–4). In the final analysis, then, Perkins's sampling technique is also beset with the fundamental problem of the independence of cases, although it is designed to resolve that problem.

Dryer (1989) puts forward a novel yet ingenious method in language sampling. One of his aims is to achieve the independence of cases at the level of large *linguistic areas*, which are continental or almost continental in size though he (1989: 267) claims that his sampling method is able to control for 'the most severe genetic bias' as well. Note that Dryer's concept of *linguistic areas* should be distinguished from the conventional concept of linguistic areas or Sprachbund. (For the sake of distinction Dryer's *linguistic areas* will appear in italics in the present book.) In Dryer's work a *linguistic area* refers to 'an area in which at least one linguistic property is shared more often than elsewhere in the world to an extent which is unlikely to be due to chance, but which is probably due either to contact or remote genetic relationships' (Dryer 1989: 266). He also invokes the concept of a *genus*, which is analogous to Bell's (1978) genetic group. Genera are supposed to be comparable to the sub-families of Indo-European, e.g. Germanic or Romance – or a time depth of 3,500 to 4,000 years. The languages of a given sample are then placed into 322 genera in total, largely in line with the genetic classification of Ruhlen (1987) (cf. Bell's (1978) 478 genetic groups; in his actual sample of 542 languages, however, Dryer (1989: 267–9) only operates with 218 genera). Thus only genera, not individual languages, are counted and taken into consideration for purposes of typological investigation. Note that the independence of cases, which is vital for all statistical procedures, is not at all demanded at the level of genera, which are only identified and utilized in order to control for the most severe genetic bias, the assumption underlying this decision being that 'languages within genera are generally fairly similar typologically' (Dryer 1989: 267; but see below). The independence of cases is required strictly at the next stage of Dryer's sampling method, where the world is divided into five large continental (or almost continental) areas: Africa, Eurasia, Australia-New Guinea, North America and South America.¹⁵ These five *linguistic areas* are then assumed to be independent of one another. It is thus at this level of the *linguistic areas* that the independence of cases is claimed to be achieved or maintained in Dryer's language sampling. The purpose of using these five *linguistic areas* is, therefore, to control not only for areal bias of the proportion that has not hitherto been understood to have a bearing on language sampling but also for remote genetic relatedness, which may not be amenable to conventional historical methods. Unless this type of areal bias – the underlying cause of which may possibly be partly or largely genetic (Dryer 1989: 266) – is controlled for in language sampling, there is indeed a distinct possibility of failing to realize when

'apparently statistically significant results ... may simply reflect areal [or remotely genetic] phenomena rather than linguistic preferences' (Dryer 1989: 283).

The ingenuity of Dryer's approach lies precisely in the fact that the independence of cases is sought – to the extent that this is possible – at the level of the five *linguistic areas*, not at the level of genera, the number of which may be rather unwieldy for purposes of such statistical manipulation. His technique thus makes it possible 'to take into consideration all of the data at hand' (Croft 1995: 91), while dealing with only the five *linguistic areas* for purposes of controlling for areal *and* genetic biases. Moreover, the most severe genetic bias is claimed to be controlled for at the level of genera particularly because '[i]n some areas of the world, these genera are the maximal level of grouping whose genetic relationship is uncontroversial' (Dryer 1989: 267). It also seems that Dryer's decision to achieve the independence of cases at the level of *linguistic areas* makes much sense because the divisions between the five *linguistic areas* 'are rather well defined physically' (Dryer 1989: 268), and, should thus be far less controversial than the divisions between the 322 genera.

For further illustration of how Dryer's method actually works, the preference of SOV over SVO can be referred back to. Recall that Tomlin (1986) does not recognize any statistical difference, or significance between the actual frequency of SOV, and that of SVO in the languages of the world. Dryer (1989: 269–70), on the other hand, provides evidence in support of SOV being a linguistic preference over SVO as exemplified below (N.B.: Afr = Africa, Eura = Eurasia, A-NG = Australia-New Guinea, NAm = North America, SAm = South America).

(16)	Afr	Eura	A-NG	NAm	SAm	Total
SOV	[22]	[26]	[19]	[26]	[18]	111
SVO	21	19	6	6	5	57

The numbers in both rows represent the number of genera exhibiting SOV, or SVO for each of the five *linguistic areas*. The larger of the two figures for each of the columns appears in square brackets (or in a box in Dryer (1989)). Though the difference between SOV and SVO in Africa is far from significant, there does clearly emerge a generalization to the effect that SOV outnumbered SVO by five *linguistic areas* to none, thereby confirming that there indeed is a linguistic preference of SOV over SVO. The logic here is that, since the five areas are assumed to be independent of one another both genetically and areally, there would only be one chance in thirty-two – one chance in sixty-four in Dryer (1992), wherein six areas are recognized, with Southeast Asia & Oceania teased out from Eurasia – for all five areas to display the given property if there were no linguistic preference for the more frequently occurring language type (Dryer 1992: 85).

Another universal claim that can easily be checked using Dryer's sampling method concerns the correlation between the position of the verb, and the distribution of adpositions, partly discussed earlier in the present chapter. (Dryer (1992) opts for OV and VO instead of verb-finality and verb-initiality.) Dryer (1989: 271) provides the following results, which, in fact, strongly support the correlation in question (N.B.: Po = postposition, Pr = preposition).

(17)	Afr	Eura	A-NG	NAm	SAm	Total
OV&Po	[13]	[27]	[15]	[20]	[12]	87
OV&Pr	2	2	1	0	0	5

(18)	Afr	Eura	A-NG	NAm	SAm	Total
VO&Pr	[14]	[23]	[5]	[15]	[5]	62
VO&Po	4	1	0	2	2	9

Note that Dryer (1989: 269, 271) takes a very conservative attitude towards interpreting his results, e.g. (16), (17) and (18). Only if and when all the five *linguistic areas* conform to the hypothesis being tested, that hypothesis is considered to be a language universal. For instance, if only four of the five conform to the hypothesis, then he prefers to speak of 'trends', short of statistical significance. By his standards, then, some of the language universals that other linguistic typologists would happily accept will have to be relegated to trends (Whaley 1997: 41).

Attractive as Dryer's sampling method may seem, there may be a number of problems associated with it, some being of a general nature, and others being unique to his approach. The use of genera in his sampling approach is to control for the most severe genetic bias especially when caused by large language families which only constitute a single genus, e.g. Bantu and Malayo-Polynesian. But, as Croft (1995: 91) points out, Dryer's technique cannot help eliminate completely genetic bias which would have existed in the linguistic situation prior to the time depth of a genus to which that technique can bring one back. As Croft (1995: 91) himself admits, however, this is 'essentially an inescapable problem'. It is almost impossible to imagine if any sampling technique or method can actually overcome this problem. Perhaps it is not really necessary to be overly concerned about the linguistic situation beyond the time depth of, say, Dryer's genus, if one adopts Nichols's (1992: 40) assumption that 'throughout human prehistory most languages have left few descendants, stock have usually had at most one sister, and at any time about half of the world's lineages have been isolates'. It follows from this that '[t]here is much less deep relatedness among the world's languages than is widely assumed'.

Whaley (1997: 41) thinks that there are two 'problems' with Dryer's approach: (i) difficulty with determining to which genus a given language belongs; and (ii) the large sample size required for the method to be effective.

The first is caused by the simple fact that there are many languages whose generic classification is 'unknown, unclear, or under dispute' (Bybee, Perkins and Pagliuca 1994: 28). This, however, cannot be said to be a problem unique to Dryer's approach. All designers of language samples will have to live with this problem, which is only emblematic of the state of the art in genetic classification (see below). The second problem is not really a problem as such. In Dryer (1989), a sample of 542 languages is used. This is, of course, a much larger sample compared with, for example, Perkins's sample of fifty languages. Tomlin's (1986) database, on the other hand, consists of 1,063 languages, and his database could potentially have been used as a sample, if not for genetic and areal biases contained in it (Tomlin 1986: 29). The point is, different researchers have access to varying amounts of resources. It makes little sense to say that one's sample is too small or too large (by whose standards, anyway?). More importantly, it should not be forgotten that Dryer only deals with a far smaller number of genera (i.e. 218 of them), not 542 individual languages in his database, in order to control for genetic bias.

In addition to the foregoing there are four specific problems with Dryer's sampling method that need to be discussed in some detail. The first problem concerns selection of sample languages. It is not entirely clear how (and which) languages are chosen for each genus. For instance, when setting up her sample Nichols (1992: 27) carefully avoids languages considered by specialists as linguistically divergent or atypical of the family so that the language(s) chosen can be representative of the whole family. There is nothing in Dryer's discussion concerning the actual selecting of sample languages. Is there a set of uniform criteria for selecting sample languages for genera? Related to this is also the issue of the minimum or maximum number of languages to be selected for each of the genera. Unfortunately, Dryer is not explicit on this point, either. It may thus depend on the actual number of languages to be chosen whether a given genus may turn out to be of type X, type Y or whatever (cf. Dryer 1989: 270). It seems that too much is left to chance. More detrimental to his method is, however, that the languages used to make up genera 'are not randomly chosen so that independence is not assured by his method but is undercut by the prior problem of lack of randomness used to choose his languages' (Perkins 1989: 299). This certainly is a serious problem from a statistical point of view alone.

It also needs to be confirmed or verified to reasonable satisfaction whether or not the five continental *linguistic areas* are comparable to one another in terms of diversity, typological and/or genetic. For instance, Nichols (1992: 39) expresses her doubt that the standard continents such as Africa, North America and Australia are the same kind of unit. She is of the opinion that they are not of the same order. For example, the entire Old World – comprising Africa, Ancient Near East, Northern Eurasia and South and Southeast Asia (Nichols 1992: 27) – behaves like the same kind of unit as the single island of New Guinea in terms of diversity (cf. 6.2.3). This leads her

(1992: 39) to suggest that in order to ensure that the units or areas to be compared really be comparable, one needs to devise a 'bottom-up' areal survey in which units or areas are first objectively defined or identified by using the same set of criteria (it is not clear what these may be), and these units or areas are then put together to 'complete' the whole world, as it were. Dryer's approach, on the other hand, is top-down in that the five *linguistic areas* are first assumed to be comparable, and each of the languages in the sample is then pigeonholed into one and only one of these areas through the medium of genera. As a result, his intention to treat these areas as independent cases can be called into question because independent cases must also be comparable.

In Dryer (1989) the existence of large *linguistic areas* is identified by one continental area patterning differently from the rest of the world (Dryer 1989: 284). For example, the putative correlation between OV and AN, which was once widely thought to be a language universal, is demonstrated to be owing largely to the dominance of that correlation in Eurasia. The data in support of this finding are presented below (Dryer 1989: 274).

(19)	Afr	Eura	A-NG	NAm	SAm	Total
OV&AN	6	[22]	5	9	6	48
OV&NA	[17]	9	[15]	[17]	[10]	68

Note that except for Eurasia there is, in fact, a clear tendency towards OV and NA in all *linguistic areas*. According to Dryer, the standing out from the rest of the world of Eurasia points to the possibility that the co-occurrence of OV and AN is an areal phenomenon which is associated only with Eurasia. He (1989: 284–5) then poses himself a hypothetical yet interesting question as to whether or not the entire world may constitute one huge *linguistic area*. This naturally leads to a further question as to whether the linguistic preferences as reported in (16), (17) and (18) may not be genuine linguistic preferences but rather due to remote genetic or areal factors (i.e. all languages deriving from a single parent language, or Proto-World). Of course, there is no way of knowing that this is not the case as Dryer (1989: 284) admits. However, one can actually go even further and extend this line of thinking to a situation such as in (19). If one really cannot determine whether the linguistic universals in (16), (17) and (18) represent linguistic preferences, or common histories, how can one be completely sure whether the correlation between OV and NA in four of the five *linguistic areas* is a linguistic preference, or is due to remote genetic or areal factors. There being – strictly speaking – nothing in Dryer's method that can in principle tell linguistic preferences and areal phenomena apart, one cannot be certain about this question. In that case only the pattern in Eurasia could well reflect the linguistic preference of OV and AN, with the other four *linguistic areas* simply exhibiting what they have inherited from their common source, or what they have assumed through diffusion, or even both.

Finally, Dryer (1989: 267) claims that in his sampling method counting genera rather than individual languages makes it possible to control for the most severe genetic bias 'since languages within genera are generally fairly similar typologically'. However, this assumption is somewhat questionable in view of a great deal of variation that does exist between different linguistic properties in terms of innovation or conservatism. For example, as Dryer (1989: 262) himself acknowledges, basic word order properties change fairly easily, whereas morphological ones may be far more resilient to change. In other words, the assumption that languages within genera are generally fairly similar typologically may not apply equally to all different types of linguistic property. Prior to the adopting of that assumption one may then be well advised to ascertain first whether or not the linguistic property being studied is a relatively stable one over time and/or in the context of contact. Furthermore, one needs to find out how stable a given linguistic property has to be in order to uphold the assumption in question. Of course, these are totally different areas of study yet to be undertaken elsewhere on a large scale (cf. Nichols 1992).

1.5.4 Determining language sample sizes

There are two additional pieces of research in language sampling that need to be reviewed here: Perkins (1989), and Rijkhoff, Bakker, Hengeveld and Kahrel (1993). These works do not concern directly language sampling *per se* but rather the determining of optimum language sample sizes.

Perkins (1989: 294) addresses the question as to what is the appropriate basis for selecting sample sizes. He puts forward a statistical method whereby (i) appropriate sample sizes can be determined with the requirements of representativeness and independence met; and (ii) the extent to which linguistic properties being studied are dependent on the variables of stratification or classification used for sampling purposes can be statistically measured. He chooses to take advantage of a statistic known as τ_{RIC} (Light and Margolin 1971; Bishop *et al.* 1975), for which a test of significance – called U^2 which has a chi-square distribution – can be employed. This statistic also generates an analysis of variance measure for categorical variables where one variable is taken to be independent, and the other dependent (Perkins 1989: 301). It thus enables one to determine, for example, whether or not in a given sample the dependence on the genetic affiliation of basic word order is statistically significant. For the sake of demonstration, he runs the statistic on Tomlin's database of 1,063 languages, and finds that the statistical significance of the association between the independent and dependent variables is substantial, suggesting that, if Tomlin's database itself were used as a sample, it would include 'cases . . . that show obvious effects due to [genetic] inheritance or borrowing' (Perkins 1989: 298). Inclusion of such cases in a sample will indeed be very detrimental to the criterion of inde-

pendence of case. The τ_{RIC} can also produce two limiting numbers of the size of the sample in which the dependent variable (e.g. basic word order) is not significantly associated with the independent variable (e.g. genetic affiliation). The mean of these two limiting numbers is then taken to be the size of an optimum language sample. Another advantage of using this statistic is that it also makes it possible to choose the level of stratification, or categorization that requires the smallest number of languages without compromising the requirements of representativeness and independence. By using the statistic, for example, Perkins (1989: 305–9) evaluates three different levels of geographical stratification – (i) Tomlin's (1986) twenty-six cells in his areal frame; (ii) Dryer's five *linguistic areas*; and (iii) the division of the world into the two hemispheres (i.e. eastern and western). The statistic of τ_{RIC} indicates that the levels of geographical stratification in (i), (ii) and (iii) require about 90, 40 and slightly over 100 languages, respectively. The smallest number of languages is required of the level of stratification in (ii): the five continental *linguistic areas* identified in Dryer (1989). Thus the highest level of strata for a language sampling frame for basic word order should be the five continents, not the twenty-six areas or cells in Tomlin (1986) or the two hemispheres.

Perkins's sampling design method can prove to be 'a heuristic suggesting plausible language sample sizes that meet the criterion of independence of cases' (Perkins 1989: 301). He (1989: 294) thus comes to the conclusion that '50 to 100 languages are most appropriate for a variety of linguistic variables [including basic word order]'. But, as Perkins (1989: 302) himself notes, it can also serve as a kind of post-mortem on language samples that have been used in typological research especially when the dependence on the type of stratification of linguistic properties investigated has not been statistically tested.

There are, however, two points that can be made about Perkins's sampling design method, one practical and the other theoretical. First, Perkins's method requires a rather large database (Perkins 1989: 301). In order to arrive at optimum sample sizes, all languages of the database must be genetically classified, and analysed in terms of structural properties. That may also depend to an extent on the nature of linguistic properties to be studied. Perhaps this may be unavoidable since, as Bell (1978: 141) perceptively notes, '[t]he investigator is always faced with the paradox that the optimum sample requires the very knowledge that he seeks'. What really is required of Perkins's sampling design method seems to be some kind of ready-made databank, wherein a large number of languages have been recorded with various linguistic properties or variables having already been identified for each of these languages without any specific research topics in mind (e.g. GRAMCATS at the University of New Mexico or Matthew Dryer's database at the SUNY at Buffalo). Most investigators do not have access to such a large databank, however.

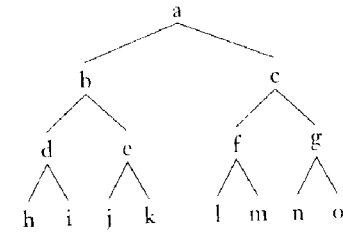
More problematic for Perkins's method is that the results of the τ_{RIC} analysis seem to hinge on the actual distribution of a given linguistic variable in the database from which an optimum language sample is to be derived as Perkins (1989: 303–5) himself demonstrates. This suggests that the optimum language sample size to be calculated may actually vary from a database which contains too many VSO languages to another one which includes too few VSO languages, for instance. To put it differently, the τ_{RIC} analysis seems to be sensitive at least to the actual constitution of the database in terms of the linguistic variable being investigated. If this is the case, the validity of 'optimum' sample sizes arrived at by means of Perkins's sampling design method cannot but be in dispute.

Rijkhoff, Bakker, Hengeveld and Kahrel (1993) describe two different types of language sample (or two different approaches to language sampling): (i) samples used to identify tendencies or correlations (i.e. *probability samples*); and (ii) samples set up to discover 'all possible realizations of a certain meaning or structure across languages' (i.e. *variety samples*). Following both Bell (1978), and Perkins (1980), they develop a sampling design procedure which falls under the second type. They (1993: 172) believe that genetic relatedness is the most important criterion in language sampling because languages that are closely related in time also tend to be closely related in space, culture and typology (but see above). In other words, they are concerned with controlling for genetic bias alone.

Rijkhoff *et al.* (1993: 172) point out that the best way to avoid genetic bias is to ensure that all languages in the sample come from different phyla. They thus agree with Perkins (1980) that a sample must include at least one representative from each phylum so that there is minimal representation of all phyla in that sample. But at the same time they note that this way to control for genetic bias will only give rise to a sample of fewer than thirty languages – assuming that there are twenty-seven phyla (cf. Ruhlen 1987). They point out that such a small sample does not make a good variety sample, the primary goal of which is 'to maximize the amount of variation in the data' (Rijkhoff *et al.* 1993: 171). A good variety sample must then reflect the greatest possible structural diversity so that even cases of the rarest structural type can have a chance of representation (Rijkhoff *et al.* 1993: 171). Clearly, a sample consisting of fewer than thirty languages will hardly be likely to achieve this. Rijkhoff *et al.*'s (1993: 179–80, 196) sampling method thus makes an attempt to incorporate Bell's (1978) call for appropriate representation of genetic diversity in a sample (i.e. variation *within* phyla), and the insight of Perkins's (1980) sampling technique (i.e. variation *across* phyla).

Rijkhoff *et al.* (1993: 171) also attempt to replace Bell's notion of genetic groups, or his 'age-criterion' (or the time depth of 3,500 years) with an 'objectively' computable measure of genetic diversity by taking advantage of the internal structure of genetic language trees. This is intended to reflect Bell's (1978) observation that '[i]f the strata [e.g. genetic groups] are not

Figure 1.1: A hypothetical language family tree



equally homogeneous, some increase in sampling efficiency may be achieved by weighting [...] samples according to strata variability'. Thus the number of languages to be selected for a phylum in a sample must be proportional to the internal genetic diversity of that phylum. In Rijkhoff *et al.*'s weighted sampling procedure the internal structure of the family tree is translated mathematically into a diversity value, which will in turn be utilized in order to calculate the exact number of languages to be selected for each of the twenty-seven phyla (including language isolates, and pidgins and creoles), based largely on Ruhlen (1987).

The diversity value (or DV) is computed on the basis of the number of nodes at the intermediate levels between the top node, and the terminal nodes at the bottom end in a language tree. Top nodes, e.g. *a* in the hypothetical family tree in Figure 1.1, are excluded from computation because they do not contribute to internal diversity. So are terminal nodes, e.g. *b, i, j, k, l, m, n* and *o* in Figure 1.1. Only the internal structure of a language family tree, or the intermediate nodes, e.g. *b, c, d, e, f, g*, in Figure 1.1, must be taken into consideration for purposes of computation because it is possible, for example, that a language family of 300 languages is far less complicated internally than a language family of fifty languages.

Rijkhoff *et al.* (1993: 181–2) also recognize that high-level splits, or branchings are more significant in terms of contribution to diversity value than low-level splits because the former preceded the latter in time or because the former had more time to develop into distinct languages than the latter. Therefore, they build into the DV formula in (20) a factor of significance which decreases as the depth of intermediate levels increases, that is by steps of $1/n$ where n is the maximum number of intermediate levels found in any phylum, e.g. sixteen in the Niger-Kordofanian phylum in the case of Ruhlen's (1987) genetic classification.¹⁴

$$(20) \quad C_y = C_x + ((n-x)/n) * (N_y - N_x), \text{ where } x = y - 1$$

C_y represents the contributing DV at any given level, which is computed by combining the contribution of the immediately preceding level (or C_x), and

the difference between the number of nodes at C_y and that at C_x (or $N_y - N_x$), which is in turn multiplied by the factor of decreasing significance referred to above (or $(n - x)/n$). The contributions of all intermediate levels of a given phylum are then computed according to the formula in (20), and then averaged out to derive the mean DV for that phylum.

Once the mean DVs of the twenty-seven phyla have been worked out, they will *invariably* be used to decide how many languages must be selected from each phylum in order to construct samples of predetermined sizes as is presented in Table 1.1. Thus Rijkhoff *et al.*'s sampling procedure is top-down. Recall that in common with Perkins (1980), Rijkhoff *et al.* recognize the need to have at least one representative from each of the phyla in a sample irrespective of actual sample sizes. In a sample of 100 languages, for example, each phyla will first be allocated at least one language (i.e. twenty-seven languages in total) regardless of its DV score, with the remaining seventy-three languages being divided up among the twenty-seven phyla according to their DVs.

Rijkhoff *et al.*'s (1993: 192, 196) quantification of internal diversity is claimed to be an improvement over Bell's age-criterion, or genetic groups in that their DV computation 'can be seen as an objectivization of Bell's language groups', the basis of which is the time depth of 3,500 years. They argue that this arbitrary time depth is difficult to apply equally to all phyla especially when the histories of many phyla are not well understood due mainly to lack of documentation. This is a fair point to make but then Rijkhoff *et al.*'s objectivization of Bell's age-criterion hardly escapes the same criticism and does rather highlight one of the intractable problems associated with all approaches to language sampling. The way language family trees are constructed is due as much to lack of understanding of, or uncertainty about, internal genetic relations as to actual internal diversity. In fact, this point has not at all been missed completely by Rijkhoff *et al.* when they (1993: 177-8) point out: '[w]hat strikes us is that the number of languages (t) per non-terminal (nt) and preterminal (pt) node is low for relatively well-explored phyla like Indo-Ihitrite (ratios 1.67 and 2.65), and rather high for phyla for which our knowledge still leaves much to be desired, such as Indo-Pacific (ratios 2.93 and 4.65)'. The reason for this difference is that phyla whose internal genetic relatedness is well understood tend to have more intermediate groups recognized, with their trees being more hierarchical or less flat, whereas phyla whose histories cannot easily be accessed tend to contain fewer intermediate groups, thereby resulting in flatter or less hierarchical trees. The problem is exacerbated by the undeniable fact that the genetic classifications on the basis of which many samples have been set up are in turn based on different sets of criteria being applied to different genetic groupings. Rijkhoff *et al.* (1993) rely heavily on the genetic classification list provided in Ruhlen (1987), with DVs computed on the basis of the internal structure of genetic language trees. But how can one lay claim to comparability of the

Table 1.1: Number of languages in samples of different sizes

Phylum/Sample size	30	40	50	60	70	80	90	100	125	150	175	200	225	250
Afro-Asiatic	1	2	2	3	4	5	5	6	8	9	11	12	14	16
Altaic	1	1	1	1	1	1	1	2	2	3	3	3	4	4
Amerind	2	5	7	9	12	14	16	18	24	29	35	40	45	51
Australian	1	2	3	4	4	5	6	7	9	11	13	15	17	19
Austrie	2	4	5	7	9	11	12	14	19	23	27	31	35	39
Caucasian	1	1	1	1	1	1	1	1	1	2	2	2	2	3
Chukchi-Kamchatkan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Elamo-Dravidian	1	1	1	1	1	1	1	1	1	1	1	2	2	2
Eskimo-Aleut	1	1	2	2	3	3	4	4	5	7	8	9	10	11
Indo-Ihitrite	1	3	5	7	8	10	11	13	17	20	24	28	32	35
Indo-Pacific	2	3	5	7	8	10	11	13	17	20	24	28	32	35
Khoisan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Sumnerian	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Ket	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Nahali	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Uralian	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Burushaski	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Meroitic	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Basque	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Farusan	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Gilyak	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Na-Dene	1	1	1	1	1	1	1	1	1	2	2	2	2	3
Niger-Kordofanian	1	3	4	5	6	7	8	9	12	15	18	20	23	26
Nilo-Saharan	1	1	2	3	3	4	4	5	6	7	8	10	11	12
Pidgins and Creoles	1	1	1	1	1	1	1	2	2	2	3	3	4	4
Sino-Tibetan	1	1	2	2	3	3	4	4	5	6	7	9	10	11
Uralic-Yukaghir	1	1	1	1	1	1	1	1	1	1	1	1	1	1
TOTALS	30	40	50	60	70	80	90	100	125	150	175	200	225	250

(Rijkhoff *et al.*, 1993: 186)

internal structure of genetic language trees across the twenty-seven phyla when the conceptual basis of that internal structure differs from one genetic grouping to another? These, of course, are not criticisms levelled at Rijkhoff *et al.*'s sampling approach *per se* but rather they are intended to emphasize the problem which Bybee, Perkins and Pagliuca (1994: 28) succinctly summarize in the following way:

[F]irst, there are many languages for which genetic classification is unknown, unclear, or under dispute. Second, more is being learned each day about genetic relations, so that some of the information published in 1978 may be incorrect. Third, different criteria were used in establishing the groupings in different parts of the world. In some cases, genetic grouping is based on extensive historical documentation and historical-comparative work (as in the case of Indo-European languages); in other cases, the groupings are based on lexicostatistical surveys; and in still others, it is admittedly only a geographical region that is being identified as a group.

Thus Rijkhoff *et al.*'s objectivization of Bell's age-criterion can only be as sound as the genetic classification on which it is based. It must also be borne in mind that what is at issue is not objectivization of the internal diversity of phyla *per se* but objectivization of the internal structure of genetic language trees with their weaknesses, flaws, gaps, and all because, as Rijkhoff *et al.* (1993: 178, 198, *passim*) themselves reiterate, only the internal structure as represented or reflected in the form of a tree diagram is 'exploited to measure linguistic diversity among genetically related languages'.

This uncertainty about genetic classification brings back to the fore perhaps the most important of the issues in language sampling. How can one ensure the independence of cases in one's sample when one cannot be completely certain whether or not some of the languages being studied are genetically independent? For example, languages that are classified into different genetic families in a given genetic classification list may turn out to be genetically related to one another; languages whose genetic relations are still in dispute may simply be assumed to belong to the same language family. Languages like these will have a greater chance of inclusion in a large language sample than in a small one because it is more difficult to maximize genetic distance in the former than in the latter. In view of this it may not always be possible to determine to which genus a language belongs. Recall that this is precisely the difficulty that Whaley (1997: 41) associates with Dryer's (1989) notion of genera.

There is also a related – equally, if not more serious – issue of whether or not languages really are 'clearly definable objects' (Perkins 1989: 295) – which is a necessary condition for all statistical procedures. Croft (1990: 22), in fact, views this as a fundamental problem that has not properly been addressed in the sampling literature (but cf. Tomlin 1986: 35). For example,

how many languages are there in Austronesian, Indo-European, or Niger-Kordofanian? We do certainly have rough figures. But is there a uniform or universally agreed on set of criteria for identifying individual languages as opposed to dialects of a single language? These questions cannot be answered because the solution to the problem of definability of languages is unlikely to be found. This is why linguists put the total number of the languages of the world at 4,000 to 7,000. As is well known, the defining of a language in contrast to a dialect is beset with a number of practical difficulties (e.g. how do we measure mutual intelligibility?), and is also susceptible to non-linguistic, e.g. political, cultural, etc., considerations.

In view of all this it may seem to be an impossible task to achieve the independence of cases required of language sampling. But it must also be emphasized that despite all these difficulties linguistic typologists have over the years made a number of significant generalizations about the nature of human language as will be demonstrated in the following chapters. Perhaps it may offer a modicum of assurance about language sampling to note, along with Bybee *et al.* (1994: 28), that a genetic classification list such as Ruhlen (1987), or Voegelin and Voegelin (1977), 'provides an objective basis for sampling that was established independently of any hypotheses that [one sets out] to test'.

1.5.5 Problems with data

As can be seen from the foregoing, linguistic typologists work with a large amount of data from a large number of languages. It is, therefore, unrealistic to expect them to have a sound firsthand understanding of all sample languages that they deal with. They normally rely on primary sources: grammatical descriptions, or grammars in monograph (e.g. *Lingua Descriptive Series* by Routledge (previously by North-Holland and by Croom Helm) or journal article form (e.g. notably *International Journal of American Linguistics*, and *Oceanic Linguistics*). The most frequent problem with this kind of data source is that grammars are not always sufficient enough in detail, and/or broad enough in scope. Far more frequently than not, grammars may just gloss over or fail to examine the grammatical phenomena that linguistic typologists wish to study, although this often depends on what grammatical phenomena are being investigated. For instance, information on basic word order can probably easily be retrieved from most grammars, whereas that on the comparative construction 'is often not found in even the most minute grammars' (Stassen 1985: 13). There may exist only texts, not grammars, for some languages. These texts may not be glossed sufficiently enough – if glossed at all – to be easily amenable to linguistic or typological analysis. In a situation like this, linguistic typologists may have to work through texts by carrying out basic linguistic analysis themselves. If they are lucky, they may manage to find good examples but, as Stassen (1985: 13) laments, 'one often

despairs of the fact that two days of deciphering a grammatical text has not resulted in finding one good and clear example of the comparative construction'. Stassen's plight is not uncommon in typological research. Finding that a grammar does not deal with the phenomenon being studied often proves to be more difficult and time-consuming than finding that a grammar does indeed deal with the phenomenon. This is because one has to study a grammar from cover to cover in order to make sure that it does not deal with the phenomenon in question, whereas discovery of the phenomenon may not require perusal of a whole grammar. Moreover, one cannot just rely on the table of contents, and the subject index – if they are provided – to ascertain whether or not a grammatical description provides information on the phenomenon because, although the table of contents, or subject index does not make mention of it, one or more good examples of the phenomenon may be hidden in a most unlikely place in the grammar. To make matters worse, some grammars may be biased towards certain grammatical aspects, e.g. morphology, thereby providing no information whatsoever on linguistic typologists' areas of interest, e.g. syntax.

Linguistic typologists may also have to rely on secondary sources. But, as Croft (1990: 25) sounds a cautionary note, these materials may already be biased by the hypothesis or theoretical orientation of original analysts. Thus a great deal of discretion must be exercised in using secondary sources. For instance, they must, whenever possible, if not always, be checked by referring to the primary sources on which they are based. The reliability of secondary sources has also not proven to be particularly commendable because errors of citation are not unheard of, and can actually be repeated in subsequent works (Mallinson and Blake (1981: 14–15) for one such perpetuating error of citation).

Linguistic typologists may also need to work with language consultants. But this also poses an enormous amount of practical difficulty. Even if one can manage to find consultants for each of one's sample languages, it will be completely unfeasible economically, given the normal size of a language sample in typological research. What is worse, it will be unwise to work with only one consultant for each language. One may need to work with more than one consultant for each language for the sake of verification, or confirmation. Working with consultants in typological research also involves the same host of problems that are encountered in any research based on interaction with (live) human subjects. Perhaps one may wish to have this option as a last resort especially when certain subtle grammatical points in primary or secondary sources need clarification.

Whaley (1997: 42) also discusses use of questionnaires (e.g. Dahl 1985; also discussion of the role of questionnaires within the Leningrad Typology Group in 7.2). A set of written questions about the grammatical phenomenon to be studied can be sent out to language specialists, or language consultants. Despite its apparent advantages (e.g. the possibility of obtaining more detailed

information about the phenomenon than is available in grammars), this kind of data collection has its own share of difficulties, too. For one thing, it may be very time-consuming and costly to implement such a questionnaire especially when one is operating with a large language sample, *and* a small budget. But with the advent of electronic technology (e.g. e-mail, or electronic discussion lists such as the *LINGUIST-LIST* and the *ASSOCIATION FOR LINGUISTIC TYPOLOGY-LIST*, to which a good number of linguists around the world subscribe), execution of a questionnaire may no longer be so time-consuming and expensive as it used to be.¹⁵ Through an electronic network, one can instantly get in touch with a number of language specialists – provided that they are willing enough to respond to one's request (within a reasonable span of time). But there may still be problems with carrying out electronic questionnaires. Often the reliability, or credibility of respondents may be in doubt or in need of confirmation. Of course, this problem can be assuaged if a selection of qualified respondents can first be made by means of careful planning, screening, etc., and if only those who pass muster can then be approached in much the same way that a written questionnaire is administered. Mention must also be made of potential difficulties with electronic transmission of data from different languages, e.g. special fonts.

For practical reasons, however, most linguistic typologists prefer to work with primary sources. When in doubt, they may seek assistance from consultants, access secondary sources or draw upon questionnaires. But, as has already been noted, the use of primary materials is not without problems. Thus Croft (1990: 25) is led to declare that 'the typologist has to rely on faith in the qualities of the materials at hand', and that 'most of those materials do not inspire faith'. Inspire faith as they may not, the situation may not be so hopeless as this. One of the virtues of working with a large number of languages in typological research is that it does to a certain extent offset the problem of faith about which Croft is pessimistic. Recall that in section 1.5.1 mention was made of the measure of recurrence in cross-linguistic comparison. If a certain pattern or structural type occurs in language after language, one can be reasonably assured that this is a real phenomenon to be identified as such. Primary materials may fail to inspire faith individually but they may inspire faith collectively. This is due to the power of the measure of recurrence in cross-linguistic research.

1.6 Partial typology vs. holistic typology

References have been made in the preceding sections to cross-linguistic comparison of a wide range of languages in typological research. In those references languages are treated as if they were units of comparison. But, as has been pointed out in passing, this does not mean that languages in their

entirety are being compared but rather that structural properties, or constructions found in languages constitute objects of comparison. Consequently, when languages are typologized or classified, it must be understood that properties, or constructions, not whole languages, are put into different types. Said differently, 'languages' are simply being used here for indexing purposes, or as 'addresses' at which the different types of construction are located. Thus specific grammatical domains, e.g. basic word order, relative clause constructions, comparative constructions, etc., are chosen for typological research, and individual languages are then studied or analysed in terms of these selected domains, and classified into different types. This is an important point to bear in mind in the context of modern linguistic typology, at least in the mainstream thereof. This kind of typology – the analytical units of which are grammatical domains, not entire languages – is referred to as partial typology. What is more frequently done in partial typology is, however, to examine concurrently a cluster of properties in one and the same grammatical domain with a view to ascertaining whether or not these properties exhibit significant connections. For example, basic clausal word order is studied in conjunction with, for example, the distribution of adpositions: the presence of verb-initiality implies that of prepositions. Investigation of a cluster of properties such as basic clausal word order, adpositions, etc. is also an example of partial typology (cf. multi-feature, polythetic typology (Ramat 1986: 6)) despite the fact that multiple properties are under scrutiny. Partial because the investigation still deals with only part of grammar or language. Most of the modern typological works including the ones to be surveyed in the bulk of the present book are, in fact, subsumed under partial typology.

As opposed to partial typology is holistic typology. This kind of typology is no longer widely practised but commanded a great deal of popularity in the nineteenth century (see Horne (1966), Robins (1997), Greenberg (1974), Skalička and Sgall (1994), Sgall (1995) and Ramat (1995) for historical perspectives on this) when the dominant mode of intellectual thinking came from natural science (especially botanical science) (Ramat (1986: 3); also see Plank (1991) for possible input from anatomy; but see below, and also discussion of the Prague School Typology in 7.4).¹⁶ Scholars of this period – e.g. August Wilhelm Schlegel (1767–1845), Friedrich von Schlegel (1772–1829), Wilhelm von Humboldt (1767–1835) and August Schleicher (1821–68) – believed that language was a (natural) organism which possessed an 'inner form' (Robins (1997: chapter 7) for an overview of linguistics in the nineteenth century). This inner form was thought to be a manifestation of the spirit (*Geist*) of the people who spoke it (cf. Greenberg (1974: chapter 3); Robins (1997: 192–5)). In the words of Wilhelm von Humboldt (Finck 1899; Lehmann's (1978c: 423) translation), '[t]he characteristic intellectual features and the linguistic structure of a people stand in such intimacy of fusion with each other that if the one were presented the other would have

to be completely derivable from it'. The inner form was in turn assumed to be reflected in 'variation in grammatical mechanisms employed in relating lexical concepts to each other [or relational meaning]' (Shibatani and Bynon 1995: 4). Thus 'each language [was] a distinct revelation of the spirit (*Geist*)' (Greenberg 1974: 38). Coupled with growing interest in etymology (i.e. comparative-historical linguistics), this point of view led directly to the emergence of the morphological (or classical) typology, wherein three different basic strategies in the encoding of relational meaning were recognized: inflectional, agglutinative and isolating – Wilhelm von Humboldt later added a fourth, incorporating, to Schlegel's tripartite classification (for an example of the classical typology, see 5.2).¹⁷ The unit of analysis in this typology was undoubtedly the word, the structure of which 'was seized upon as in some sense central to the attempt to characterize the language as a whole' (Greenberg 1974: 36) so that 'the description of the entire grammatical system [could] be annexed to an exact description of the structure of the word in every language' (Lewy 1942: 15, cited in Greenberg 1974: 36). They believed that it was possible to characterize the *entire* language on the basis of a single grammatical parameter, or even a small number of grammatical features in much the same way that, for example, one could reconstruct the entire skeleton of an animal on the basis of a fossil jaw (Ramat 1986: 3). (This intellectual wish is what Shibatani and Bynon (1995: 16) term *the Gabelentzian ideal* since the formulation of this ambitious typological approach is generally attributed to Georg von der Gabelentz (1901).) In other words, the single property of morphology was assumed to constitute the ultimate basis of typological classification of all languages of the world. The classical typology is, therefore, a true exemplar of holistic typology.

But the adequacy of holistic typology as a classificatory scheme was subsequently called into question (Sapir (1921); cf. Greenberg (1954) for a quantitative approach to this type of classification). It was soon realized that languages in their entirety are not amenable to neat pigeonholing because most languages are in fact mixed types or hybrids in that they employ more than one type of morphological mechanism, e.g. Polynesian languages better characterized as 'agglutinative-isolating', and Cambodian as 'inflectional-isolating' (Shibatani and Bynon 1995: 5–9). As a consequence, holistic typology began to give way to far more modest partial typology in the twentieth century.

As Ramat (1986: 4) points out, however, holistic typology is 'perfectly understandable and reasonable' from a theoretical point of view. There are no theoretical or logical reasons why the Gabelentzian ideal cannot be retained as the ultimate – albeit probably unattainable – goal in typological research. The possibility – however remote that may be – of reducing a host of grammatical phenomena at all linguistic levels to a single underlying principle or even a handful of them is extremely attractive and tempting especially to the theoretically minded because, admittedly, '[t]he best

possible typology ... would be one that refers all manifestations of language to one single underlying principle'. With such a holistic typology in hand, for example, one should be in a position to predict on the basis of basic word order what type of relative clause construction is used, whether or not morphological causativization of transitive, as opposed to intransitive, verbs is permitted, what type of case marking system is employed, whether or not verbs have aspect rather than tense, and so forth (Shibatani and Bynon 1995: 12). But, as Comrie (1989: 40) puts it succinctly, 'experience to date is rather against this possibility: while we can state often wide-ranging correlations among logically independent parameters, these correlations are not sufficiently strong or sufficiently wide-ranging to give holistic types rather than cross-classification of languages on different parameters'. This probably explains the paucity of holistic typology in the current typological literature.

That is not to say that there have recently been no attempts at holistic typology. Since the bulk of the present book only deals with partial typology, it perhaps is worthwhile very briefly surveying two holistic works in the present chapter. (Members of the Prague School of linguistics or the Prague School Typology do also practise holistic typology by continuing with the nineteenth-century classical typology; but their work deserves separate discussion (cf. 7.4 and 7.5) in the context of Chapter 7).

W.P. Lehmann (1973, 1978b, 1978c), for instance, entertains the possibility of drawing certain phonological and morphological implications from basic word order typology: many OV order languages are claimed to have certain phonological characteristics such as (C)CV (or open) syllable structure, vowel harmony and pitch accent. To the best of the present writer's knowledge, however, Lehmann's claim has never seriously been put to the test. Interesting as it may be, it still remains nothing more than a hypothesis, if not a total conjecture (refer to Ramat's (1986: 5) sceptical view on this holistic typology).

Perhaps better known in the context of modern holistic typology is the work of the Russian linguist G.A. Klimov (1977, 1983). The basis (or the underlying principle) of his holistic typology is 'a language's predications and its categorization of basic nominal and verbal notions'. (Nichols 1992: 7–11). There are four basic types: (i) the accusative type, (ii) the ergative type, (iii) the active type and (iv) the class type. The types in (i)–(iv) are based on subject-object relations, agent-patient relations, an active/inactive distinction and referential properties of nominals, respectively. For example, the active type is claimed to be associated with a cluster of grammatical properties of different linguistic levels: lexical properties such as binary division of verbs into active and inactive, inclusive/exclusive distinction in first person pronouns, etc., morphological properties such as alienable/inalienable possession distinction, more verbal inflection than nominal inflection, etc., and syntactic properties such as SOV basic word order, direct object incorporation into verb, etc. (also see Klimov 1974). Interestingly enough, Nichols

(1992: 11) reports that her own results partly confirm some of the predictions drawn from, or the claims made in, Klimov's holistic typology although she argues that the morphological distinction between head and dependent marking (see 3.13) is a better predictor of other features than Klimov's notion of type.

The theoretical attractiveness of the Gabelentzian ideal notwithstanding, holistic typology is beset with at least one fundamental problem as Ramat (1986: 8–9) correctly identifies it: diachronic dynamics of language. Languages do shift from type X to type Y while retaining some features of type X. This fundamental fact of language alone casts much doubt on the feasibility of holistic typology, with there necessarily being languages of mixed types. A single underlying principle, which in holistic typology is used to classify the whole of a language into a given type, will not easily take adequate account of such mixed languages. This problem indeed harks back to the difficulty inherent in the classical typology as discussed above (Sapir 1921).

1.7 Organization of the rest of the book

In Chapter 2, beginning with Greenberg's (1963b) ground-breaking article, a number of cross-linguistic word order studies will be discussed and critically examined. In so doing, word order patterns and relevant correlations between different word order parameters will be investigated not only at the clausal and phrasal levels but also at the morphological level (e.g. suffixing vs. prefixing). Reference will also be made to a number of factors ranging from structural ones to processing efficiency as plausible avenues of explanation of observed word order patterns and correlations.

In Chapter 3 attention is directed to the ways in which the languages of the world deal with the fundamental 'problem' of expressing 'who is doing X to whom'. A comprehensive survey of case marking systems will first be provided with a view to understanding each of these systems and also to gaining insight into possible functional factors underlying not only these case marking systems but also their distribution. Moreover, correlations between case marking and word order types will be explored.

In Chapter 4 various relativization strategies will be exemplified with special reference to the expression of the head NP. In addition a brief cross-linguistic survey of accessibility to relativization will be carried out based on the hierarchy of grammatical relations. Furthermore, discourse and processing motivations will be appealed to in attempting to explain a number of observed cross-linguistic patterns or propensities in relativization. Possible correlations between relative clause and basic word order types will also be discussed as a prelude to functional explanation of cross-linguistic patterns in relativization.

In Chapter 5, following discussion of the morphologically based typology of causative constructions, the hierarchy of grammatical relations (cf. Chapter 4) will be revisited in slightly modified form with particular reference to the syntax of the causee NP, especially in the context of morphological causativization. Also included in this chapter is discussion of causative types, and causation types, the interaction between which is most clearly reflected in the actual case marking of the causee NP, for instance.

In Chapter 6 possibilities of applying linguistic typology to other areas of linguistics will be explored, where and if possible, referring back to relevant discussions in the preceding chapters. The areas of linguistics to be discussed in this chapter are historical linguistics (especially linguistic reconstruction), first language acquisition and second language acquisition, the latter two in the context of accessibility to relativization – perhaps the most researched topic of language acquisition, with direct reference to linguistic typology. Discussion is also provided of potential contributions of linguistic typology – i.e. what Nichols (1992) refers to as population typology – to a better understanding of linguistic prehistory, the time depth of which may be inaccessible by means of the classic comparative-historical method.

Finally, in Chapter 7 a survey of non-Greenbergian approaches to modern linguistic typology is offered with a view to helping the reader to become aware of the basic aims, assumptions and achievements of three prominent European schools of linguistic typology: the Leningrad Typology Group, the Cologne UNITYP Group and the Prague School Typology. A comparison of the European schools, and the Greenbergian/American tradition of linguistic typology will also be provided with a view to highlighting both strengths, and weaknesses of these different approaches to modern linguistic typology, all having their roots in the nineteenth-century European tradition of linguistic typology.

Notes

1. The distinction between languages and dialects is a notoriously difficult one to draw. For instance, one of the most widely used criteria is *mutual intelligibility*. But that notion itself is a difficult one to define, adding to the difficulty of defining the distinction in question.
2. This example is often cited as an exceptionless language universal. But it is correct to say that there are a few verb-initial languages with postpositions, e.g. Yagua (Comrie 1988: 146). Dryer (1991: 448) adds three more counter-examples: N. Tepehuan, Cora and Guajajara. But it will be kept here as an exceptionless language universal for the sake of illustration.
3. Nichols (1992: 42) describes universal tendencies as properties or correlations favoured in languages independent of geography and genetic affiliation, and thus as universal preferences of the languages of the world.
4. It must be borne in mind that languages are not being compared and classified in their entirety, but only in terms of structural properties in question (see 1.6 for discussion of partial and holistic typology).
5. Examples of a complement clause with a verbal head, and that with a nominal head are exemplified within square brackets in (i) and (ii), respectively (Comrie and Horie 1995: 65–6).
 - (i) The teacher knows [that the student bought the book].
 - (ii) the declaration/knowledge/fact [that the student bought the book].
6. In fact, Warlpiri has one additional construction that is employed consistently for the expression of relative clause function, i.e. the nominalized non-finite clause (Mary Laughren, personal communication). Being subject to aspectual or temporal restrictions, however, this construction seems to be marked as opposed to the adjoined clause (Hale 1976: 83).
7. This principle was first introduced into the study of language by Neogrammarians from the natural science thesis of Hutton and Lyell. Karl Brugmann is quoted as saying (Collinge 1994: 1561): '[t]he psychological and physiological nature of man as speaker must have been essentially identical at all epochs'.
8. A related question will be: which stage of evolution in human language should be chosen as the 'target' stage?
9. One may choose to use the descriptive label 'an escape hatch', rather than 'a frame of reference'.
10. Perkins (1980: 56, 1992: 123–4), however, points out that this is not the best way of studying language universals because the assumption of the independence of cases (see below) is clearly not met in the universe of languages.
11. Perkins (1980) uses a universe of cultures in order to derive a sample of languages because 'it is more reasonable to expect that linguistic materials exist for cultures that have been studied by ethnographers than those that have not[;] the appearance of a culture on Murdock's list makes it considerably more likely that the corresponding linguistic materials exist than for languages chosen from a language list [e.g. Voegelin and Voegelin (1966)]' (Perkins 1992: 125).
12. Perkins (1992: 179–81) argues that Dryer's criticism is far from damaging to his sampling procedure because, although possibly related, the Mon-Khmer languages, for example, diverge in terms of the relationship between deictic grammaticalization and cultural complexity in the way that is predicted by his hypothesis. In other words, this divergence despite genetic relatedness is claimed instead to strengthen the proposed association between the linguistic, and cultural properties. Be that as it may, Perkins's argument is beside the point because Dryer's criticism is directed at Perkins's sampling procedure in general, not to a particular application

of that procedure (i.e. Perkins's own investigation). Therefore, Dryer's criticism remains valid.

13. In his subsequent work (1992), Dryer removes Southeast Asia & Oceania from Eurasia, and treats them as an independent *linguistic area*. Thus in Dryer (1992) there are six, not five, large areas.
14. The reason why the highest number of intermediate levels is taken as the basis of the factor of decreasing significance in (20) is for corresponding levels in each phylum to be treated alike (Rijkhoff *et al.* 1993: 200).
15. The *LINGUIST-LIST*, and its related resources are now accessible on the internet (<http://www.linguistlist.org/>), and the *ALT-LIST* and its resources are also accessible on the internet (<http://148.88.14.7:80/alt/>).
16. Shibatani and Bynon (1995b: 16), however, see the difference between partial and holistic typology to be largely a matter of degree.
17. Comrie (1989: 45), on the other hand, adopts 'fusional' in lieu of 'inflectional' because 'both [agglutinative] and fusional languages, as opposed to isolating languages, have inflection, and it is [...] misleading to use a term based on (*in*)*flexion* to refer to one only of these two types'.