Linguistic Universals

edited by

Ricardo Mairal

and

Juana Gil

2   Linguistic typology

Kees Hengeveld

1   Introduction

This chapter discusses the contribution of linguistic typology to the study of language universals. Language universals research and linguistic typology are closely related fields, and are often not distinguished very clearly. Yet the difference between them can be characterized in the following way (Comrie, 1989): language universals define the restrictions on cross-linguistic variation; linguistic typology studies the restrictions on cross-linguistic variation; so typological research can be seen as the primary method used in uncovering language universals.

After a brief description of the way in which language samples are selected in cross-linguistic research in Section 2, the basic concepts and the methodology used in linguistic typology are introduced in Section 3, in which the notion of *implicational hierarchy* is taken as the point of departure. The implicational hierarchies uncovered through typological research are generally assumed to reflect true language universals, which means that they may be expected to show up in other linguistic domains as well, such as the historical development of languages, the process of language acquisition, language contact phenomena, and the distribution of linguistic phenomena within a single language. Section 4 of this chapter is dedicated to this issue. Conclusions are presented in Section 5. Wherever possible, the examples I present are taken from my own work, since this allows me direct access to the primary data.

2   Language sampling

Typological investigations make use of representative samples of the approximately 6,000 languages of the world. In order to guarantee the representativeness of the sample, three factors have to be taken into account: a sample should be representative from a genetic, an areal, and a typological perspective.

Languages may share properties because they belong to the same language family. For this reason the languages of the sample have to be selected in such a way that the genetic distance between them is maximal. The classification of languages through genetic reconstruction is complicated, particularly in the case of languages with an exclusively oral tradition. The existing classifications are therefore always tentative and reveal important differences of opinion between specialists.

Languages may also share properties because they are spoken in contiguous or (partially) coextensive areas. For this reason, the languages of the sample have to be selected in such a way that the geographical distance between them is maximal.

Finally, languages may share properties because they are of the same linguistic type. This factor is very difficult to control, since it is still unclear which typological properties should be decisive in assigning a language to a certain type. For the moment, the typological features most often used in controlling typological bias are the morphological type and the basic word order of a language.

One method of drawing representative language samples of various sizes in terms of genetic distribution is that proposed in Rijkhoff *et al.* (1993) and Rijkhoff and Bakker (1998). This method assigns a numerical value to a language family on the basis of the structure and complexity of its genetic tree. This value is then used to calculate the number of languages by which that family should be represented within a sample of a given size. The method is applied recursively to determine the number of languages by which a sub-family should be represented within the sample, etc. An advantage of this method is that it can be applied to any genetic classification. An example of a language sample[1] created using this method, applied to Ruhlen's (1987) genetic classification of languages of the world, is given in Table 2.1.

3   Implicational hierarchies

3.1   Introduction

The main tool that is applied in the study of language universals in linguistic typology is the *implicational hierarchy*. The general properties of implicational hierarchies are presented in Section 3.2, their application to various components of the language system in Section 3.3, and the level of abstraction of these hierarchies in Section 3.4. Sections 3.5 and 3.6 then deal with the interaction between hierarchies and the nature of the explanations given for them, respectively.

[1] This is the sample that is used in Hengeveld *et al.* (2004).

Table 2.1 *Example of a language sample*

Phylum	Subgroup			Language
Afro-Asiatic (2)	Chadic (1)			*Gude*
	Cushitic (1)			*Oromo, Boraana*
Altaic (1)	Northern (2)			*Turkish*
Amerind (7)		Almosan-Keresiouan (1)		*Tuscarora*
		Penutian (1)		*Koasati*
	Andean (1)			*Quechua, Imbabura*
	Equatorial-Tucanoan (1)			*Guarani*
	Ge-Pano-Carib (1)			*Hixkaryana*
	Central Amerind (1)			*Pipil*
	Chibchan-Paezan (1)			*Warao*
Australian (3)	Gunwinyguan (1)			*Ngalakan*
	Pama-Nyungan (1)			*Kayardild*
	Nunggubuyu (1)			*Nunggubuyu*
Austric (6)	Austro-Tai (4)	Daic (1)		*Nung*
		Austronesian (3)	Malayo-Polynesian (2) Western (1)	*Tagalog*
			Central Eastern (1)	*Samoan*
			Paiwanic (1)	*Paiwan*
	Austroasiatic (1)			*Mundari*
	Miao-Yao (1)			*Miao*
Basque (1)				*Basque*
Burushaski (1)				*Burushaski, Hunza*
Caucasian (1)				*Abkhaz*
Chukchi-Kamchatkan (1)				*Itelmen*
Elamo-Dravidian (1)				*Tamil*
Eskimo-Aleut (1)				*West Greenlandic*
Etruscan (1)				*(Etruscan)*[2]
Hurrian (1)				*Hurrian*
Indo-Hittite (2)	Indo-European (1)			*Polish*
	Anatolian (1)			*Hittite*
Indo-Pacific (5)	Trans New Guinea (1)			*Wambon*
	Sepik-Ramu (1)			*Alamblak*
	East Papuan (1)			*Nasioi*
	West Papuan (1)			*Tidore*
	Torricelli (1)			*Arapesh, Mountain*
Kartvelian (1)				*Georgian*
Ket (1)				*Ket*
Khoisan (1)				*Nama Hottentot*
Korean-Japanese-Ainu (1)				*Japanese*
Meroitic (1)				*(Meroitic)*
Na-Dene (1)				*Navajo*
Nahali (1)				*(Nahali)*
Niger-Kordofanian (4)	Niger-Congo (3)	Niger-Congo Proper (2)	Central Niger-Congo (1)	*Babungo*
			West Atlantic (1)	*Kisi*
		Mande (1)		*Bambara*
	Kordofanian (1)			*Krongo*
Nilo-Saharan (2)	East Sudanic (1)			*Lango*
	Central Sudanic (1)			*Ngiti*
Nivkh (1)				*Nivkh*
Pidgins and Creoles (1)				*Berbice Dutch*
Sino-Tibetan (2)	Sinitic (1)			*Chinese, Mandarin*
	Tibeto-Karen (1)			*Garo*
Sumerian (1)				*Sumerian*
Uralic-Yukaghir (1)				*Hungarian*

[2] Parentheses around language name indicate that insufficient data are available for a language that should be in the sample according to the sampling method.

3.2   *Universal implications and implicational hierarchies*

Language universals are generally expressed in the form of universal implications, which are most often unidirectional. An abstract example of a unidirectional universal implication is given in (1):

(1)   A > B

This universal implication defines the possible combinations of properties A and B listed in (2):

(2)   A   B
                 +   +
                 +   −
                 −   −
          *   −   +

Since the presence of property B in a language implies the presence of property A in that same language, but the absence of B does not imply the absence of A, the only combination of features that is excluded by the implication is the absence of A in the presence of B. Thus, of the logically possible combinations ($2^2 = 4$), one (i.e. 25%) is excluded.

Some concrete examples of universal implications are the following, taken from Keenan and Comrie (1977) and Comrie (1989) (see also Bakker and Hengeveld, 1999):

(3)   (a) Subject   > Direct object
         (b) Direct object   > Oblique object
         (c) Oblique object > Possessor

These implications were proposed to account for, among other things, the variation in the degree to which languages permit the relativization of constituents that fulfil the relevant syntactic and semantic functions within the relative clause. For instance, languages that permit the relativization of direct objects will also allow the relativization of subjects (3a); those that do not permit the relativization of direct objects will not allow the relativization of oblique objects either (3b); and those that permit the relativization of possessors will allow the relativization of oblique objects (3c).

The implied category in (3c) is the implying category in (3b), and the implied category in (3b) is the implying category in (3a), so that the universal implications in (3) can be combined into the chain in (4):

(4)   S   >   DO   >   OO   >   Poss
         +   +   +   +
         +   +   +   −
         +   +   −   −
         +   −   −   −
         −   −   −   −
    *   +   −   +   −
    *   etc.

A chain of universal implications such as the one in (4) is called an "implicational hierarchy." In this case, 11 (67.5%) out of all logically possible combinations are excluded. This shows that the higher the number of categories that can be combined within a single hierarchy, the more constrained the typological description of cross-linguistic variation will be.

In general, two types of implicational hierarchy are distinguished: absolute and statistical ones. Absolute implicational hierarchies are valid for all languages, statistical ones for a significant proportion of all languages. It is important to note that absolute implicational hierarchies are absolute only in relation to our current knowledge of linguistic variation, since, on the one hand, as mentioned in Section 2, typological research is based on samples of the world's languages, and, on the other hand, many of the languages of the world lack documentation. Another limitation that has to be taken into account is that implicational hierarchies are normally based on a sample of oral languages: the study of the typology of sign languages is still in its infancy.

3.3   *Implicational hierarchies and the components of the language system*

3.3.1   *Introduction*

Implicational hierarchies are a useful tool in capturing linguistic generalizations pertaining to the different components of the language system. This will be illustrated in this section through the analysis of examples from the areas of phonology (3.3.2), morphology (3.3.3), syntax (3.3.4), semantics (3.3.5), and the lexicon (3.3.6).

3.3.2   *Phonology*

In Section 3.4 a detailed example of a possible phonological universal will be given. I therefore limit myself here to an example of an implicational hierarchy concerning the distribution of nasal phonemes across languages, presented in (5), which concerns dental/alveolar, bilabial, and palatal voiced nasals, respectively:

(5)   /n/ > /m/ > /ɲ/

This hierarchy defines the following possible combinations of nasal consonants in the phoneme inventory of a language:

(6)   /n/ /m/ /ɲ/
        /n/ /m/ –
        /n/ –   –
        –   –   –

Note that this hierarchy, like most typological hierarchies, is not exhaustive. It does not predict, for instance, under what conditions a language might have the voiced velar nasal /ŋ/, which does not form part of the hierarchy in (5).

3.3.3   Morphology

Important implicational hierarchies in the area of morphology concern the order of inflectional morphemes expressing tense, aspect, and modality. Bybee (1985) proposes the generalization in (7):

(7)
```
     3      2      1     stem     1      2      3
      \      \      \____ aspect ____/      /      /
       \      _____ tense _____/      /
        _____ mood _____/
```

This generalization predicts that aspect will be expressed closer to the (verb) stem than tense, which in turn will be expressed closer to the stem than mood.

Elaborating on Bybee's work, Hengeveld (1989) proposes a more detailed classification, partly inspired by Foley and Van Valin (1984). This classification is shown in (8):

(8)   5   4   3   2   1   **stem**   1   2   3   4   5
        1. qualitative aspect / agentive modality
        2. tense / realis–irrealis / quantitative aspect / negation
        3. evidentiality
        4. illocution
        5. mitigation–reinforcement

The classification in (8) situates the categories of aspect and mood in various positions, depending on the subcategories involved. The position of these subcategories within the hierarchy reflects their scope within the sentence. Thus, qualitative aspect occupies the position closest to the predicate because it affects just the predicate, while quantitative aspect occupies the second position, since it modifies the event described within the sentence.

The morpheme order in example (9) obeys the hierarchy in (8):[3]

(9)   Turkish (Altaic)
        Anli-y-abil-ecek-miş-im.
        understand-*y*-ABIL-IRR-INFER-1.SG
        "I gather I will be able to understand."

In this example the suffix -*abil* expresses ability (ABIL), an agentive modality; the suffix -*ecek* expresses irrealis (IRR), which in this case is translated as a future tense; and the suffix -*miş* expresses inference (INFER), an evidential modality. Applying the classification given in (8) to these categories, example (9) reflects the hierarchy in (8) in the following way:

(10)   **stem** -1   -2   -3
        Anli-y-abil-ecek-miş-im.
        1. agentive modality
        2. realis–irrealis
        3. evidentiality

Example (11) shows that the hierarchy in (8) can be applied to a language that is structurally very different:

(11)   Hidatsa (Siouan; Matthews, 1963)
        Wíra i   apaari ki   stao   ski.
        tree it grow   INCH   REM PAST   CERT
        "The tree must have begun to grow a long time ago."

In Hidatsa the categories of tense, mood, and aspect are expressed through grammatical particles, rather than through affixes as in Turkish. In example (11), the particle *ki* expresses inchoative aspect (INCH), which is a qualitative aspect; the particle *stao* expresses remote past (REM.PAST); and the particle *ski* expresses certainty (CERT) on the part of the speaker, an evidential modality. This example thus reflects the hierarchy in (8) in the following way:

(12)   **stem**   1   2   3
        Wíra i apaari ki stao ski.
        1. qualitative aspect
        2. tense
        3. evidentiality

The translation of example (11) into English shows that even in a language in which these categories are expressed through auxiliary verbs the hierarchy in (8) is respected, albeit in the opposite order:

---

[3] The -*y*- is meaningless. It is inserted for phonological reasons.

(13)   The tree must have begun to grow.
                3   2   1   **stem**

The predictive force of the hierarchy is further corroborated by the fact that the order of the auxiliaries cannot be altered, as can be observed in the following examples:

(14)   (a) *The tree must begin to have grown.
        (b) *The tree has must begin to grow.
        (c) *The tree has begun to must grow.
        (d) *The tree begins to must have grown.
        (e) *The tree begins to have must grow.

The examples from Hidatsa and English serve to illustrate that universals such as the one in (8), which was originally proposed in the area of morphology (affix order), can be applied in the area of syntax (particle and auxiliary order) too, since this universal is semantically conditioned.

3.3.4   Syntax

Many of the studies that have appeared in the field of syntactic typology are dedicated to establishing generalizations with respect to word order. As an example of a universal in this area consider the order of verb and object in relation to the order of standard and quality in comparative constructions. As shown by Greenberg (1966), the relation between these two parameters is as in (15):

(15)   (a) Object–Verb ⬦ Standard–Quality
        (b) Verb–Object ⬦ Quality–Standard

In this case the implication is bi-directional, i.e. the properties specified within the implication always occur together. It is furthermore important to note that although all languages have ways of expressing comparison, not all languages have a comparative construction. The generalizations in (15) are only valid for languages with such a construction.[4]

The examples in (16) and (17) illustrate the two combinations of properties given in (15). Turkish (16) has the combination of properties of type (15a), English (17) of type (15b):

(16)   *Turkish*
        (a) Object–Verb
            Ressam bize resimlerini$_{Object}$ *gösterdi*$_{Verb}$.
            artist   us   his.paintings   showed
            "The artist showed us his paintings."

---

[4] This statement may be generalized to all implicational hierarchies; they are only valid to the extent that the categories contained within them are relevant to the languages under consideration.

        (b) Standard–Quality
            kurşun-dan$_{Standard}$ *ağır*$_{Quality}$
            lead-ABL   heavy
            "heavier than lead"

(17)   *English*
        (a) Verb–Object
            The artist showed$_{Verb}$ us his paintings$_{Object}$.
        (b) Quality–Standard
            heavier$_{Quality}$ than lead$_{Standard}$

The correlation between the features illustrated in these examples appears to be absolute or nearly absolute, and is one of the main indicators of the basic word order of a language.

3.3.5   Semantics

One of the best known hierarchies in the area of semantics is the color term hierarchy proposed in Berlin and Kay (1969), which is given in (18). This hierarchy says that if a language has just two basic color terms, these will be the equivalents of "white" and "black" (or rather "light" and "dark"); if it has three, the third one will be the equivalent of "red"; etc. For each of the different systems defined by the hierarchy, an example of a language manifesting the relevant combination of basic color terms is given.

(18)   *The color term hierarchy*
        white   > red   > yellow   > yellow   > blue   > brown
        and black   or green   and green
        Dani   >
                    Koromfe >
                            Babungo >
                            Kobon   >
                                        Wari   >
                                            Nkore-Kiga >
                                                        Punjabi

In more recent studies in semantic typology, a new method has been introduced that is based on the concept of "semantic map." A semantic map organizes sets of related meanings in such a way that the different means that languages use to express those meanings consistently occupy contiguous areas on that map. By way of illustration, consider the following series of constructions and the verbs that are used within them:

Table 2.2 *Example of a semantic map*

Locative/+pres.			Possessive/+pres.
Locative/–pres.	Property	Status	Possessive/–pres.

Table 2.3 *Turkish*

Locative/+pres.			Possessive/+pres.
Locative/–pres.	Property	Status	Possessive/–pres.

Table 2.4 *Babungo*

Locative/+pres.			Possessive/+pres.
Locative/–pres.	Property	Status	Possessive/–pres.

(19)   Locative/+presentative
         "There *is* a woman in the garden."

(20)   Locative/–presentative
         "The woman *is* in the garden."

(21)   Possessive/+presentative
         "I *have* a book."

(22)   Possessive/–presentative
         "The book *is* mine."

(23)   Property
         "The woman *is* beautiful."

(24)   Status
         "Charles *is* president."

Hengeveld (1992) shows that these constructions can be organized in the semantic map given in Table 2.2.

Languages subdivide this semantic space in different ways. In Turkish (Table 2.3), the two presentative constructions are expressed in one way, and all the non-presentative ones in another.

In Babungo (Table 2.4) locative constructions use one strategy and all other constructions another.

Table 2.5 *Nasioi*

Locative/+pres.			Possessive/+pres.
Locative/–pres.	Property	Status	Locative/–pres.

Table 2.6 *Spanish*

Locative/+pres. *Haber*				Possessive/+pres. *Tener*
Locative/–pres. *Estar*	Property *Estar*	Property *Ser*	Status *Ser*	Possessive/–pres. *Ser*

Nasioi (Table 2.5) draws the line between adjectival and nominal constructions on the one hand, and locative and possessive ones on the other.

Spanish uses a large number of verbs in expressing the various construction types. All of these occupy contiguous spaces, as shown in Table 2.6.

More detailed examples of the application of the semantic map methodology may be found in Haspelmath (1997), van der Auwera and Plungian (1998), and Croft (2002).

3.3.6   Lexicon

Not only lexical semantics (3.3.5), but the organization of the lexicon in terms of categories and subcategories has been the object of typological study too. In particular, the presence and absence of certain parts of speech have drawn the attention of typologists. Within this field, Hengeveld *et al.* (2004), following up on Hengeveld (1992), propose the hierarchy in (25):

(25)   Parts-of-speech hierarchy

Head of	>	Head of	>	Modifier of	>	Modifier of
predicate		referential		referential		predicate
phrase		phrase		phrase		phrase

This hierarchy should be interpreted in the following way. The part of speech that languages are most likely to have is a class of verbs, i.e. a set of lexemes that are exclusively used in predicative function. Languages are least likely to have a class of manner adverbs, i.e. a set of lexemes that are exclusively used to modify the head of a predicate phrase. Since languages may furthermore either have lexemes that can be used within various syntactic slots (flexible systems) or simply lack lexemes for certain syntactic slots (rigid systems), the hierarchy in (25) defines seven different parts-of-speech systems, which are presented in Table 2.7.

Table 2.7 *Parts-of-speech systems*

Parts-of-speech system		Head of predicate phrase	Head of referential phrase	Modifier of referential phrase	Modifier of predicate phrase
Flexible	1	lexeme			
	2	verb	non-verb		
	3	verb	noun	modifier	
Differentiated	4	verb	noun	adjective	manner adverb
Rigid	5	verb	noun	adjective	–
	6	verb	noun	–	–
	7	verb	–	–	–

By way of illustration, consider the following examples:

Warao (Chibchan-Paezan; Romero-Figueroa, 1997, 50, 119)
(26)   *yakera*
beauty
"beauty"

(27)   *Hiaka yakera auka   saba tai   nisa-n-a-e.*
dress beauty daughter for   her   buy-SG-PF-PAST
"She bought a beautiful dress for her daughter."

(28)   *Oko yakera nahoro-te.*
we   beauty eat-NONPAST
"We eat well."

In Warao, there is a class of verbs – e.g. *nisa-* in (27) – and a class of non-verbs. One and the same lexeme, *yakera*, is used in (26) as the head of a referential phrase, in (27) as a modifier within a referential phrase, and in (28) as a modifier within a predicate phrase. For this reason, Warao may be classified as a type 2 language in terms of the classification in Table 2.7.

3.4   Level of abstraction

Implicational hierarchies can only be successful predictors of linguistic variation to the extent that the categories, functions, or constructions that they try to capture can be identified cross-linguistically. As a result, the categories, functions, and constructions under investigation should be defined at a sufficiently abstract level. On the other hand, implicational hierarchies have to be sufficiently concrete in order to be applicable to actual linguistic data. The ideal

Table 2.8 *Three consonants*

	labial	coronal	dorsal	glottal
plosive	**X** (p)	**X** (t)	**X** (k)	
fricative				
nasal				
liquid				

Table 2.9 *Six consonants*

	labial	coronal	dorsal	glottal
plosive	X (p)	X (t)	X (k)	
fricative		X (s)		
nasal	**X** (m)	**X** (n)		
liquid				

implicational hierarchy therefore reflects both a sufficient degree of abstractness and a sufficient degree of descriptive potential. An example from the area of phonology may help to illustrate this point.

The consonant inventories at the disposition of languages are notoriously variable. A typologist who tries to define implicational hierarchies based on specific phonemes will encounter few generalizations: it is difficult to decide which consonants are less marked than others. A famous exception is the hierarchy of nasal consonants presented in Section 3.3.2. At a higher level of abstraction, however, generalizations may be formulated in terms of combinations of distinctive features of consonant phonemes available in the languages under investigation. On the basis of a small sample, at this higher level of abstraction the following generalizations seem to hold. If a language has just three consonants, it is very likely that these will be characterized by the features represented in Table 2.8. Note that in each case an example of a consonant is given that exemplifies the combination of features under consideration. This specific consonant is, however, not necessarily a phoneme of the language under investigation. For instance, both /b/ and /p/ are instances of the combination of the features /labial/ and /plosive/.

If a language has three more consonants, it is probable that these will display the combinations of features printed in bold in Table 2.9.

The next two consonants would be likely to display the combinations of features that are given in Table 2.10.

And the next two might very well display the features in Table 2.11.

Table 2.10 *Eight consonants*

	labial	coronal	dorsal	glottal
plosive	X (p)	X (t)	X (k)	
fricative		X (s)		**X** (h)
nasal	X (m)	X (n)		
liquid		X (l)		

Table 2.11 *Ten consonants*

	labial	coronal	dorsal	glottal
plosive	X (p)	X (t)	X (k)	
fricative		X (s)	**X** (χ)	X (h)
nasal	X (m)	X (n)	**X** (ŋ)	
liquid		X (l)		

This tentative example shows that a phenomenon that at first sight seems to defy generalization does show systematicity at a higher level of abstraction.

3.5   *Interaction between implicational hierarchies*

In many cases, various implicational hierarchies are simultaneously operative within the same domain of grammar, and give rise to conflicting results. Depending on the weight languages assign to each of these counteracting hierarchies, the final outcome of their application may differ from language to language. This interaction between hierarchies unavoidably leads to counterexamples at the level of the individual hierarchy, which is the main reason for the existence of statistical hierarchies.

As an example of the interaction between hierarchies, consider the following two generalizations in the area of word order, as formulated in Dik (1997, 403–404):

(29)   The Principle of Functional Stability: constituents with the same functional specification are preferably placed in the same position.

(30)   The Principle of Increasing Complexity: there is a preference for ordering constituents in an order of increasing complexity.

As an illustration of the interaction between these hierarchies, consider the following examples:

Table 2.12 *Two-dimensional representation of parameters*

		Time dependency	
		+	−
Presupposedness	+	Dependent verb forms most likely	
	−		Independent verb forms most likely

(31)   *That he left was a pity.*

(32)   **Was a pity that he left.*

(33)   *It was a pity that he left.*

According to principle (29), a construction with a clausal subject in English should be expressed as in (31), in which the subject appears in preverbal position. This sentence is grammatical, although it does not display the preferred constituent order. According to principle (30), the same construction should be expressed as in (32), in which the clausal subject appears in final position. The preferred construction type in (33) can be seen as a compromise between the two counteracting principles: the expletive pronoun *it* occupies the subject position, following principle (29), and announces the subject clause that appears in final position, following principle (30).

In order to account for the interaction between implicational hierarchies one needs a multidimensional model, which is difficult to represent graphically in those cases in which more than two parameters are involved. For a simple example involving two parameters, consider Table 2.12, adapted from Hengeveld (1998, 379).

In Table 2.12 two parameters that exert influence on the expression of the verb forms used in subordinate clauses are represented. Verb forms are classified as dependent or independent. A dependent verb form is one that cannot be used in a main clause, while an independent verb form is one that could also be used in a main clause. The parameter represented horizontally states that subordinate clauses with dependent time reference, i.e. which depend for their temporal interpretation on the main clause, are more likely to be expressed by means of dependent verb forms than those with independent time reference. The parameter represented vertically states that subordinate clauses with a presupposed content are more likely to be expressed by means of dependent verb forms than those with a non-presupposed content. The combination of these

Table 2.13

	Time dependency		
		+	−
Presupposedness	+	Dependent verb forms	Independent verb forms
	−		

Table 2.14

	Time dependency		
		+	−
Presupposedness	+	Dependent verb forms	
	−		Independent verb forms

Table 2.15

	Time dependency		
		+	−
Presupposedness	+	Dependent verb forms	Independent verb forms
	−		

two parameters predicts that presupposed time-dependent subordinate clauses are most likely to be expressed by means of dependent verb forms, while subordinate clauses with the opposite values are most likely to be expressed by means of independent verb forms.

Depending on the weight languages assign to the different parameters, the resulting systems could, for example, be of any of the types represented in Tables 2.13–2.16.

3.6   *The interpretation of implicational hierarchies*

An important general property of implicational hierarchies is that, all other things being equal, features more to the right on the hierarchy are less likely to occur in language systems. A question that arises immediately when an

Table 2.16

	Time dependency		
		+	−
Presupposedness	+	Dependent verb forms	
	−		Independent verb forms

implicational universal is discovered is therefore how this increasing markedness of features can be explained in terms of the constraints that are imposed on possible human languages. Explanations given to language universals generally make reference either to (i) cognitive constraints (e.g. ease of processing, innate faculties, iconicity, saliency in perception), or to (ii) communicative needs (e.g. saliency of information, disambiguation, economy).

Another important aspect of the interpretation of typological generalizations concerns the extent to which these are used in the elaboration of a comprehensive theoretical model of language. Often linguistic typology is seen as a goal in itself: the research has a purely descriptive aim. In other cases linguistic typology is seen as a method that can be applied in developing (part of) a linguistic theory, which incorporates typological results to give rise to new predictions. In the first case, the primary goal of the research is an exhaustive characterization of languages; in the second case, an exhaustive characterization of language.

4   Linguistic typology and other branches of linguistics

4.1   *Introduction*

Typological hierarchies reflect universal properties of language and may therefore be assumed to define restrictions on cross-linguistic variation. But if these hierarchies really have a universal value, they should not only manifest themselves in cross-linguistic variation, but also restrict intralinguistic variation, such as, for instance, diachronic variation, the different phases in language acquisition, the effects of language contact, and the quantitative distribution of phenomena within a single language. In this section, I present a number of examples of intralinguistic phenomena that may be interpreted as the result of the influence of some of the same implicational hierarchies that were introduced earlier in this chapter in relation to cross-linguistic facts.[5]

[5] For a more detailed study of the relation between cross-linguistic and intralinguistic variation, see van Lier (2005).

4.2   Linguistic typology and historical linguistics

In (8) in Section 3.3.3 I presented an implicational hierarchy in the area of morphology that defines restrictions on the order in which various grammatical categories may be expressed with respect to the predicate. The categories are defined in semantic terms. From a diachronic perspective this hierarchy is relevant too, in the sense that it defines the possible changes in meaning that a grammatical element may undergo in the course of time. A concrete example of this is the diachronic development of the auxiliary verb *will* in English, as described in Bybee *et al.* (1991). In a first stage, this verb only expressed the agentive modalities Intention and Obligation, which pertain to the first category of the hierarchy. In the second stage, the verb acquired the temporal meaning Future, which belongs to the second category of the hierarchy. In the third phase the verb expressed the evidential modality Prediction as well. This modality falls in the third category of the hierarchy. Thus, the changes in meaning that this verb underwent concern contiguous categories of the hierarchy, and follow the order predicted by the hierarchy.

4.3   Linguistic typology and language acquisition

The same hierarchy is relevant in the area of first language acquisition too. Various studies concerning the acquisition of Turkish (Ekmekci, 1979; Aksu-Koç and Slobin, 1985) show that the first category in the area of tense, mood, and aspect that Turkish children acquire is the durative aspect, which pertains to the first category of the hierarchy in (8). The past tense, which belongs to the second category of the hierarchy is acquired later, but before the inferential mood, which falls into the third category of the hierarchy. Thus, in this case the process of first language acquisition respects the order and the contiguity of categories in a morphological hierarchy that was developed to capture constraints on cross-linguistic variation.

4.4   Linguistic typology and language contact

Hierarchy (25) in Section 3.3.6 predicts the possible combinations of parts of speech in a language. These combinations vary from one to four of the parts of speech that the hierarchy takes into consideration. An interesting question resulting from this typological observation is what happens when languages with less than these four parts of speech enter into contact with a language with all these four parts of speech and start incorporating loans from this language. If the hierarchy is correct, it is to be expected that in such a situation languages will tend to incorporate items pertaining to the first category of the hierarchy that is lacking in their native lexicon in a significantly high degree.

This hypothesis may be verified, for instance, by studying the effects that contact with Spanish has had on various typologically different indigenous languages of Central and South America. A few isolated examples seem to indicate that the typological difference between Spanish and the indigenous language determines which part of speech is borrowed more frequently.

In Quechua there is no clear-cut distinction between Nouns and Adjectives, so that one can say that only the first two categories on the hierarchy are covered. Hekking and Muysken (1995) show that the number of borrowed Spanish Adjectives in this language is surprisingly high, which indicates that the lexicon is being extended with elements from the third category of the hierarchy.

In Pipil there is no clear-cut distinction between Adjectives and manner adverbs in the indigenous vocabulary, so in this language the first three categories of the hierarchy are covered. Campbell's (1985) dictionary shows clearly that the number of manner adverbs borrowed from Spanish is surprisingly high, which indicates that the lexicon is being extended with elements from the fourth category of the hierarchy.

4.5   Linguistic typology and descriptive linguistics

Table 2.12 in Section 3.5 presented two hierarchies that influence the choice of dependent or independent verb forms in subordinate clauses. Many languages of Western Europe allow the use of both dependent and independent verb forms in the same type of construction. The following examples show that one and the same type of adverbial construction may be expressed with a dependent verb form (34a) or an independent verb form (34b):

(34)   a. *Apart from doing the cooking, I look after the garden.*
         b. *Apart from the fact that I do the cooking, I look after the garden.*

Under the hypothesis that typological hierarchies are reflected in intralinguistic variation, it is to be expected that the selection of one of the two alternatives available for expressing a subordinate clause will correspond quantitatively to what is predicted in Table 2.12.

On the basis of a large-scale corpus investigation of Modern English, Pérez Quintero (2002, 132) presents the results given in Table 2.17 for the choice of dependent and independent verb forms in adverbial clauses of the types under consideration. The percentages refer to the proportion of dependent verb forms used.

The percentages in boldface clearly show that dependent verb forms are used much more frequently in constructions that, from a cross-linguistic perspective, are most likely to be expressed by dependent verb forms, and the opposite is true for independent verb forms.

Table 2.17

		Time dependency	
		+	–
Presupposedness	+	**54.39%**	14.58%
	–	40.43%	**6.98%**

5   Conclusion

Linguistic typology studies the restrictions on cross-linguistic variation. These restrictions are formalized as implicational hierarchies, which do not only manifest themselves in the differences between languages, but also in the restrictions that they impose on intra-linguistic variation, since implicational hierarchies represent language universals that manifest themselves in all language systems.

3   Universals in a generative setting

Cedric Boeckx

Scientific endeavors tend, by their very nature, to focus on universal statements like (1), as opposed to existential statements like (2).

(1) $\forall x \ldots [\ldots x \ldots]$

(2) $\exists x \ldots [\ldots x \ldots]$

Statements like (1) give rise to universal formulations or "laws." In the context of the scientific study of language, universal statements (henceforth, "universals") have been formulated along the axes that Gould (2002, 259) identifies as influences underlying the genesis of natural objects: historical, functional, and formal.

In this chapter, I will focus exclusively on the formal axis. This is not to suggest that other axes are unimportant. It is simply that universals have a much more precise meaning in the context of formal approaches to language, by which I mean "Generative Grammar" (GG), hence they allow for clearer exposition.[1]

Let me start by pointing out that GG is a topic, a perspective on language, one that, following a long tradition ("Cartesian linguistics," see Chomsky, 1966), takes language to be a species-specific property, an object of the natural world to be studied like any other ("methodological naturalism," see Chomsky, 2000a). This makes linguistics a part of cognitive science, and ultimately biology.

This biolinguistic approach, as Jenkins (2000) has aptly called it, emerged more and more clearly as research in GG advanced.

The first chapter of *Aspects of a Theory of Syntax* (Chomsky, 1965; hereafter, *Aspects*) firmly places the study of language in a cognitive, and ultimately biological, setting, and arguably remains to this day the clearest statement of the generative enterprise as a whole. Here Chomsky argues that the central problem of linguistics is to account for how children are able to acquire their native languages.

[1] I also think that the formal approach, inspired by the goals of Generative Grammar, takes epistemological priority over other aspects of language. But this is not a theme that I will be able to develop here. For some discussion, see Chomsky (1986).