

7 Language typology

A **TYPOLGY** is simply a categorization of some range of phenomena into various types. To “typologize” something is to group its parts into types. For example, we often hear jokes like the following: “There are three kinds of people – those who can count, and those who can’t.” Typological linguists are people who like to group languages into well-defined and useful types.

But what makes a typology useful? A typology is useful when it makes “predictions” about multiple characteristics of the items being typologized. For example, suppose we were to typologize motorized vehicles. Which would be the most meaningful typology, A or B?:

- Typology A: bus, van, automobile, tractor
- Typology B: red ones, green ones, blue ones, white ones

If you know that a motor vehicle is a bus, what else do you know about it? Quite a lot actually – it is probably going to be a large vehicle, with lots of seats, designed primarily to carry people, etc. If, on the other hand, you know some random motor vehicle is blue in color, there is not much else you can guess about its characteristics. Therefore, typology A is more useful, because it reflects “clusters” of structural and functional characteristics that go together, rather than simply indicating isolated properties.

Turning to a linguistic example, we could say that there are two kinds of languages in the world – those that have the sound [r] in their phonetic inventory and those that don’t. However, knowing whether a language has an [r] is not likely to have many repercussions in other parts of the language, therefore this is not a particularly interesting or useful typology. However, there are several other linguistic typologies that have been very helpful to people interested in exploring the characteristics of the human mind. These are typologies that identify *clusters* of characteristics that languages are likely to possess.

The value of typologizing languages is that it helps linguists understand the range and limits of possible variation among human languages. If logically possible types are found to be very rare or nonexistent, that may provide some insight into how the human mind works. Thus language typology can give us a “window” on the mind and communication. To extend our non-linguistic example, if we typologized all the motorized vehicles in the world according to number of wheels, we might find that there are no, or extremely few, vehicles with five wheels. This fact would invite us to investigate *why* motorized vehicles are restricted in exactly

this respect. What is it about the origin, history, or function of motor vehicles that seems to rule out the existence of five-wheeled vehicles?

Several typologies of language have been proposed in the history of linguistic science. In this chapter, we will discuss morphological and syntactic typology. In later chapters we will discuss a typology of grammatical relations (chapter 8), voice and valence (chapter 9), and clause combining (chapter 10). Syntactic typology has proven particularly fruitful in stimulating the subfields of **TYPOLOGICAL LINGUISTICS**, and **FUNCTIONAL LINGUISTICS**.

Morphological typology

There are two parameters by which the morphological typology of a language may be measured. These are described by Comrie (1989) as the **INDEX OF SYNTHESIS** and the **INDEX OF FUSION**. The index of synthesis refers to how many morphemes tend to occur per word in a language, while the index of fusion refers to how many meanings tend to be associated with each morpheme.

The index of synthesis defines a continuum from **ISOLATING** languages at one extreme to highly **POLYSYNTHETIC** languages at the other. Figure 7.1 illustrates this continuum.



Figure 7.1 *The index of synthesis (the number of morphemes per word)*

A strictly isolating language is one in which every word consists of only one morpheme. The Chinese languages come close to this extreme. A highly polysynthetic language is one in which words tend to consist of several morphemes. The Quechuan and Eskimo-Aleut languages are good examples of highly polysynthetic languages. The following is an example of a polysynthetic structure in Central Yup'ik (thanks to Eliza Orr):

- (1) Tuntussuqatarniksaitengqiggtuq
 tuntu-ssur-qatar-ni-ksaite-ngqiggte-uq.
 reindeer-hunt-FUT-say-NEG-again-3SG.IND
 'He had not yet said again that he was going to hunt reindeer.'

The index of fusion (figure 7.2) describes a continuum between highly **AGGLUTINATIVE** languages to highly **FUSIONAL** languages. A highly agglutinative language is one in which most morphemes express one and only one meaning. A highly fusional language (sometimes called “inflectional,” but since this has other connotations, we will use the term fusional) is one in which morphemes often express several meanings. For example, in Spanish the suffix *-ó* in a word like *habló* expresses at least five conceptual categories: indicative mood, third person, singular, past tense, and perfective aspect. If any one of these conceptual

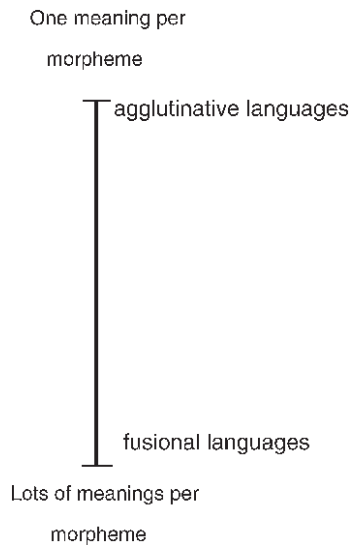


Figure 7.2 *The index of fusion*

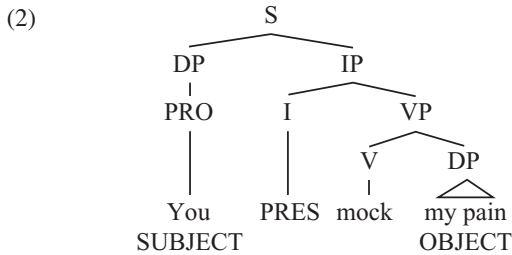
categories changes, the form of the suffix must change. Turkish is a language for which each lexical meaning and conceptual category is, in general, expressed by its own morpheme. Therefore, Turkish is a highly agglutinative language. For highly isolating languages, the index of fusion just doesn't apply. If anything, English is agglutinative rather than fusional, e.g., in *anti-dis-establish-ment-arian-ism* each morpheme has a specific and fairly clear meaning. But then, such words in English are mostly of Latin origin. Fusion is apparent in English in the present tense, third person, singular suffix *-s*, as in *he walks the line*, and in the paradigm for the verb *be*, but not much else.

There is no generally accepted quantitative method for precisely establishing the indices of synthesis and fusion for a given language. A rule of thumb for the index of synthesis is that if the language can express a whole sentence with just a verb, it is polysynthetic. If it can't, then it is isolating. Adjectives such as "somewhat" or "highly" can then be added in order to give a sense of where a language falls on each continuum, e.g., English is "somewhat isolating," Mandarin is "highly isolating." Turkish is "somewhat polysynthetic and highly agglutinative" while Yup'ik is "highly polysynthetic and somewhat fusional." Knowing something about the morphological typology of a language helps linguists make better hypotheses about the likely meanings of various structures and helps tremendously in understanding the historical roots and development of a language.

Syntactic typology

Linguists have long noticed that some languages tend to place the verb at the end of a clause, others at the beginning, still others place it somewhere in the middle. Finally, many languages seem to place the verb just about anywhere.

Among the nominal (“noun-like”) elements in a clause, an important distinction has traditionally been made between subject and object (abbreviated S and O in early typological research).¹ In terms of tree diagrams, you can think of the subject as the DP that is directly under the S node, and the object as the DP that is directly under the VP node:



It turns out that there is a very major **TYPOLGICAL** distinction between languages in which the object follows the verb (VO languages), and those in which the object precedes the verb (OV languages; Greenberg 1963, *inter alia*). In terms of phrase structure rules, this can be thought of as a distinction between languages, like English, in which the VP rule has an optional DP following the head verb, and others, like Japanese, in which the optional DP precedes the head verb:

- (3)
- | | | | |
|----|---|--------|--|
| VP | → | (DP) V | OV Languages (Japanese, Finnish, Hindi . . .) |
| VP | → | V (DP) | VO Languages (Mandarin, Indonesian, English . . .) |

What is interesting about this typology is that the order of object and verb in the verb phrase tends to correlate with other aspects of the syntax of the language. For example, if a language has OV order, it will almost certainly have postpositions, rather than prepositions. Conversely, if a language has VO order, it will almost certainly have prepositions. Also, in OV languages, inflected auxiliaries almost always come after the verb, whereas in VO languages, auxiliaries usually precede the verb. In short, there are, generally speaking, two major types of languages in the world: those in which syntactic heads normally precede their complements, and those in which syntactic heads follow their complements:

- (4)
- | VO languages: | | | OV languages: | |
|---------------|-------|------------|---------------|------|
| | Head | Complement | Complement | Head |
| VP | → V | DP | DP | V |
| IP | → AUX | VP | VP | AUX |
| PP | → P | DP | DP | P |
| DP | → D | NP | NP | D |

Because every language is always in a state of change, and the order of head and complement in a particular phrasal category is one variable that may change over time, these correlations are not absolute. However, they are highly significant from a statistical point of view. It is certainly not mere coincidence that languages correlate in this way. The problem for linguistic theory is *why* this should be the case. Many linguists have approached this problem from different directions, and

Table 7.1 *Summary of Greenberg's Universals (from appendix 2 of Greenberg 1963)*

Greenberg's Universal	Parameter	correlation	
#1	Main clauses	V-O	O-V
#3,4	Adpositions	Prepositions	Postpositions
#2	Genitive (possessor) and head noun	N-G	G-N
#17	Head noun & adjective	N-Adj	Adj-N
#24	Relative clauses and head noun	N-RelCL	RelCL-N
#22	Comparatives	Adj-Mkr-Std	Std-Mkr-Adj
#16	Inflected auxiliaries	Aux-V V-Aux	
#9	Question particles	Sentence-initial	Sentence-final
#12	Question words	Sentence-initial or elsewhere	Sentence-initial
#27	Affixes	Prefixes	Suffixes

we will not try to summarize these here. Rather, we will simply present the findings of some very important foundational research, and then give some examples of languages that represent each of the major types.

The foundational work in syntactic typology was done by Joseph Greenberg in the early 1960s. Greenberg compared the syntactic characteristics of thirty languages and found several interesting correlations. In particular, he noticed that the languages in his sample tend to have a basic, or unmarked, syntactic structure, and that the order of certain elements in this basic structure correlate with the orders of other elements. Table 7.1 summarizes the correlations that Greenberg (1963) observed for VO and OV languages. These have come to be known as “Greenberg’s Universals,” since they were assumed to represent correlations that hold true universally, i.e., for all languages.

It is important to recognize that Greenberg simply observed certain correlations. He did not attempt to provide a reason for (i.e., to “motivate”) those correlations, or even to test them for statistical significance. In this sense, Greenberg did not attempt to *predict* constituent orders in as yet unstudied languages. Since 1963, much research has revealed problems with Greenberg’s original typology. Significant revisions, criticisms, and extensions of Greenberg’s work are found in Hawkins (1983), D. Payne (1985), Mithun (1987), Dryer (1988, 1992), and Hawkins (1994). In an important correction, Dryer (1988) shows that Greenberg’s Universal number 17 (the order of adjective and head noun) does not hold when a larger sample of languages is considered. Nevertheless, Greenberg’s work stimulated the field of typological linguistics and has continued to be very influential.

In the following sections we will explain and illustrate some of the correlations described in table 7.1, using examples from two typologically distinct languages – Japanese and Malagasy.