1 *The study of voice*

1.0 Preliminaries

In the history of grammar, the study of voice dates at least to the Sanskrit grammar of Pāṇini (*ca* 500 BC). In this work, entitled *Aṣṭā-dhyāyī* ('Eight-Chaptered'), are described the distinctions of inflectional paradigms and of meanings associated with the opposition of active and middle in the Sanskrit verb. Voice analysis is thus one of the most ancient topics in the tradition of descriptive grammar. Nevertheless, many recent writers seem to invoke the concept of voice in an intuitive or pretheoretical sense, rarely offering clearcut standards to determine whether specific behaviors are to be included in or excluded from the category.

Grammatical voice* refers to a category of the verb. Its status is thus comparable with that of other verbal categories such as tense, aspect, and mood/modality.

Grammatical voice is manifested in systems in which alternations in the shapes of verbs signal alternations in the configurations of nominal statuses with which verbs are in particular relationships. Voice differs in this respect from case*, a nominal category whereby the relationship of some particular nominal to some verb is signaled.

The present work is intended to broaden and lend clarity to the current understanding of grammatical voice and voice systems by advancing a typological scheme for their description; that is, by proposing grounds, behavioral and structural, for characterizing voice phenomena as conforming to types. The present chapter furnishes a descriptive overview of a variety of voice behaviors. The chapter presupposes some familiarity with voice accounts in current grammatical theory, although it does not address the content of recent proposals for the formal treatment of voice.[1]

Chapter 1's objective is to overview classes of natural language

1

2   *The study of voice*

behaviors which descriptive grammars often handle under the rubric of
voice. Section 1.1 concerns historical foundations for analyzing these
behaviors, i.e. traditional and posttraditional conceptions and defini-
tions of voice. Sections 1.2–1.4 survey three kinds of behaviors which
match traditional and posttraditional views of voice alike; these are
termed derived voice* systems, basic voice* systems, and pragmatic
voice* systems. Section 1.5 concerns some concepts of grammatical
organization foundational to the further study of voice.

   An outline of the contents of the complete work is deferred until the
close of the present chapter.

1.1 Voice in traditional grammar

Current treatments of voice reflect earlier traditions of grammatical
scholarship. For the purposes of the present discussion, one of the most
significant traditions is that of classical language studies. Works
representative of this tradition include descriptive and pedagogical
treatments of the structures of ancient European languages, chiefly
Greek and Latin.

   In the classical tradition, the structures of these languages are usually
analyzed in terms of two factors. The first is the language items them-
selves, which are organized according to lexical classes (parts of
speech). The second factor is the paradigmatic variation in the forms of
items according to certain functions called grammatical categories*. As
noted above in Section 1.0, voice is recognized, along with tense,
aspect, and mood/modality, as a grammatical category of the lexical
class of verbs. Since this view has its origins in traditional grammar, it is
fair to say that voice is traditionally regarded as a parameter of morpho-
logical variation in the verb.[2]

   Some terms for the category of voice reflect this. That is, traditional
terminology alludes to the different ways a verb might be, so to speak,
sounded (Latin *vox* 'voice'), i.e. to its repertoire of forms; or to vari-
ation in the verb's disposition (Greek *diathesis*) according to alterna-
tions in its relations with sentential arguments.[3]

   English grammarians usually treat the active/passive alternation
under the rubric of voice, although there are noteworthy exceptions.
Pre-eminent among them, the great Otto Jespersen speaks of the active
and passive "turn" in the English verb, refusing to identify this alterna-
tion with voice "as found, for instance, in Greek" (1965: 168). In

Jespersen's view, the latter term is more appropriate to the description of verbal alternations in classical Indo-European languages.

In part, Jespersen's may be a reaction to an inaccuracy perpetuated in many treatments of voice in traditional grammars. In classical literary languages such as Greek and Latin, voice alternations are not restricted to transitive verbs. However, there is a tradition of describing voice alternations in terms of relations within transitive structures.[4]

Moreover, many traditional grammars associate the functions of voices with alternations in viewpoint encoded in structures that are both formally as well as logically transitive. Transitivity* in the logical sense means that the situation denoted in a predication involves two participants; while formal transitivity (for purposes of the present discussion) can be regarded as meaning that the predication includes nominals in at least two semantic roles*, such as Agent and Patient (as these roles are often termed traditionally). According to this view of voice, a transitive situation can, in principle, be projected grammatically from either of two viewpoints, corresponding to two voices of the sentential verb.

One of these voices is said to encode the doing of an action (Michael 1970: 374–5). This voice is called active* because the action notionally devolves from the standpoint of the most dynamic, or active, party involved in the situation, typically the Agent. The second voice encodes action which notionally devolves from the standpoint of a nondynamic, typically static participant in the situation, such as the Patient of a transitive verb. This voice is called passive* because the verb is portrayed as "signifying the *state* of 'being acted upon' or 'suffering the effects of the action'" (Lyons 1968: 372; emphasis in original).

Traditional grammars also recognize a third or middle* voice category. Originally, the middle seems to have been conceived as a compromise category displaying characteristics of both the active and the passive. In a middle construction, the viewpoint is active in that the action notionally devolves from the standpoint of the most dynamic (or Agent-like) participant in the depicted situation. But the same participant has Patient-like characteristics as well, in that it sustains the action's principal effects.

A typical instance is the Classical Greek *loúomai khitôna* 'I wash (middle) the shirt (for myself),' i.e. 'I am washing my shirt' (Lyons 1968: 373). Here the middle is appropriate, since the Agent both performs and benefits from the action. According to one writer, "The Greek construction is called middle because it is semantically intermedi-

4 *The study of voice*

ate between the active and the passive: the subject does the action to or for him/herself . . ." (van Oosten 1977: 469, fn. 1)

In posttraditional linguistics, many treatments of voice dispense both with the notion of participant viewpoint (the idea of action devolving from some participant's standpoint) and with the concept of a middle voice, in the sense of a construction which displays both active-like and passive-like characteristics. Some recent accounts suggest a mapping between clause-level verb–nominal structure and logical-level predicate-argument structure* roughly as shown in Figure 1.1. This figure depicts a correspondence between a clause-level configuration, consisting of a verb or verbal element in a network of syntactic configurations with nominals, and a logical-level configuration, consisting of a predicate in a network of relations with arguments. These logical-level arguments are of two kinds: those essential to form a predication, or core arguments* (called "inherent arguments" by Marantz 1984); and other arguments that are optional (noncore). (In the figure, optional or noncore arguments correspond to numbered xs enclosed in parentheses.)

The number and type(s) of its core or essential argument(s) is an idiosyncratic property of an individual predicate and plays a role in determining the valence* of the corresponding (clause-level) verb or verbal element. A one-place (univalent) predicate, or predicate in an essential relation with exactly one argument, corresponds to an intransitive* verb; a two-place (bivalent) predicate corresponds to a transitive* verb; a three-place predicate to a ditransitive* verb; and so on. In Figure 1.1, following a convention introduced by Dixon 1979 and reaffirmed by Comrie 1978, core or essential clause-level nominals/logical-level arguments are indicated as follows: the symbol S corresponds to the core argument of a typical univalent predicate/intransitive verb, while A and P correspond to the two core arguments of a typical bivalent predicate/transitive verb.

Here we do not pause to discuss the assignment of semantic roles* to core arguments/nominals, although obviously, several of the symbols in Figure 1.1 do suggest specific semantic content (for instance, "A" is reminiscent of Agent and "P" of Patient). For present purposes, it is merely assumed that S is a univalent structure, A and P in a multivalent structure are each in a direct relation with the predicate/verb. A structural configuration which meets this description may be called normal.

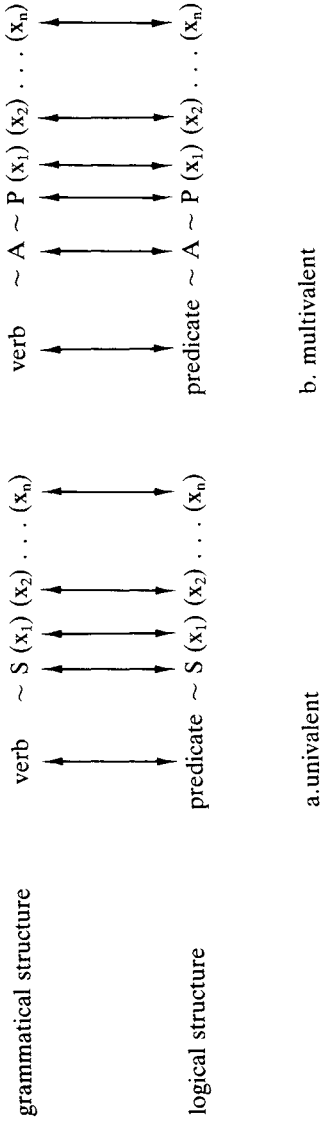Given this background of assumptions, a normal structural configura-

grammatical structure

verb $\sim$ S (x_1) (x_2) . . . (x_n) verb $\sim$ A $\sim$ P (x_1) (x_2) . . . (x_n)

logical structure

predicate $\sim$ S (x_1) (x_2) . . . (x_n) predicate $\sim$ A $\sim$ P (x_1) (x_2) . . . (x_n)

a. univalent   b. multivalent

Figure 1.1   Clause-level and logical-level structure correlations

6   *The study of voice*

tion, e.g. that of a clause, corresponds to one which is in an unmarked voice. Moreover, a marked voice results from any alternation in or deviation from the normal relations – depicted in Figure 1.1 – between a predicate/verb and its core arguments/nominals.

This conception of voice and voice alternations has many proponents. One of the best known is Fillmore 1968. In his view (and that of writers who share a common perspective), the function of voice marking, or overt verbally encoded manifestations of voice, is to signal the intactness or disruption of the basic relation(s) of a verb to its core nominal(s).[5]

This view of voice is consistent with the possibility of several distinct types of disruptions in verb–nominal relations, corresponding to several marked voice configurations. For instance, a core argument of an unmarked or basic configuration may be eliminated, resulting in its being suppressed in a corresponding marked voice construction. When this occurs, the marked voice may be said to encode or reflect omission of a core argument.

Examples (1)–(3) below illustrate this point. In Bengali (ex. 1), omission of the core argument, the S, of an intransitive (*se* in 1a) results in a marked construction (1b). Here the basic verb is nominalized and followed by a finite form of the verb *jaa-* 'go.' However, in (1b), this verb does not have the conventional lexical sense; rather, it serves as a grammatical marker, signaling omission of the basic S. The result is a structure in which the S is suppressed, or a subjectless clause. Moreover, syntactically speaking, clause (1b) has zero valence, or is said to be impersonal*.

Omission of either core argument, A or P, of a transitive may be signaled by distinct marked voices. (2) illustrates P-omission in the Mayan language Tzeltal (from Jacobsen 1985: 182).[6] In (2b), the omission of the P is signaled both by the addition of the verbal suffix *-awan* and by the loss of the P-agreement morpheme (the suffix *-on* in 2a).

A-omission is illustrated in the Spanish examples of (3). Example (3b) is marked relative to (3a), reflecting omission of the A (*ellos* 'they' in 3a). Moreover, while the argument corresponding to the P in the basic structure of the clause (*la ventana* 'the window') appears in (3b), the basic A is suppressed, or obligatorily unexpressed. In (3b), the element *se* marks this omission. (Taken up in Chapter 2 is the fact that this *se* also has other functions, particularly in marking the reflexive*.)

(1) a. Se okhaane bɔse
         he there   sits
         'He sits there'
     b. Okhaane bɔsaa   jaay
         there   sitting goes
         '(literally) It is sat there/It can be sat there'
(2) a. La   s-   tiʔ   -on   te   ȼ'iʔ e
         tense 3SG A bite 1SG P the dog demonstrative
         'The dog bit me'
     b. Tiʔ   -awan -∅   ȼ'iʔ e
         bite *-awan* 3SG A dog demonstrative
         'The dog was biting'
(3) a. Ellos rompieron la   ventana
         they   broke-3PL the window
         'They broke the window'
     b. Se rompió   la   ventana
         se   broke-3SG the   window
         'The window broke/was broken'

   Omission of a core argument is just one type of disruption in basic
verb–nominal relations. Another type is core argument rearrangement.
This occurs when a core argument's basic structure relation to the verb
is ceded, or reassigned to another nominal, usually entailing an alter-
ation of basic verbal valence* (defined as the number and types of
nominal positions with which a verb is lexically associated). The Ger-
man example (4b) is based on a univalent (intransitive) structure cor-
responding to (4a); while the English example (5b) is based on a
transitive structure corresponding to (5a).

(4) a. Die Kinder   schlafen
         the   children-NOM sleep
         'The children sleep'
     b. Es wird (von den Kindern)   geschlafen
         it   is   (by   the children-DAT) slept
         '(literally) It is slept by the children'
(5) a. Babe Ruth throws baseballs
     b. Baseballs are thrown (by Babe Ruth)

   First consider the alternation in (4a, b). The S in (4a), *die Kinder* 'the

8   *The study of voice*

children', may or may not appear in (4b). Either way, the verb in (4b) appears in a special, or marked, shape preceded by an auxiliary, and the S, if expressed, fails to agree with it. This is evidence that the relation of the S to the verb in (4a) is not preserved in (4b), or to put it another way, the corresponding nominal is no longer a core nominal. An additional piece of evidence for this is that, if expressed, the S appears in an oblique shape; the parenthesized form in (4b) is in a nondirect (non-nominative) case governed by a preposition (*von*). Moreover, it is in a marked position within the sentence, appearing clause-finally (compare the S's clause-initial positioning in ex. 4a).

In addition to all this, (4b) includes a clause-initial element *es*. It is what some writers term a dummy*, or a syntactic placeholder with neither thematic nor referential content (similar to *there* in the English *There were children sleeping*). Expressions such as (4b) are often termed impersonal passive*; they will be further considered below (Section 1.3; Section 2.3). For the present, it may be noted that (4b) is impersonal because, like the earlier (1b), it is nonvalent, having no nominal in a core relation to the verb.

Now consider the alternation in (5a, b). Example (5a) corresponds to a basic transitive configuration with a verb in relation with two core nominals, A and P. The corresponding marked construction (5b) has one core nominal corresponding to the P in (5a), but in (5b) this nominal has evidently assumed the syntactic relation basic to the A, or the relation borne by the A in (5a). Evidence for this is the fact that the (logical-level) P in (5b) both occupies the usual linear position of the A and also takes its case form, the nominative. For its part, the nominal corresponding to A is either suppressed in (5b) or, if expressed, occurs in a marked position – sentence-finally – and in an oblique shape (governed by the postposition *by*). The totality of evidence indicates that the A's basic relation to the verb is disrupted in (5b), being, in fact, reassigned to the basic P.

The sort of marked voice construction illustrated in (5b) is, of course, commonly called passive*. Notice that in (5b) the verb, which is basically transitive (*throws*), appears in a marked shape and is intransitive (*are thrown*). Examples (4b) and (5b) reaffirm the point made earlier that the basic valence is typically altered when a voice alternation disrupts the verb's basic relations with core nominals. In particular, one typically observes a reduction in transitivity*, here meaning the number of nominal positions associated with a lexical verb.[7]

To summarize the present discussion, one posttraditional view of voice is based on the assumption that in basic structural configurations, verbs occur in relations with core nominals that are normal, or unmarked. Nonbasic, marked structural configurations arise by alterations in these normal relations. The function of voice marking is to signal intactness or disruption of normal relations.[8]

Disruptions are of two major types: omissions (of core nominals) and rearrangements (of core nominal relations within clausal configurations). Figure 1.2 presents a scheme for the disruptions – or marked voice configurations – covered in the above discussion.[9]

The scheme summarized in Figure 1.2 presupposes an explicit theory of mapping between grammatical and logical levels of structural analysis; this has been provided for earlier in Figure 1.1.

In evaluating the above model of voice alternations, one issue to take into consideration is whether it accurately summarizes the range of types into which voice behaviors are organized cross-linguistically. A perusal of various language-particular accounts suggests that it may not. In fact, there seem to be entire classes of behaviors that conform either partially or not at all to the scheme just outlined, although they are widely regarded as voice behaviors. As it turns out, the above model is of questionable validity particularly for classical languages, those languages for whose description voice was originally posited as a grammatical category. There are even certain limitations in the above view of voice when it comes to languages to which it is relatively suited, as will be noted in the following section.

One recent writer comments on the current state of voice analysis as follows: "In traditional Indo-European grammar, the voices of a verb are morphologically distinct forms which indicate various relations between the subject and the verbal idea; however, the manner in which the term 'voice' should be applied outside of that language family is not always a straightforward, uncontroversial issue" (Ayres 1983: 20). In fact, the term voice is currently invoked in reference to systematic alternations in predicate–nominal relations of at least three distinguishable types.

In the first place, the term voice is perhaps most widely invoked today in reference to systems in which an alternation in verbal shapes signals alternate assignments of nominals to positions in structural configurations. This has been reviewed above. However, there is also a second, more traditional use of the term voice in reference to systems in which

	a. basic univalent configuration	b. basic multivalent configuration
basic configuration	$VI \sim S$ (exs. 1a, 4a)	$VT \sim A \sim P$ (exs. 2a, 3a, 5a)
omission of core argument	V_0 (ex. 1b)	$VI \sim S$ ($= A$) (ex. 2b) $VI \sim S$ ($= P$) (ex. 3b)
rearrangement of core argument(s)	V_0 (oblique S) (ex. 4b)	$VI \sim P$ (oblique A) (ex. 5b)

Key
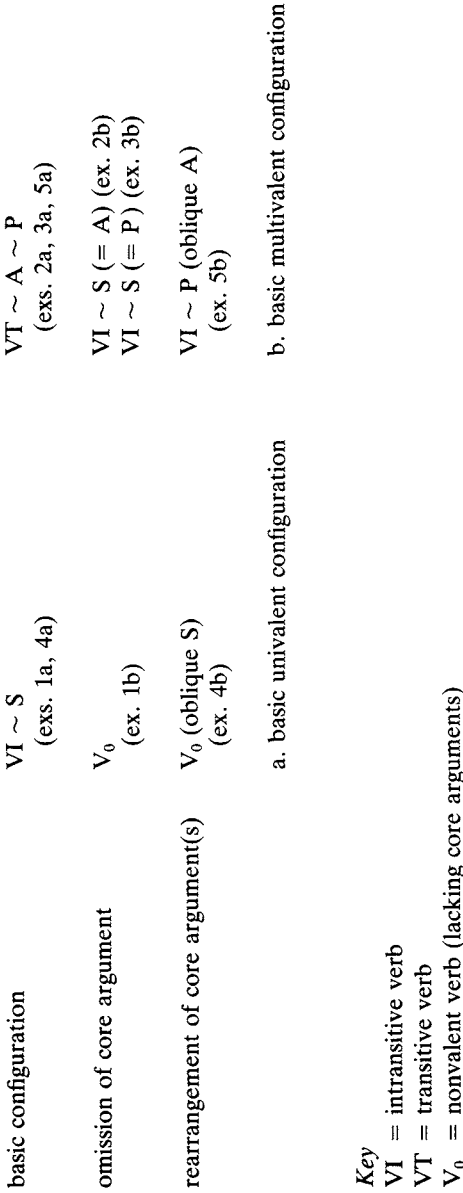VI = intransitive verb
VT = transitive verb
V_0 = nonvalent verb (lacking core arguments)

Figure 1.2   Unmarked and marked voice configurations