

Part I

INTRODUCTION

In dealing with problems which can be solved in more ways than one, the solutions themselves are of less interest than the reasons for making one choice rather than another.

Newman (1967: 192)



I Preamble

This study was intended, at the outset, to have two complementary aims. The first was to present a detailed synchronic description of Latin verb inflection, and to justify this description so far as it seems possible. This task will (it is hoped) be completed by a separate descriptive volume, but selected examples are foreshadowed, in a simplified way, throughout the chapters which follow. The second aim is to develop a suitable linguistic theory: a framework, that is to say, within which the description of inflectional systems may be attempted. This is mainly the task of the present initial phase of the investigation, though several problems of detail will be taken further as the descriptive phase proceeds. The motive for both phases as a whole is expressed succinctly by the quotation on the part title, from a brief article by a well known anthropological linguist. They are intended, it will be seen, to approach the problem from opposite directions: the first from the standpoint of a general morphological theory, the second through the close examination of specific morphological rules. But these are designed as complementary sides of the same investigation.

The scope of inquiry will be restricted, in the main, to inflectional morphology in its usual sense. One cannot, of course, discuss the inflections of Latin without referring also to the various syntactic and semantic categories which they represent. Nor can one avoid excursions into the theory of syntax on the one side and into phonological theory on the other. The work as a whole will, however, say little that is systematic or original on any of these bordering topics. In particular, two partial or virtual exclusions seem to be justified. First, so far as the facts of Latin are concerned, we shall accept without ado the scheme of Verbal categories (Tense, Aspect, Mood, etc.) which is normally adopted in philological discussions. The arguments for this scheme will be familiar to many readers; for the rest, a non-Latinist may be assured that there is no alternative scheme which would lead to a simplification of the theoretical problems. In the second place the theory developed is a

¹ See Appendix 1 for a tabular comparison between the Tense/Aspect scheme and the scheme of Tenses normally taught in the schoolroom.



Preamble

theory of inflectional morphology alone, and not what may be called an 'integrated' theory of morphology and syntax. This does not mean that an integrated theory is eschewed; indeed it is assumed to be necessary, as we shall make clear later in these introductory chapters. But the field of syntax is at the time of writing in such turmoil that the pursuit of an integrated theory would lead us into issues which are quite irrelevant to our descriptive problem. It therefore seems wise to restrict ourselves to morphology and to the link, via the word, between the two sides of grammatical description, and to show how our proposals might relate, perhaps, to at least some of the suggestions now current in the syntactic field. There seems no reason, at any rate, to believe that there are sound syntactic arguments for excluding the approach to morphology which we shall put forward.

With these provisos, therefore, a problem which is typical of those investigated in this study may be illustrated as follows. We have in Latin a word which may be transcribed, in terms of the phonological units of the language, as ferre, and we accept that it may be characterised, in terms of its syntax and semantics, as an Imperfective Infinitive translatable by English 'to bring' or 'to carry'. How, then, should a grammar relate the phonological form on the one hand to its grammatical form a specific lexical item in the Infinitive, etc. - on the other? A traditional answer, of course, might be in terms of exemplary paradigms. More recently, however, a common answer would be to divide the phonological form into various smaller parts, ascribing a complex fer to the 'bring' verb itself, the gemination of the r to the category 'Infinitive', and the final e, perhaps, to a further category 'Active'. In doing so, our aim would be to relate the fer complex in this particular form to the recurrent fer of fer-o:3 'I bring', fer-t 'He brings', etc., and the remainder of the word to the corresponding termination of ama:r-e 'to love', da-r-e 'to give', etc., and the partly different termination of fer-r-i: 'to be brought', ama:-r-i: 'to be loved' or da-r-i: 'to be given'.

The principles underlying such an analysis are no doubt familiar to

¹ See, in particular, the discussion of so-called 'syntactic features' etc. in ch. 7, p. 146 et passim.

Traditionally 'Present Infinitive'; for the term 'Imperfective' see Appendix 1,

already referred to.

³ Latin words will normally be cited in a reconstructed broad transcription, with hyphens marking relevant morphological boundaries. Orthographic forms, which will be used occasionally for explanatory or other purposes, are cited in inverted commas. For the transcription concerned see Appendix 2; it will be seen that I have used I.P.A. symbols with their normal phonetic values, to assist readers who may be ignorant of Latin phonetics.



1 Preamble

5

most readers of this volume. The particular interest of Latin resides, however, in the complexity and 'irregularity' of the data and in the nicety of the analytical judgments which a linguist is obliged to make. In the case of *ferre*, for instance, there is at least one alternative solution which ought to be compared with the one which we have just given: namely to treat the whole of the *re* termination as a 'marker' or overt manifestation of 'Infinitive' and to leave the 'Active' category (as elsewhere throughout the verbal paradigms) non-manifested or 'unmarked'. The *ri*: termination of *ferri*:, etc. would then be regarded as a sandhiform or coalescence of *re* followed by *i*:. Neither of these treatments can be rejected on obvious grounds of simplicity or elegance: what, then, are the criteria which might allow one to choose between them?

The answer to such a question commonly involves a variety of considerations. In part, naturally, it may depend on wider acquaintance with the structure of the language; in this case, one might particularly want to correlate the phonetic characteristics of Active and Passive Voice-forms with the more abstract syntactic or semantic status of the same opposition. It is not easy, however, to separate the empirical evidence adduced in such investigations from the quite different sort of argument which derives from theoretical and in some sense a priori commitments. In the case before us, a linguist's view of the syntactic status of Active and Passive might well be determined, in advance of any specific empirical consideration, by the limitations of his particular model or blueprint for human language. The ancient grammarians, for example, implied a model of language in which the oppositions between grammatical categories are of a uniformly 'equipollent' character; 3 in terms of this model, one must establish 'Active' as a category whose status is logically opposed to that of 'Passive' in precisely the same way, let us say, that '1st Person' is logically opposed to '2nd Person' and '3rd Person', or 'Past Tense' to 'Future', and so forth. Some modern writers, on the other hand, have experimented with a model in which all such oppositions are 'privative';4

¹ For the concept of sandhi see ch. 6, p. 71, and references cited at that point.

² The second analysis is, in fact, that suggested by Hill (1958: 472) and Householder (1957: 777); it also conforms to the history of the *re* and *i*: elements, insofar as this may be reconstructed (e.g. Palmer, 1954: 278-9). The first, on the other hand, was supported by S. C. Dik and H. Pinkster in correspondence on this chapter. We shall use this example for several illustrative purposes throughout this volume.

³ The terms 'equipollent' and 'privative' (see below) are borrowed from the 'Prague' school of phonology; cf. Trubetzkoy, 1939: 67 (=Trubetzkoy, 1949: 77) or excerpts in Vachek, 1960: s.v. 'opposition équipollente', 'opposition privative', etc.

⁴ The leading account of marked/unmarked oppositions in morphology is that of Jakobson (1932) for Russian (compare also Trubetzkoy, 1934: 5ff.). For an applica-



6 1 Preamble

in terms of this approach either the so-called 'Active' forms must be analysed as 'unmarked' with respect to the 'marked' category Passive (i.e. they may be described as lacking some characteristic which for the Passive is positively specified) or, conversely, the so-called 'Passives' must be analysed as unmarked in opposition to the Active. If one accepts this there is probably good evidence for saying that it is indeed the 'Active' category which is the syntactically 'unmarked' term in the opposition; one might then expect, if possible, that it should be 'unmarked' in the phonetic sense also.

It seems clear, from the foregoing discussion, that a question such as 'How should we analyse Latin ferre or ama:re?' cannot readily be divorced from wider questions concerning the validity of descriptive models and, indeed, the justification of as yet even vaguer analytical preferences. To evaluate such models, and where possible to clarify these preferences, must, accordingly, be a major part of our investigation. In the theoretical chapters which follow, the structure of the argument is essentially as follows. The balance of the present part (Part I) is intended as a brief metatheoretical introduction: an attempt to characterise our objectives more precisely, to explain our use of certain terms, and to make clear our position with regard to some of the major controversies affecting linguistics as a whole. In illustration, we will introduce a number of particular issues which will be taken up more seriously at various later stages. Part II then explores the nature of inflectional morphology in general, commenting especially on the various forms of description that have been advocated in the literature and on the arguments which might be advanced in favour of one form or another. This is the heart of our theoretical study, and the type of description which is eventually chosen supplies the setting for the two more technical parts which follow. Of these, Part III concentrates in detail on the format of an inflectional grammar: how do we give an appropriate and explicit expression to the sort of analysis which has been thought desirable? At the same time, it will continue to throw up a number of more general issues. Finally, Part IV will turn to the problem of choosing between alternative particular grammars - assuming already, that is to say, that the format in Part III has been accepted. The divisions between Parts II, III, and IV spring, in effect, from the characterisation of a linguistic theory which is summarised briefly in the following chapter (2.1-3).

tion to Latin Tense, Mood and Aspect see Bergsland, 1942 (a reference which S. C. Dik and H. Pinkster have kindly drawn to my attention).



2 The scope of a linguistic theory

So far as a study of this kind is concerned, the term 'linguistic theory' or 'theory of language' may be taken to refer to a system or assemblage of constructs with essentially the following characteristics.

2.1 Models of description

In the first place, a theory will be expected to specify a MODEL OF DESCRIPTION in the sense implied in the preceding paragraphs. This has been aptly described as 'a frame of reference within which an analyst approaches...a language and states the results of his investigation'; more precisely, however, such a model may be thought of as a specification of certain types of unit (for example, words or sentences) and of certain relations which will obtain between units of these various types in any relevant description. Rudimentary examples may be found already in the work of the ancient grammarians. Thus the following quotation from Priscian:

quemadmodum literae apte coeuntes faciunt syllabas et syllabae dictiones, sic et dictiones orationem²

expresses a model, or partial model, in which the primitive units designated by the Latin terms 'litera' (Letter), 'syllaba' (Syllable), 'dictio' (Word) and 'oratio' (Utterance or Sentence)³ enter into a relationship indicated here by the verbal forms 'coeuntes' ('coming together') and

¹ Hockett, 1954: 210 (=Joos, 1958: 386); Joos, 1958, will henceforth be referred to as *RiL*. For a survey of the diverse senses of 'model' see Chao, 1961.

² 'Just as letters, when they come together in an appropriate way, form syllables and syllables form words, so also words form sentences': ed. K. (Keil, 1855-70) III: 108. The relationship of letter to syllable and word to sentence is already implied by the definitions of συλλαβή and λόγος in Dionysius Thrax (ed. Uhlig, 1883: 16f., 22); for a more elaborate hierarchy, with a fifth term διάνοια between 'word' (λέξις) and 'sentence' (λόγος), cf. Stephanus (Scholia Vaticana on Dionysius Thrax), ed. Hilgard, 1901: 211.

³ 'Śyllaba' apart, the precise translation of all these terms might be disputed. For 'dictio' in relation to modern theories of the word see ch. 9, p. 160f. For the precise properties of the 'litera' see Priscian's own remarks (K. II: 6-9), and discussion of this and subsequent concepts of the letter in Abercrombie, 1949. The distinction between 'utterance' and 'sentence' is, as far as I know, quite modern.



8 2 The scope of a linguistic theory

'faciunt' ('form' or 'make'). This is, of course, a quite unsatisfactory frame of reference by modern standards: even where the concept of successive size-levels is preserved, it is customary to emphasise that the Syllable and the so-called 'Phonological Word' belong to a different series or hierarchy of size-units from the Sentence and the Word in a syntactic sense. 1 But the principle which it illustrates remains an abiding preoccupation of linguistic theory. The object, clearly, is to express a generalisation about the way in which languages (or at least a certain type of language) might be said to hang together.² The units and relations which are specified will accordingly concern precisely those properties which are believed to hold for more than one individual case; by the same token, naturally, a model may be amended or qualified whenever some feature leads to unacceptable descriptive consequences. In the case of the ancient model, the most interesting assertion concerns the status of the Word ('dictio' or λέξις) as the basic unit in the formation of sentences; to cite another passage from Priscian, 'dictio est pars minima orationis constructae'.3 This part of the hypothesis is indeed a matter of continuing controversy.4

2.2 Generative grammars

A model in the sense described is undoubtedly a large part of what linguists have understood by a linguistic theory. There are, however, at least two further, more controversial assumptions which it will also be necessary to take for granted in the chapters which follow. The most important of these concerns the concept of a GENERATIVE GRAMMAR or, more generally, of a grammar as the application to a language of some kind of formalised RULE-SYSTEM. This term 'generative' is commonly used in the literature in both a narrow and a somewhat wider sense.⁵

- See, for example, Bendor-Samuel, 1966: 30; Pike, 1967: particularly 569ff. For hierarchical models in syntax see further Halliday, 1961; for phonology, e.g., Shorto: 1960: 544. For phonological and grammatical criteria for the word see ch. 6, pp. 96ff.
- pp. 96ff.
 ² Cf. the metaphor 'how language works', e.g. in Halliday, 1961: 242. For the qualification in parenthesis see ch. 3, pp. 32f.
- ³ Freely, 'the word is the minimal unit in sentence-structure': K. II: 53. For the earliest Greek original see Dionysius Thrax (ed. Uhlig, 1883: 22): Λέξις ἐστὶ μέρος ἐλάχιστον τοῦ κατὰ σύνταξιν λόγου.
- 4 See ch. 4, p. 44 et passim.
- ⁵ The original sense is the narrow and precise one, as applied to a grammar as a whole. This concept was first introduced seriously into linguistics by Chomsky (1957: 13ff. for 'generate'; 1961a; 1961b; etc.), though brief suggestions may be traced back at least to Harris, 1951a: 372f. (and subsequently Harris, 1954a: 260);



2.2 Generative grammars

I. In the narrow sense it refers to any set of rules - or formalised statements about a language - which may be interpreted as defining, specifying or 'generating' some particular set of formal objects. A complete grammar has thus been seen, in this light, as an assembly of rules which will define the set of all 'possible sentences' in the language concerned. Alternatively, an individual section of a grammar might be 'generative' in the sense that it defines some set of objects of a more elementary kind; for example, we shall refer in Part III to the feasibility of rules which would generate the 'possible forms of a phonological word'. As applied to a grammar as a whole this is not, of course, incompatible with the aims which many less technically minded grammarians have set themselves in the past. For example, a grammar of Latin (for a reader who did not know the language) could be expected to indicate that 'one can say' puella moritur 'The girl is dying'2 but 'one cannot say', e.g., puellae mortua est (compare a putative English sentence 'The girls is dead'); the former, that is to say, is among the set of sentences which should be generated by the grammar, but the latter is not. The crucial rule, of course, is the one which states that a Plural Subject must govern a verb which is itself in the Plural. Similarly, a grammar can be said to exclude, as 'ungrammatical', a putative sentence such as puella morita est; i.e. there should be rules which specify that the relevant Participial form is mortua, not morita. What is new, however, in the modern concept of generative grammar is the continued search for mathematically precise and exhaustive presentation.

II. The second and wider sense refers simply to the commitment to precision which we have just referred to. Most grammars of the traditional sort are strictly neither wholly explicit nor wholly exhaustive; they can be used and understood only with the help of the reader's imagination, perhaps with some prior knowledge of the language or of

Harris was, presumably, Chomsky's supervisor at the time. It is also normal in textbooks of Chomsky's school (Bach, 1964: 13ff.; Koutsoudas, 1966: 1f.; Ruwet, 1968: 32f.). For sense II (below) the term 'generative' is not very well motivated, but it is easy to see how the usage has arisen; see especially passages such as Chomsky, 1965: 4, 8–9. The two different senses are clearly and usefully distinguished by Lyons in all his recent writings (1968: 155ff.; 1970a: 43f.; 1970b: 24). Finally, 'generative' is also used, often in a confusing way (cf., e.g. ch. 10, p. 199 et passim for 'generative phonology') to refer to anything which emanates from Chomsky's school.

¹ Ch. 10, p. 222 et passim in 10.2.

_

² For readers unfamiliar with Latin this example may be glossed as follows: 'Girl/ Nominative-Singular Die/Imperfective-Present-Indicative-3rd-Singular'; in the two examples which follow, *puellae* is the Nominative plural corresponding to *puella*, *est* is a 3rd Person singular form of the Verb 'be', and *mortua* is a Feminine singular Participle corresponding to *moritur*.



10 The scope of a linguistic theory

some other language which is similar, perhaps with some ad hoc exegesis, and so on. Often, of course, there is no reason why they should be otherwise; in many contexts, it would be foolish and tedious to spell out matters in a mechanically literal way. A generative grammar, however, should place no reliance on vague external aids of this kind. All that one should need, in order to understand it, is a knowledge of its formal structure and of the abstract principles on which its rules are constructed.1

This is not the place to justify or argue the merits of the generative approach. It is enough to acknowledge that though it is now widely accepted it is still, nevertheless, controversial.2 In the present study we are concerned, of course, with only one part or area within a generative grammar of Latin. But within this area, we shall require a linguistic theory to provide not only a model of description in the sense of the preceding subsection, but also a specification both of the precise format of grammatical rules and of the precise way in which such rules are to be interpreted. For any section of a grammar, this may be referred to as a specification of the RULE-SYSTEM which is considered appropriate. For illustration, it may help if we examine an earlier (though admittedly inadequate) account of essentially the same data. According to Books IXx of Priscian's Institutiones Grammaticae,3 the various forms which make up each verbal paradigm are derived, directly or indirectly, from a single 'leading form' for each lexical element. This 'leading form' is itself one particular member of the paradigm concerned. Thus for the 'Imperfect Indicative'4 Priscian formulates the rule:

praeteritum imperfectum...a praesenti fieri sic: in prima quidem et in secunda coniugatione et quarta in 'eo' desinente a secunda persona ablata s

- The incomplete and inexplicit character of traditional grammars is stressed, for
- example, by Chomsky, 1964: 937; see also Ruwet, 1968: 33ff.
 ² For the earliest critiques of Chomsky's approach see Haas, 1958 (and more sharply Haas, 1966: 119f.), Reichling, 1961 and Uhlenbeck, 1963 (more recently, Uhlenbeck, 1967), Hill, 1961 (but see Chomsky, 1961b: 227ff.), Dixon, 1963 (but see, for example, Lyons, 1963b and Matthews, 1965a). In most of these works other aspects of Chomsky's teaching are also under consideration; for a further reply see Chomsky, 1966. A recent critique which bears directly on the notion of a 'generative grammar is that of Hockett (1968); much of Weinreich et al. 1968, is also pertinent. Finally, it is as well to admit that I myself have argued against Chomsky's theories in one or two places (Matthews, 1967c: init; briefly in Matthews, 1968a; more sharply again in Matthews, 1970a); however the grounds of disagreement are largely irrelevant to the present study.
- ³ K. II: 452ff. The reasons for regarding this approach as unsatisfactory are discussed below, pp. 27ff; for the moment it is introduced merely as a convenient and familiar illustration.
- 4 = Imperfective Past Indicative; see Appendix 1.



2.2 Generative grammars

finali et addita 'bam': 'amas amabam', 'doces docebam', 'is ibam'; in tertiae vero omnibus verbis et quartae in 'io' desinentibus prima persona mutat o in e productam et assumit 'bam': 'lego legebam', 'facio faciebam', 'venio veniebam'.

Such statements could undoubtedly qualify as grammatical rules, in the special sense in which we shall use the term. That is, they could (though they do not in Priscian's case) form part of a total generative description of Latin. The relevant theory, or partial theory, would then have to supply the following information concerning the type of rule which they represent.

First, it would have to indicate their place and function within the overall scheme of a grammar; how do they connect, for example, with the basic description of sentence-structure or with concordial rules of the type applicable to puella mortua est? It is clear, of course, that the rules of syntax handle the external relations of one word to another within the sentence, whereas these further, 'morphological', rules are concerned with their internal relations within the paradigm.² In a rather different sense, however, such statements might be said to supply the answers to a certain kind of question: namely, that exemplified by the specimens 'What is the 1st Person Singular Imperfect Indicative of AMO "love"?' (Answer, ama:bam), 'What is the 3rd Person Singular Perfect Indicative of ROGO "ask"?' (Answer, roga:wit), and so forth.³ Given, that is to say,

- 'The past Imperfect is formed from the Present in the following way: for [verbs of] the first and second conjugation and [verbs of] the fourth conjugation ending in eo; the final s is deleted from [the form of] the 2nd Person and bam is added: [thus] ama:s \rightarrow ama:bam, doke:s \rightarrow doke:bam, i:s \rightarrow i:bam; for all verbs of the third [conjugation], however, and for [those of] the fourth which end in io; the o: of the 1st Person is changed to a long e: and bam is added to this: [thus] lego: \rightarrow lege:bam, fakio: \rightarrow fakie:bam, wenio: \rightarrow wenie:bam'. K. II: 457-8.
- This is perhaps the most usual account of the morphology/syntax division in the classical and similar European languages; cf., e.g., the straightforward note by Belič (1949) and the more elaborated definition by Kurylowicz (1949b: 286) in response to the questionnaire of the Sixth International Congress. The answers to this part of the questionnaire (Lejeune, 1949: 261–302; Report by Trnka, 1949; interventions, 473–96) are still worth studying as a whole. As other contributions made very clear, the problem with this definition is that it hinges on the often problematic concept of the 'word' (see 6.4.3). There have therefore been various attempts to generalise some vaguely related distinction: notably in terms of a system/text dichotomy (e.g. Togeby, 1949b for the same congress) or in terms of a strict schema of 'vertical'/'horizontal' relations (most recently by Pittman, 1959; cf. paradigmatic/syntagmatic in ch. 7, p. 119). In this way a distinction can be applied to all languages, but the TERMS 'morphology' versus 'syntax' are no longer appropriate (cf. Martinet's contribution, 1949c: 295). Our own resolution of this dilemma is to sacrifice universality (7.5).
- ³ I shall use orthographic forms in small capitals whenever I want to refer to Latin verbs *qua* lexical items. Thus AMO symbolises the item translated 'to love' in the dictionary or lexicon, whereas *amo*: represents the particular member of its paradigm translated 'I love'. Here we anticipate a theoretical distinction which is not clear in