The change from SOV to SVO in Ancient Greek

Ann Taylor
University of Pennsylvania

Abstract

Although the order of major sentence constituents at all stages of Ancient Greek is quite free, the distribution of clause types is not random over time, but changes from predominantly verb-final to predominantly verb-medial, suggesting a change in progress. Using the paradigm of Kroch (1989), in which it is assumed that syntactic change involves competition between grammatical systems, two models are constructed, one for the verb-final grammar of the Homeric period (pre-800 B.C.) and one for the verb-medial grammar of the Hellenistic Koine (c. 100 A.D.). The language of an intermediate stage (Herodotus, c. 450 B.C.) is shown to pattern in part like Homer and in part like the Koine. More strikingly, the ratio of verb-final to verb-medial structure that best fits the distribution of clause types found in Herodotus is extremely close to an independent measure of this ratio obtained from the distribution of weak pronouns and clitics.

The basic order of a verb and its arguments in all stages of Ancient Greek (AG) has been a subject of dispute for some time. Whereas most scholars agree that the subject is initial, both OV and VO order have been argued for with a slight bias in favor of OV (VO: Short, 1890; Kieckers, 1911; Friedrich 1975; OV: Ebeling, 1902; Fischer, 1924; Frisk, 1932; Dover, 1968). Most of these studies were based on statistical counts, though Friedrich also used type logical arguments. The difficulty of coming to any uniform conclusion based on simply counting is illustrated by the conflicting results that have been reported. Fischer concluded that Homeric Greek (HG) is SOV based on Th Iliad (5:512–909), whereas Friedrich found more or less equal numbers of SVO, SOV, and OVS in Th Iliad (5:1–296), but concluded that HG is SVC. Dover found Herodotus to be SOV in one passage (8:1–48), but SVO in another (3:61–78). An exhaustive study of the first book of Herodotus (Dunn 1988) found that both the subject and indirect object favor a position preceding the verb (approximately 70% and 65% of the time, respectively) whereas the direct object precedes and follows the verb about equally (pre

I thank Ann Bates, Lynn Guindon, Tony Kroch, Susan Patzuck, Owen Rambow, Don Ringe, and Beatrice Santorini for discussion and comments. Part of the work presented in this article was supported by a Social Sciences and Humanities Research Council of Canada Dissertation Fellowship.
cedes: 49%; follows: 51%). The most recent work on the subject (Cervin, 1990) concluded that the VP in Classical Greek (CG) is unordered.

In this article, I claim that the distribution of clause types across the AG period makes sense if it is assumed that the data reflect an ongoing change from verb-final to verb-medial structure beginning at or before Homer and nearing completion 800 years later with the Koïnè of the Hellenistic period. I follow Kroch (1989) in assuming that syntactic change may involve competition between grammatical systems. Under this hypothesis, incremental linguistic change is accomplished by the gradual replacement of one grammatical system (or subsystem) by another. While the change is in progress, speakers can make use of forms from either grammar, but as time passes, a higher and higher percentage of the output is derived from the innovative grammar at the expense of the one it is replacing. Finally, the paucity of evidence for the old grammar triggers a reanalysis in which all structures are generated by the new grammar and the change is complete.

Ancient Greek is a good testing ground for this hypothesis because the available data cover very nearly the entire span of the change from verb-final to verb-medial structure. Thus, the earliest extensive Greek texts, the Homeric poems, can be shown to be almost categorically verb-final, whereas, by the Koïnè period about 800 years later, the change is virtually complete, and the grammar is essentially verb-medial. On the basis of these two data sets, therefore, I construct a simple model of the relevant aspects of the verb-final grammar that predates the change and the verb-medial grammar that takes over. This information is then used to test the data from Herodotus, which falls between Homer and the Koïnè in time. As expected under the hypothesis that grammatical systems gradually replace one another through competition, a fraction of the Herodotean data patterns like Homer, whereas the rest patterns like the Koïnè. More strikingly, the ratio of verb-final to verb-medial structure that best fits the distribution of clause types found in Herodotus is extremely close to an independent measure of this ratio obtained from the distribution of weak pronouns and clitics. Thus, evidence from two independent types of data confirms the hypothesis that Herodotus made use of forms from two different grammars, one verb-final and one verb-medial.

First, I introduce the necessary theoretical apparatus. Then, I discuss the problem of determining underlying structure in a language that exhibits every possible surface order of a verb and its arguments. It is shown that, although most clauses are structurally ambiguous, an estimate of the ratio of verb-final to verb-medial structure can be obtained from the kind of diagnostic tests for verb position commonly used in Germanic syntax. In the third section, I discuss the model of grammar and usage adopted and outline the two grammars that constitute the beginning and end point of the change. Finally, I demonstrate that the distribution of clause types found in Herodotus indeed supports the hypothesis that syntactic change involves phrase structure competition.

SOV TO SVO IN ANCIENT GREEK

THEORETICAL ASSUMPTIONS

In what follows, I rely on the framework of Government and Binding (GB) as laid out in Chomsky (1981, 1986) and the wide range of work done by others within this theory.

X-bar theory

I assume, following Chomsky (1986), that all categories—both lexical (noun (N), verb (V), adjective (A), and preposition (P)) and nonlexical (complementizer (C) and inflection (I))—project in the same way: that is, according to the schema in (1). (Order is not significant and XP = X'.

\[(1) \text{a. } XP \rightarrow XP \ X'. \]
\[(1) \text{b. } X' \rightarrow X XP \]

This gives phrasal structures of the shape illustrated in (2), where the XP dominated by XP is the specifier and the XP dominated by X' is the complement.

\[(2) \]
\[\text{XP} \]
\[\text{XP} \ X' \]
\[X \ XP \]

The presence or absence of specifiers and complements for any particular category is determined by other aspects of the grammar (theta-theory, case theory, and the Extended Projection Principle). In addition, the X-bar constraints hold only at D-structure and may be altered by adjunction at other stages in the derivation.

In this system, IP (= S) is a projection of l or INFL (= AUX), and CP (= S') is a projection of C (= COMP). The structures of these two phrases, IP and CP, are illustrated in (3).

\[(3) \]
\[\text{CP} \]
\[\text{specCP} \ C' \]
\[C \ IP \]
\[\text{IP} \]
\[\text{specIP} \ I \ VP \]

C' takes an IP as its complement, and IP in turn takes a VP. I assume that the subject is base-generated in the specifier position of VP—a position from which it may move to the specifier of IP (Fukui & Speas, 1986; Kitagawa, 1986; Sportiche, 1988).

The head of IP is the site of inflectional elements, including tense, agreement, and modals. Head-complement order for each category is a language-
specific matter and is not necessarily uniform. For example, German is generally considered to have a head-initial CP, but a head-final IP, giving an underlying structure, as in (4).¹

\[
\begin{align*}
\text{CP} & \\
\text{specCP} & \text{C'} & \text{IP} \\
\text{C} & \text{IP} & \text{VP} & \text{I}
\end{align*}
\]

The theory of movement

According to Chomsky (1986), there are two types of movement: substitution and adjunction. Substitution takes two forms: (1) movement of XP to a specifier position (as in topicalization), and (2) movement of an X category to a head position (as in movement of V to I). X’ categories are never affected by substitution.

Adjunction takes a maximal projection XP or a head X° and Chomsky-adjoins it to another category of the same type (XP to XP, which may never be an argument [Chomsky, 1986], head to head), creating structures like those in (5). As with substitution, X’ level categories do not participate in adjunction, and it is generally assumed that they may not be adjoined to.

\[
\begin{align*}
\text{XP}_2 & \quad \text{XP}_1 \\
\text{specXP} & \quad \text{specXP} \\
\text{X} & \quad \text{X} \\
\text{XP}_2 & \quad \text{XP}_2 \\
\text{I} & \quad \text{I}
\end{align*}
\]

Verb movement and the position of INFL

The head of IP, INFL (or I) is the locus of such information as tense, aspect, agreement, and mood. This information, if present, must merge with the verb, a process that takes one of two forms: either the inflections lower to the verb, as in English, or the verb raises to INFL, as in French and many of the Germanic languages. When IP is head-final, the movement of the verb from its base-generated position into INFL is string-vacuous, and thus it is difficult to determine whether movement takes place (see (6)).

When IP is head-initial, as in French, a comparison of the position of such diagnostic elements as the negative pas with respect to tensed and untensed verbs indicates that the tensed verb, whether main or auxiliary, is in a position different from that of the untensed verb. Thus, in (7a), pas follows the main verb, which is tensed, whereas, in (7b), it follows the tensed auxiliary rather than the untensed main verb.

\[
\begin{align*}
\text{IP} & \\
\text{specIP} & \text{I'} & \text{VP} \\
\text{specVP} & V' & \text{X} & \text{VP} \\
\text{I} & \text{I} & \text{I} & \text{I}
\end{align*}
\]

(7) a. Je ne mange pas
   I ne eat not
   I ne have not eaten

   b. Je n'ai pas mangé

In a language such as Ancient Greek, which does not make much use of auxiliary verbs, it is much more difficult to determine either the position of INFL or whether the tensed verb raises. Since the question of whether the verb occupies its base-generated position in VP or has moved to INFL is immaterial to the issue at hand, I simply use the term "verb-medial" to indicate a clause that may be interpreted in structural terms as occupying the head of either a head-initial VP or a head-initial IP, as in (8). "Verb-final" covers either of the structures in (9). For my purposes, it is sufficient that there is some grammatical difference between these two structures, whatever its exact nature may be.

Verb-medial?

(8) Verbm-едial? 

\[
\begin{align*}
\text{IP} & \\
\text{specIP} & \text{I'} & \text{VP} \\
\text{V'} & \text{V'} & \text{V'} & \text{V'} \\
\text{X} & \text{X} & \text{X} & \text{X}
\end{align*}
\]
This study is based on works by three Ancient Greek authors widely spaced in time: The Iliad of Homer, the Histories of Herodotus, and the Acts of the Apostles by Luke. The Iliad is the earliest surviving extensive Greek text; originally an oral epic poem, it is generally thought to have been committed to writing for the first time around the 8th century B.C. Although it is clear that at least some aspects of the language of these poems are considerably older than the 8th century, that does not pose a problem in this particular case because Homer is the starting point of the study; that is, it is not necessary to place the poem relative to an earlier stage, but only relative to a later one. The bulk of the sample is taken from Books 1, 3, and 8, but data on the ditransitive verb, which is much rarer than the simple transitive type, are drawn from all 24 books.

Herodotus’s Histories are the earliest surviving Greek prose texts (Ionic dialect). Herodotus was born around 484 B.C., and thus the Histories presumably date from the middle or latter part of the 5th century. Herodotus is generally considered the most “natural” (i.e., close to speech) of the early prose writers (Dover, 1968, p. 10). The sample is taken from the beginning of Book 3.

The Koine Greek of the latest period was used as a lingua franca over a large part of the Mediterranean world during the Hellenistic and Roman periods. Many who used it were not native Greek speakers, and since it is unclear to me how much effect this would have on the data, I have tried to avoid the problem by taking the sample from the Acts of the Apostles, written in the last half of the 1st century A.D. This work is traditionally ascribed to Luke, a Greek doctor (who also is supposed to have written the Gospel of Luke). The sample is from the first 18 chapters.

BASIC WORD ORDER
The data
Word order freedom in all stages of AG is fairly extreme. All but the rarest orders of major constituents are represented in even a small sample from the texts of any given period. Examples of clauses with two arguments from each of the three authors included in the corpus are given in (10) through (18). Because AG allows empty subjects, these clauses may contain either a subject and a complement or two complements. Subjects are symbolized by S, and complements by X or Y. The (a) examples are from Homer, the (b) examples from Herodotus, and the (c) examples from Luke.
Achāiatoi / sumpantas
Achāiatoi-G all-ACC
‘indeed sometime longing for Achilles will come on all the sons of the Achāiatoi’ (II.1.240)

b. hoi Aiguptoi men autika apo paidiōn
because Egyptians-NOM PTC straightaway from childhood-G
arksenomai kourotistai tas kephalas
beginning-NOM shave the-ACC heads-ACC
‘because the Egyptians beginning straightway from childhood shave their
heads’ (Hdt.3.12.3)

c. ho theos tòn patérōn himon égeires Iason
the-NOM god-NOM the-fathers-G our-G raised Jesus-ACC
‘the God of our fathers raised Jesus’ (Acts.5.30)

(14) X v S
a. tounektō ar' alge' edoken hekēbolos
because of this PTC pain-ACC gave far-shooter-NOM
‘because of this the far-shooter gave pain’ (II.1.96)

b. to Persoi nomima epistēmati kai Aiguptoi
the-ACC Persian-G laws-ACC know especially Egyptians-NOM
‘the Egyptians especially know the Persian laws’ (Hdt.3.2.2)

c. toutan ton Iason anestēsen ho theos
this-ACC Jesus-ACC resurrected the-NOM god-NOM
‘God resurrected this Jesus’ (Acts.2.32)

(15) X v Y
a. apha gerontos apósmen agrion andra
so that old-man-G we-might-push-away wild-ACC man-ACC
‘so that we might push away wild-ACC the old man’ (II.8.96)

b. homōn tois ge paiōn autōn apodidousi tēn
nevertheless the-D PTC sons-D their-G they give back the-ACC
arkhēn
rule-ACC
‘nevertheless they gave back the rule to their sons’ (Hdt.3.15.2)

c. hina dedomenous autous agaγεῖ ἐπὶ τοὺς
so that being-bound-ACC them-ACC he might lead to the-ACC
arkhierēs
chieft-priests-ACC
‘so that he might lead them to the chief priests bound’ (Acts.9.21)

(16) v X
a. gnōi de kai Atreides eurō kreiōn
may recognize PTC also son-of-Atreus-NOM wide-NOM ruling-NOM
Agamenon / hein atēn
Agamenonom-NOM his-ACC madness-ACC
‘Atreus’ son wide ruling Agamenomon also may recognize his madness
(II.1.411)

b. ekhei de ho moskhos houzas ho Apis
has PTC the-NOM calf-NOM this-NOM the-NOM Apsis-NOM
kaleōmenos sēmeia
being-called-NOM markings-ACC the-following-ACC
‘this calf which is called Apsis has the following markings’ (Hdt.3.28.3)

(17) v X
a. all' ei me dōtousi gerai megarkumō
but if PTC would give prize-ACC great-hearted-NOM
Achāiatoi
Achāiatoi-NOM
‘but if the great-hearted Achāiatoi would give me a prize’ (II.1.135)

b. epempsan gar dē penteiakosias mēnas argyriōi hoi
sent PTC PTC 500-ACC minae-ACC silver-G the-NOM
Kurēnaioi
Kurēnaioi-NOM
‘the Kurēnaioi sent 500 minae of silver’ (Hdt.3.13.4)

c. hōs de ekousan touts logous tous houtos ho te
when PTC heard the-ACC words-ACC these-ACC the-NOM and
pragmatōs tou hierou kai hoi arkhierēs
captain-NOM the-G temple-G and the-NOM chief-priests-ACC
‘when the captain of the temple and the chief priests heard these words’
(Acts.5.24)

(18) v X
a. herdon d' Apollonī telēssas hekatombas
they sacrificed PTC Apollo-D perfect-ACC hecatombs-ACC
taurōn éa' aigōn para thin' halos atragoros
bulls-G and goats-G beside shore-ACC sea-G barren-G
‘they sacrificed perfect hecatombs of bulls and goats to Apollo beside the
shore of the barren sea’ (II.1.316)

b. apekrinē tou stratou hōs pente muryias
he split off the-G army-G about five-ACC thousands-ACC
‘he split off about 50,000 men from his army’ (Hdt.3.25.3)

c. anh' hōn ouk edoken tēn doksan τοίς theōi
because not he gave the-ACC glory-ACC the-D god-D
‘because he did not give the glory to God’ (Acts.12.23)

The data used in this study include all the tensed clauses with nonpronoun-
inal arguments from the portions of text mentioned earlier, with two excep-
tions. First, I have left out clauses with the verb eimi 'to be' or any of the
other verbs that can be used with this meaning (eignomai 'to come into
being/to be', ἐπιθύμω 'to begin/to be in existence', etc.) because these often act
differently from simple transitive clauses. Second, I have avoided clauses with split arguments like those
illustrated in (19).

(19) a. ἀποτελεῖτο Ἀκαθάρια Τρωών εὐκερτείς eu naloimenos
when Achāiatoi-NOM Trojans-G sacked well-built-ACC
pietolethron
city-ACC
‘when the Achāiatoi sacked the well-built city of the Trojans’ (II.1.163)
<table>
<thead>
<tr>
<th>Type</th>
<th>Homer</th>
<th>Herodotus</th>
<th>Luke</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verb-final</td>
<td>0.38 (41)</td>
<td>0.25 (34)</td>
<td>0.07 (7)</td>
</tr>
<tr>
<td>X Y v</td>
<td>0.06 (7)</td>
<td>0.02 (2)</td>
<td>0.01 (1)</td>
</tr>
<tr>
<td>Total</td>
<td>0.44 (48)</td>
<td>0.27 (36)</td>
<td>0.08 (8)</td>
</tr>
<tr>
<td>Verb-medial</td>
<td>0.28 (31)</td>
<td>0.44 (59)</td>
<td>0.56 (57)</td>
</tr>
<tr>
<td>S v X</td>
<td>0.12 (13)</td>
<td>0.12 (16)</td>
<td>0.03 (5)</td>
</tr>
<tr>
<td>X v Y</td>
<td>0.04 (4)</td>
<td>0.01 (1)</td>
<td>0.01 (1)</td>
</tr>
<tr>
<td>Total</td>
<td>0.44 (48)</td>
<td>0.57 (76)</td>
<td>0.62 (53)</td>
</tr>
<tr>
<td>Verb-initial</td>
<td>0.07 (8)</td>
<td>0.15 (20)</td>
<td>0.25 (25)</td>
</tr>
<tr>
<td>v S X</td>
<td>0.05 (5)</td>
<td>0.02 (2)</td>
<td>0.06 (6)</td>
</tr>
<tr>
<td>Total</td>
<td>0.12 (13)</td>
<td>0.17 (22)</td>
<td>0.31 (31)</td>
</tr>
</tbody>
</table>

N: 109 134 102

b. autika hoi Aiguptii ailema epheoreon
at once the-NOM Egyptians-NOM clothes-ACC put on

the-ACC most-beautiful-ACC

'at once the Egyptians put on their best clothes' (Hdt. 3.27.1)

This type of construction declines in frequency from Homer to Herodotus and is virtually nonexistent in Luke. It is unclear what the proper analysis of these clauses is, and as their inclusion in the model complicates matters considerably, I have left them out for the present.

The distribution of clause types

Although these word-order patterns occur throughout AG, the frequency with which they appear in the different periods is not the same, as can be seen in Table 1. This table gives the distribution of the clause types given in (10) through (19). It includes tokens with full NP (not pronominal or PP) arguments only. Note also that the order of the arguments on either side of the verb is not taken into account. Thus, S X v includes both S X v and X S v, and v S X both v S X and v X S. As this table shows, the distribution of clause types changes over time with the percentage of clauses with medial and initial verbs rising at the expense of those with final verbs. Thus, although the same range of clause types appears at each stage, their distribution is changing, reflecting a change in the underlying structure of the language.

In the next section, I discuss structural evidence that indicates both that a change from verb-final to verb-medial structure takes place during the AG.

The extreme freedom of AG word order, it is necessary to address the of nonconfigurality. A number of proposals have been put forward entert years that in one way or another attempt to account for free word by appealing to "flat" or "nonconfigurational" phrase structure (Hale, 1982, 1983; Lapointe, 1981; Pullum, 1982; Stucky, 1981; Zwicky, 1983). The strongest form of this proposal (Hale, 1981, 1982, 1983) would all languages into two types — configurational, those with hierarchica structure, and nonconfigurational, those with flat structures — by means of configurality parameter. The alternative to this approach is to derive possible word orders from a fixed order base by movement, a process referred to as "scrambling" (den Besten & Weibelhuth, 1990; Saito, 1989; Thibaut, 1989).

From a conceptual terms, the scrambling approach has the advantage of unifi-city. It assumes that all languages share a common, configurational base, that the assignment of case and theta-marking, which determines interpretation of sentences, is also universal. All deviations from the base are generated by the same mechanism — movement rules — whether they in a fairly restricted way, as in English, or much more freely, as in AG. Movement rules are necessary independently of the free word-order phe-nomenon (to derive wh-questions, etc.), the adoption of scrambling does not any fundamental change in core areas of the grammar. By contrast, the assumption of nonconfigurational structure simplifies consid-erably the derivation of alternate word orders (especially when these are of the same paradigm), it necessarily complicates interpretation, because some other argument must be assigned case and theta-roles correctly. Thus, the assumption of a nonconfigurational base structure simplifies one of the grammar of free word-order languages, this gain is offset by com-plexities elsewhere.

In addition, as this article demonstrates, if we assume a grammar in which word orders are derived from a configurational base by means of movement rules, it is possible to construct a very simple model of usage that is observed data. I therefore adopt a model of grammar in which the generated order is fixed by language-particular parameters and alternates are derived by the generalized movement rule move-alpha.

Structural ambiguity

Though, under the assumption of configurality, the base order of argument is fixed, it is not always obvious what this base order is when there is
a great deal of movement. Given that the head position is either initial or final in the phrase (Kayne, 1983; Speas, 1990; Stowell, 1981), the verb may either precede or follow all its complements, but not appear between them. Clearly, then, examples like (20), in which the surface order is S X v Y, must be derived, either from a verb-final base by postponing one complement, as in (21), or from a verb-medial base by preposing one, as in (22).

(20) autar ho g’hérós Talthubói kéruki didou
    but the-NOM PTC hero-NOM Talthyblos-D herald-D gave
    perikales aethlon
    lovely-ACC prize-ACC
    ‘but the hero gave his lovely prize to the herald Talthyblos’ (II.23.897)

(21) Verb-final structure with postponing
    1o autar ho g’hérós [1o, Talthubói kéruki didou] [ perikales aethlon ],

(22) Verb-medial structure with preposing
    2o autar ho g’hérós [ Talthubói kéruki ], [1o, didou t, perikales aethlon ]

In cases like (20), which have two complements, there is clear evidence of movement, although we may not be sure of its direction. In those cases with only one complement, such as S X v or S v X, it is not even certain that movement has taken place. Thus, the S v X example in (23) may reflect an underlying verb-medial structure, as in (24), or a verb-final structure with postponing, as in (25).

(23) e pot’ Akhiléítos pothē hiketai huisas
    indeed sometime Achilles-G longing-NOM will-come sons-ACC
    Akhaión / sumpanías
    Achaians / all-ACC
    ‘indeed sometime longing for Achilles will come on all the sons of the Achaians’ (II.1.240)

(24) Verb-medial structure with no movement
    1o e pot’ Akhiléítos pothē [1o, hiketai [1o, huisas Akhaión sumpanías ]]

(25) Verb-medial structure with postponing
    1o e pot’ Akhiléítos pothē [1o, t, hiketai [1o, huisas Akhaión sumpanías ],

Distinguishing verb-final from verb-medial structure

Dealing with structural ambiguity of this sort is a common problem in historical syntax, and a number of diagnostic tests have been devised to disambiguate verb position. One often used in Germanic syntax is the relative position of the verb with respect to pronouns or verb particles. Thus, Santorini (1989) noted that, although most verb-medial subordinate clauses in early Yiddish are structurally ambiguous because they can be derived either by verb movement to medial INFL or by complement postponing past a final verb, there is one subset of the data in which this ambiguity is not present. Unlike other constituents that can appear after a final verb, as illustrated in

(26) pronouns and verb particles always precede an unambiguously final verb in Yiddish, as in (27). The tensed verb is underlined in each case.

(26) dz ikh reyn verde fun der ashin (Purim-shpil, 1004)
    that I clean become of the ash
    ‘that I may become clean of the ash’ (Santarini, 1989)

(27) ven du mir meyn kop ab shneydikt (Magen Abraham 2, 1624)
    if you me my head off cut
    ‘if you cut my head off’ (Santarini, 1993)

It can be concluded, therefore, that particles and pronouns do not postpone in Early Yiddish, a restriction that in fact applies generally in Germanic languages. Thus, any superficially verb-medial clause like (28) containing a postverbal pronoun or particle must be underlyingly INFL-medial, with a structure like that in (29). In underlyingly INFL-final clauses, on the other hand, a pronoun or particle will always precede the tensed verb, and it can be assumed that any constituents that follow the verb have been postponed.

(28) vi es ist mir zu kilt (Purim-shpil, 424)
    how it is me so cold
    ‘how I feel so cold’ (Santarini, 1989)

(29) levi I le et l, I et, l, mir zu kilt l, ll

Postverbal pronouns and particles in Yiddish, therefore, provide crucial evidence that at least some verb-medial clauses reflect verb-medial phrase structure and are not derived by postponing. Consequently, phrase structure competition is established. Furthermore, because clauses containing a diagnostic are not ambiguous, an estimate of the rate of verb-medial structure in the language can be obtained by dividing the number of cases with a postverbal diagnostic by the total number of clauses containing a diagnostic (see Santorini, 1993).

The evidence from structural diagnostics

Pronouns. The diagnostic tests just discussed are a little more difficult to apply in AG, largely because of changes in the language that are not related (at least directly) to the change at hand. For instance, although at least some verb particles are independent in Homer, by Herodotus they have largely fused with the verb and thus provide no useful evidence. The pronoun test is more promising, but still has its difficulties.

AG makes extensive use of both empty and clitic pronouns, in addition to full pronouns, and consequently the full pronouns tend to be used in emphatic or contrastive situations. The three possibilities are illustrated in (30).
The marked discourse function of full pronouns potentially affects their position, and they cannot therefore be safely used as a diagnostic. This problem can be avoided in the later two stages by using the third person pronoun *auton* because it is never emphatic (emphatic third person reference being provided by the demonstrative). Thus, the clause in (31a), with an initial demonstrative referring to a previously mentioned person, would never appear as in (31b) with *auto*.

### Forms of *auton* do not appear after a final verb (i.e., a verb preceded by more than one constituent) in any period of AG, and thus, as in Early Yiddish, it is clear that they do not postpone as other NPs do, a fact presumably to be attributed to their relatively light weight, as well as their status as given information. Some typical examples of the use of *auton* are given in (32).

### (32) a. *Periandros* men toutoisi *auton* katelambane

*Periandros*-NOM P these-D him-ACC coaxed

"*Periandros* coaxed him with these things" (Hdt.3.52.6)

b. *sugkaleutontes* Persias hoi megoi

collecting-together-NOM Persians-ACC the-NOM magi-NOM

*ameibibosan auton* epi purgon

mounted him-ACC on tower-ACC

"having collected together the Persians the magi mounted him on a tower" (Hdt.3.75.1)

### (33) Position 1

a. *le autor* [le = hoi Proetus kata mēsato thumō]

but him-D Proetus-ACC evil-ACC planned heart-D

"but Proetus planned evil against him in his heart" (II.6.157)

b. *le hoi* [le =min houtō prothumōs Harpagos metepempasato]

why him-ACC so eagerly Harpagus sent-for

"why Harpagus sent for him so eagerly" (Hdt.1.111.2)

c. *eles* [le = hoi =de leukatē =mei epélulhe nuxs]

but when PTC tenth-NOM me-D came night-NOM erebhēnē]

dark-NOM

"but when the tenth dark night had come to me" (II.9.474)
These two positions can be derived in the following way. Clitics are base-generated in the same position as their full NP counterparts: that is, as complements of the verb, either before it in verb-final clauses or after it in verb-medial clauses. Because of their clitic status, however, they do not remain in this position, but "float" to the left edge of some phrasal projection to which they attach by Chomsky-adjunction, as illustrated in (35). In addition, a rule of prosodic inversion (PI) may apply in a phonological form (PF) that moves the clitic one phonological word to the right, as in (36). Thus, the clitic will always surface as the first or second word in the phrasal unit in which it appears.

When the verb is final in its phrase and therefore governs to the left, the clitic floats leftward as far as possible (i.e., to the IP), as in the examples in (36). The relevant structure is shown in (37). When the verb is medial, on the other hand, I assume the VP acts as a barrier and the clitic must remain within it, as in (38).
This floating of the clitic to the maximal VP created by adjunction is optional, however, as shown by example (41) in which the clitic appears in postverbal position and thus must be adjoined to the minimal VP.

\[(41) \text{hoti e\'g\'o \{ro, ei\, etn\'e \{ro, makran \{ro, ekstrapol\'e} \]
\]
\[\text{since}\ \text{I-NOM to gentiles-ACC far away send =ze} \]
\[\text{you-ACC} \]
\[\text{since I send you far away to the Gentiles} \] (Acts. 22.21)

Thus, I take the adjunction site of the clitic (IP or VP) to provide a diagnostic for the underlying position of the verb. As Table 3 shows, in Homer clitics adjoin almost categorically to IP, whereas in the two later authors both types are found, although they differ in frequency. The data in this table are based on The Iliad 1–12, Herodotus 1–4, and the entire New Testament corpus. The proportion of IP to VP clitic attachment in Herodotus and the Koine is very close to that of pre–versus postverbal pronouns. The difference is in fact not significant (Herodotus: $\chi^2 = .896, p < .5$; Luke: $\chi^2 = .881, p < .5$). The best estimate of the proportion of verb-final to verb-medial structure in these three authors, derived by combining the pronoun and clitic evidence in Tables 2 and 3, is given in Table 4. By this measure, it appears that Homeric Greek is essentially verb-final, whereas the Koine is almost verb-medial. Herodotus, between the two in time, reflects a mixed type in which both phrase structures are well represented.

**QUANTITATIVE EVIDENCE FOR PHRASE STRUCTURE COMPETITION**

The pronoun and clitic data give structural evidence for verb-final/verb-medial competition in Homeric Greek. Here, I present a further quantitative argument for phrase structure variation that is based on the changing distribution of clause types (illustrated in Table 1) and that relies on the hypothesis that usage patterns are stably associated with a particular grammar. The model of usage associated with the verb-final grammar can be based directly on the

Homerian data, since we know from the structural evidence that it reflects virtually no verb-medial structure. The model for the verb-medial grammar can likewise be based on the Koine evidence with some adjustment for the 5% of the data derived from the old verb-final base. The hypothesis that the changing distribution of clause types in AG is due to phrase structure competition is confirmed when the two models are brought to bear on the Herodian data. Although neither model alone fits the data well, a good fit is obtained if a proportion of the data is derived from each model. Strikingly, the ratio of verb-final to verb-medial structure that gives the best fit does not differ significantly from the best estimate of the ratio obtained on the basis of the structural diagnostics.

**Constructing the models**

The model of usage I adopt here is very simple and rests on four basic assumptions.

1. At least some types of syntactic rules apply in the generation of sentences. That is, I assume a fairly direct connection between grammar and usage such that the organization of the grammar is reflected in the patterns of usage (Kroch, 1989). As Guy (1991) noted in his examination of this issue in the area of phonology, the connection between theories of competence and production data is not a necessary one. However, the assumption that there is such a connection, and that it is reflected in usage in a direct way, is the first step in the development of empirically testable theories of language use.

2. Grammatical rules apply independently of one another. A model in which there is no interaction between the application of different rules or multiple applications of the same rule is of the simplest kind and is to be preferred unless evidence to the contrary forces the assumption of interaction.

3. The application rate of grammatical rules on average and across texts is constant. I assume that the factors that determine the application of such optional rules as argument postposing are stylistic or pragmatic (as, e.g., information status, heanness, focus, etc.). Previous studies (e.g., Kroch, 1989; Noble, 1985; Oliveira e Silva, 1982) have shown that the effect of factors of this type is stable: that is, the extent to which each factor favors application of the rule does not change from one context to another.

4. The application rate of these rules is stably associated with the grammar to which they belong. Thus, as long as the grammar does not change, the prob-
ability of rule application will not change, and conversely, an apparent change in the application rate of a rule indicates a change at the level of the grammar.

As I demonstrate, the data strongly support these four assumptions.

The verb-final grammar. Earlier, I exemplified the clause types that are attested in the AG texts and that the grammar must therefore be able to generate. Assuming on the basis of the clitic evidence (Table 3) that Homeric Greek is essentially verb-final, the simplest way of deriving all the alternate orders is by postponing either the subject or the complement or both. We can obtain an estimate of the rate at which this rule applies to subjects and complements from the distribution of clause types observed as follows. Let s equal the probability that the subject postpones and p equal the probability that the complement postpones. The probability of the opposite outcome in each case (i.e., that the subject or complement does not postpone) is 1 − s and 1 − p, respectively, since postponing and not postponing exhaust the possible outcomes. The probability for each clause type is defined by multiplying the probabilities of the appropriate outcome for each argument in the clause. The expected distribution for the sample (e.g., as in Table 6) is then obtained by taking the probability for each clause type times the number of clauses containing the relevant arguments. Thus, for example, the probability of the order S v X is defined by p(1 − s): that is, the probability of the subject not postponing multiplied by the probability of the complement postponing. The formula for each clause type is given in Table 5.

As an example, let’s assign s and p each the value .5. The expected distribution is obtained by substituting .5 for s and p in the formula and multiplying the outcome by the N for clauses with subjects (if the pattern contains an S) or by the N for clauses without subjects (if the pattern does not contain an S). The expected distribution of clause types in the Homeric sample, based on the assumption that both postponing and preposing occur 50% of the time, is given in Table 6, along with the actually observed distribution from Table 1.

As can be seen in Table 6, the observed and expected distributions are not at all close, and thus it can be concluded that the probability of application of postponing in Homer is not .5. In order to minimize the difference between the observed and expected distributions, the estimates of s and p must be adjusted. The best fit is obtained when s = .23 and p = .43, as shown in Table 7.

Some test cases

Previously, I used the distribution of clauses with two arguments to get an estimate of the probability of subject and complement postponing in the verb-final grammar. The validity of the estimates obtained on the basis of these clauses can be checked against additional Homeric data sets. As I demon-
trate here, the results bear out my key assumptions that rules apply independently and at a constant rate.

Clauses with three arguments. Although the number of clauses is small, all possible orders of a verb with three full NP arguments are represented in the sample, with the exception of v S X Y, as illustrated in (42). The lack of an example of the v S X Y pattern is attributable to the general rarity of clauses with three full NP arguments and the low probability of all three postposing, rather than to the ungrammaticality of this one type. I have included an example with a PP, rather than a full NP, to show that this type does indeed occur as expected. As before, the order of the arguments on each side of the verb is not taken into account.

12a. S X Y v
    hûs Akhileus Trœessi ponon kai ἵκτο eπέκα
so Achilles-NOM Trojans-D toll-ACC and care-ACC put 'so Achilles put toll and care on the Trojans' (II.21.525)

b. S X v Y
    autar ho g' hêrôs Talthubôi kéruki didôu
    but the-NOM PTC hero-NOM Talthybios-D herald-D gave
    perikalês aethôn
    lovely-ACC prize-ACC
    'but the hero gave his lovely prize to the herald Talthybios' (II.23.897)

c. X Y v S
    dôru d' ap' allêloi periulôa doûmen amphô
    gifts-ACC PTC come one-another-D lovely-ACC we give both-NOM
    'we both shall give lovely gifts to one another' (II.7.299)

d. S v X Y
    idôn d' Hélenos, Priamoio phîlos paih
    their-G PTC Helenos-NOM Priam-G beloved-NOM son-NOM
    synthetho thumôî houlên
    put heart-D council-ACC
    'Helenos, Priam's beloved son, put their council in his heart' (II.7.44)

e. X v S Y
    Tudoideî men dôke meneprolemos
    son-of-Tydeus-D PTC gave stubborn-in-battle-NOM
    Thrasymerês phagûanon amphiêkes
    Thrasymeres-NOM sword-ACC two-edged-ACC
    'Thrasymeres, stubborn in battle, gave his two-edged sword to the son
    of Tydeus' (II.10.255)

f. v S X Y
    éôthen gar thêsonai peri astu makhên
    at-dawn PTC put about city-ACC battle-ACC
    helikôpes Akheîoi
    glancing-eyed-NOM Achaians-NOM
    'at dawn the glancing-eyed Achaians put battle about the city' (II.24.402)

Table 8 gives the observed and expected values for this data set, minimiz-
Thus, $s = .47$ and $p = .41$. Although the value for $p$ is in the same range as our past results, the value for $s$ is much higher than the previously obtained high value of .29. A closer examination of the clauses with subjects, however, reveals that the probability of postponing varies among types of subjects, some of which favor postponing, whereas others disfavor it. The subjects can be divided into two categories: (1) proper names, either alone or with a formulaic epithet; and (2) other subjects. These two types are illustrated in (44) and (45).

(44) Proper name subjects
a. $kai\,\omega$ thoi peirithoiv Akhaioi and him-ACC obey Achaians-NOM
   'and the Achaians obey him' (II.1.79)
b. $e\,\dot{e}$ pou kogkhalióddoi karé having-abundant-hair-NOM
   indeed surely laugh
   Akhaioi, Achaians-NOM
   'for surely the flowing-haired Achaians laugh' (II.3.43)

(45) Other subjects
a. olekontio de laoi perished PTC people-NOM
   'and the people perished' (II.1.10)
b. atar en pote dasmos hikétai but if-PTC ever distribution-NOM comes
   'but if ever distribution comes' (II.1.166)

Table 10, which compares the distribution of pre- and postverbal subjects in the two categories for clauses with one argument and clauses with two arguments, makes it clear that the major difference between the two data sets lies in the proper name subjects. Whereas the difference in distribution between the other subjects is not significant, $\chi^2 = .961$, $p < .5$, the difference between the proper name subjects is significant at the .01 level, $\chi^2 = 7.387$. This result appears to run counter to the claim that the combined effect of the factors that influence movement is constant across the text. It is important to note, however, that this claim refers to internal linguistic factors such as heaviness, focus, information status, and so forth. The deviant behavior of the single argument clauses, on the other hand, is the result of formal pressures imposed by the structure of the Homeric text.

The Homeric poems are written in dactylic hexameter: that is, lines composed of six feet with the metrical shape $UU$ or $\ldots$. Thus, metrical requirements necessarily have some effect on the language of the poem. The most obvious effect is a process called "localization," whereby certain words or phrases are limited either completely or to a large extent to a certain position in the line (O'Neill, 1942; Porter, 1951). However, the effect of localization on word order is not as extreme as might be expected. First of all, localization is rarely absolute; there are generally two or three positions available for any given item. Second, localization is actually a restriction on metrical types rather than individual words, and because Greek is an inflected language, a lexical item may be localized in a number of positions, depending on how it is inflected. Verbs are especially mobile, since their inflected forms vary widely in metrical type. A related point is that Homeric Greek contains many interchangeable pairs of words with different metrical shapes which may be variants of the same word (Odysseus/Odyseus, where the former is scanned $\textit{U} \ldots$, whereas the latter is $\textit{UU} \ldots$), forms from different dialects (amnes [Aeolic]/hémis [Ionic] 'we', scanned $\textit{U} \ldots$ and $\ldots$), or more or less synonymous forms (doru/egkhas 'spear', scanned $\textit{UU}$ and $\ldots$).

Third, although clauses are often accommodated within a single line, this is by no means necessary. A clause may begin on one line and end on the next, a construction traditionally called "enjambement." A word localized to line-end, therefore, may in general be related to a clause on either its own line or the following line. Compare (46a) and (46b).

(46) a. Enjambed
   ... ophr an Akhaioi
   until PTC Achaians-NOM
   huios emon teisónin...
   son-ACC my-ACC house...
   'until the Achaians honor my son' (II.1.509)

b. Unenjambed
   ... $kai\,\omega$ thoi peirithoiv Akhaioi
   ... him-ACC obey Achaians-NOM
   'and the Achaians obey him' (II.1.79)

Therefore, there are a number of compositional techniques that generally serve to mitigate any extreme effect localization might otherwise have on
word order. The mitigating effect of these factors, however, is greatly reduced in short clauses, especially in short clauses containing proper name subjects. The reasons for this are twofold. First, enjambement is heavily disfavored in clauses made up of only two constituents; thus, this type of clause is generally contained on a single line. Second, whereas localization affects all words to some degree, it is especially evident in formulaic constituents, a class to which the proper name subjects belong; moreover, these formulaic constituents are most often localized to the end of the line (Parry, 1971). Thus, in the majority of subject-verb clauses in which the subject is localized at line-end, the word order will be v S.

The source of the skewed results in Table 10 is the inclusion of clauses with exactly two constituents (i.e., the verb and one argument with no adjunct material, such as adverbs, nonargument NPs or PP's, etc.) because this category includes a type of clause in which structural pressures permit little if any word order variation. Table 11 compares the distribution of pre- and postverbal subjects in the two categories for clauses with one argument and clauses with two arguments, leaving out all clauses with exactly two constituents.

With the exclusion of clauses made up of only two constituents, the difference between the clauses with one argument and those with two disappears (the difference is not statistically significant in either case). Although proper name subjects continue to differ slightly from other subjects, this difference is constant in the two data sets, and the overall probability of subject postposing in the one-argument clauses (.27) is now in the same range as the probability of subject postposing found in the other two data sets (.23 and .29). Modeling all three NP argument data sets together gives equally good results, with the $\chi^2$ reaching a minimum of 8.51 (assuming the values $s = .25$ and $p = .42$).

Clauses with prepositional phrases. The final data set to be tested is made up of clauses with two arguments, one of which is a prepositional phrase (PP). These include both clauses with a subject and a PP complement and those with no subject in which one of the complements is an NP and the other

(47) a. S P v
hoi men Akhaion keres epi kkhoni pouloboteirei
the-NOM PTC Achaian-G fate-NOM to earth-D prospering-D
hezathen
settled
‘the fate of the Archaians settled to the prospering earth’ (II.8.73)

b. P S v
Tryoban anth hektosn te dekosiun te hekastos steter
Trojans-G against 100-ACC and 200-ACC and each-NOM stood
en polemi in battle-D
‘each [of you] stood against 100 and 200 of the Trojans in battle’ (II.8.233)

c. X P v
all hote de opa te megallen ek stethos heit
but when PTC voice-ACC and great from chest-G he-sent
‘but when [he] sent the great voice from his chest’ (II.3.211)

d. P X v
kai eis hala lamata ballon
and into sea-ACC washings-ACC they-threw
‘and [they] threw the washings into the sea’ (II.1.314)

e. S v P
leukheo men hoi keita epi kkhoni poulobotereirei
armor-NOM PTC his-D lies on earth-D prospering-D
‘his armor lies on the prospering earth’ (II.3.195)

f. P v S
thi gar epi phrasi theke thea leukolenos
his PTC in heart-D put goddess-NOM white-armed-NOM
Hera
Hera-NOM
‘the goddess white-armed Hera put [it] into his heart’ (II.1.55)

g. X v P
nun d’ age nea melainen erussamen
now PTC come ship-ACC black-ACC we-drag
eis hala dian
to sea-ACC divine-ACC
‘but come now, let us drag a black ship down to the divine sea’ (II.1.141)

h. P v X
aps d’ es kouleon ose megas kasthios
back PTC into scabbard-ACC he thrust great-ACC sword-ACC
‘he thrust the great sword back into its scabbard’ (II.1.220)

i. v S P
etha de Athene ouranotthen
came PTC Athena-NOM heaven-from
‘Athena came from heaven’ (II.1.194)
TABLE 12. Best fit of observed and expected distribution of clauses containing PP arguments in Homer

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Formula</th>
<th>Observed Distribution</th>
<th>Expected Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>S P v</td>
<td>(1 - s)(1 - m)</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>X P v</td>
<td>(1 - p)(1 - m)</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>S v P</td>
<td>m(1 - s)</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>P v S</td>
<td>s(1 - m)</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>P v X</td>
<td>p(1 - m)</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>X v P</td>
<td>m(1 - p)</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>v S P</td>
<td>ps</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>v X P</td>
<td>pm</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

*N = 58, n for clauses with subjects = 36, n for clauses without subjects = 22; s = .28, p = .48, m = .37; $\chi^2 = 2.78, p > .98$.

TABLE 13. Probability of subject and NP complement postposing in four Homeric data sets

<table>
<thead>
<tr>
<th>Subject Postposing</th>
<th>Complement Postposing</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 NP argument</td>
<td>.27</td>
<td>.41</td>
</tr>
<tr>
<td>2 NP arguments</td>
<td>.23</td>
<td>.41</td>
</tr>
<tr>
<td>3 NP arguments</td>
<td>.29</td>
<td>.41</td>
</tr>
<tr>
<td>1 NP/1 PP argument</td>
<td>.28</td>
<td>.48</td>
</tr>
</tbody>
</table>

TABLE 14. Distribution of verb-final clauses in Luke assuming values of s and p from Homer

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Formula</th>
<th>Estimated Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>S X v</td>
<td>(1 - s)(1 - p)</td>
<td>2.02</td>
</tr>
<tr>
<td>S v X</td>
<td>p(1 - s)</td>
<td>1.5</td>
</tr>
<tr>
<td>X v S</td>
<td>s(1 - p)</td>
<td>.67</td>
</tr>
<tr>
<td>v S X</td>
<td>sp</td>
<td>.49</td>
</tr>
</tbody>
</table>

*N = 5.1; s = .25, p = .42.

In order to fit the data in (47), we need to add another variable, m, to stand for the probability of PP postposing. Table 12 gives the best fit for these data, and once again the values of s and p are quite close to those obtained from the other data.

Table 13 compares the probability of postposing for subjects and NP complements obtained from each of the four data sets tested here. In each case the probabilities are quite close. This convergence of results from different samples gives strong support for the model of usage assumed here. The fit of the model to the data under the assumption of independence is good in each case. Once again the probability of application in the case of subjects and complements is fairly stable.

The verb-medial grammar. On the evidence of the pronouns and clitics discussed earlier, the estimate of the amount of verb-final structure in Luke is about 5%. The first step in constructing the usage model associated with the verb-medial grammar, therefore, is to remove the verb-final tokens from the sample. Although it is generally impossible to tell for any individual token whether it is underlying verb-final or verb-medial, we know from the clitic/pronoun data that about 5% of the sample (0.05 * 102 = 5 tokens) is verb-final. These five tokens can be distributed according to the values of s and p associated with the verb-final grammar (s = .23 and p = .43, obtained by averaging the probabilities for the Homeric data in Table 13). The distribution of these five tokens is given in Table 14. As clauses with three arguments, and even clauses with two complements, are rare in the Koine data, I have only used clauses with a subject and a complement here.

Having calculated the distribution of the verb-final clauses, the verb-medial estimates are obtained simply by subtracting the verb-final clauses from the observed distribution. Table 15 gives the results.

The verb-medial grammar must generate the same range of clause types as the verb-final grammar discussed in the previous section. Because complements are base-generated to the right of the verb, however, a rule of preposing is needed to derive apparently verb-final orders. If we let s equal the probability of complement preposing and p continue to stand for the probability of subject postposing, and solve for the variables as before, we obtain the results in Table 16. The estimated verb-medial distribution from Table 15 has been rounded in Table 16. As usual, the $\chi^2$ is based on the actual estimates.
The assumption of independence and a fixed probability of application for each rule again produces an expected distribution that is close to that observed. A comparison of Tables 13 and 16 reveals that, whereas subjects postpone at about the same rate whether the clause in question is generated by the verb-medial or the verb-final grammar, the complement postponing associated with the verb-medial grammar applies at a much lower rate than the complement postponing associated with the verb-final grammar.

The intermediate period

Having established the probabilities for the usage models associated with the verb-final and verb-medial grammars, the Herodotean data can now be tested. First, it is clear that neither the usage model of the verb-final grammar nor of the verb-medial fits these data. Table 17 gives the expected distribution of the Herodotean data under the hypothesis that it is generated by the verb-final grammar of Homer, and Table 18 gives the same under the hypothesis that Herodotus is generated by the verb-medial grammar deduced from the Koiné. In both cases, the difference between the observed and expected distribution is significant.

In Tables 17 and 18, it is crucially assumed that the usage model is stably associated with its grammar; that is, the clauses derived from the verb-final grammar will continue to be distributed according to the values of \( s \) and \( p \) obtained from the Homeric data, and the clauses derived from the verb-medial grammar will be distributed according to the values of \( s \) and \( q \) obtained on the basis of Luke. An alternate way to approach these data, however, is to assume that Herodotus is verb-final like Homer, and that the difference between Homer and Herodotus simply reflects a rise in the rate of complement postponing. Solving for the values of \( s \) and \( p \) that give the best fit for the Herodotean data, we get the results in Table 19.

As the fit of the data to this model is quite good, it might be claimed that it is not the grammar that has changed, but rather that the probability of
TABLE 19. Observed and expected distribution of clauses in Herodotus assuming a verb-final grammar and allowing s and p to vary*  

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Formula</th>
<th>Observed Distribution</th>
<th>Expected Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>S X v</td>
<td>((1-s)(1-p))</td>
<td>34</td>
<td>38</td>
</tr>
<tr>
<td>X Y v</td>
<td>((1-p)^2)</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>S v X</td>
<td>(p(1-s))</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>X v S</td>
<td>(s(1-p))</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>X v Y</td>
<td>(2p(1-p))</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>v S X</td>
<td>(sp)</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>v X Y</td>
<td>(p^2)</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

*\(N = 134, \, n\) for clauses with subjects = 129, \(n\) for clauses without subjects = 5; \(s = .28, \, p = .60; \chi^2 = 3.38, \, p > .9\).

complement postposing has risen from about .43 in Homer to around .6 in Herodotus. There are two arguments against taking this approach, however. First, it provides no account of the behavior of the pronouns and clitics. If the Herodotean grammar is the same as the Homer, why do the clitics behave so differently? They are not affected by postposing in any way, and yet their distribution changes. Likewise, although pronouns do not postpose, their distribution changes between Herodotus and Luke. These facts remain unaccounted for under the hypothesis that change is simply in the probability of postposing.

Second, evidence from other studies examining the effect of various factors on a number of different types of language change have clearly shown that the probabilities associated with these factors do not change. A number of examples of this type are discussed in detail in Kroch (1989). In addition, in a directly parallel case, Santorini (1993) found that the rate of postposing in Yiddish remains stable over time (see her Table 5). Thus, the assumption that the change in distribution of clause types between Homer and Herodotus is the result of an increase in the rate of postposing runs counter to evidence provided by other changes in the language, as well as to our present understanding of the mechanism of language change based on a number of studies in other languages.

I turn now to the hypothesis that the differences in the distribution of clause types among Homer, Herodotus, and Luke are the result of competition between verb-final and verb-medial structure. Because we know the rates of pre- and postposing associated with each grammar from Homer and Luke, we can now test the Herodotus data to see if a good fit can be obtained by assuming that some proportion of the data is verb-final and the rest verb-medial. This is done by calculating the distribution of a part of the data as if it were verb-final (using \(s\) and \(p\) from Homer) and distributing the rest as if it were verb-medial (using \(s\) and \(q\) from Luke). The estimated tokens for each clause type are then added together and compared to the observed total. Varying the proportion of verb-final/verb-medial tokens in order to minimize the difference between the estimated total (estimated verb-final tokens + estimated verb-medial tokens) and the observed total results in the estimate that 62% of the data is verb-final (see Table 20). Adjusting the proportion of verb-final/verb-medial structure any further in either direction makes the fit worse. To illustrate this point, Table 21 gives the chi-square for a spread of possible verb-final/verb-medial mixes.

From these results it is clear that Herodotus cannot be derived solely from either the verb-final grammar of Homer or the verb-medial grammar of the Koine. On the other hand, a model in which a little over 60% of the data is from a verb-final base and the rest from a verb-medial one gives a fairly good fit. Striking and independent confirmation of the quantitative analysis and the assumptions it is based on comes from the near-identity of the best estimates of the percentage of verb-final structure that are obtained from it (62% verb-final) and from the structural analysis summarized in Table 4 (60% verb-final).

CONCLUSIONS

It has been proposed (Kroch, 1989; Pintzuk, 1991; Santorini, 1989) that syntactic change may be achieved through competition between subsystems of grammar, rather than by a real analysis of ambiguous clauses at the end of the change, as is often assumed (Lightfoot, 1991). The AG change from SOV to SVO gives further support to this hypothesis. The distribution of clitics and pronouns provides structural evidence that, whereas Homer is essentially verb-
final, both verb-final and verb-medial phrase structures occur in Herodotus. Synchronic phrase structure variation is also reflected in the changing distribution of clause types between the Homeric and Herodotean periods. The Herodotean data fit neither the verb-final Homeric model nor the later, largely verb-medial Koine. But if part of the data is modeled according to the original verb-final grammar and part according to the innovative verb-medial model, a good fit is obtained: the percentage of the data that must be attributed to each grammar in order to obtain the best fit is extremely close to the percentage expected on the basis of the structural evidence.

The hypothesis that the change in the distribution of clause types reflects the change in the grammar (rather than a change in usage) rests on a number of assumptions about the nature of the association between grammar and usage. Two of these—that optional rules apply independently, and that they apply at a constant rate—are explicitly tested against the Homeric data and found to hold. Thus, the model used—one that assumes no interaction between one application of a rule and another—gives a good fit for all four of the Homeric data sets tested. In addition, all the data sets (with one exception, which is clearly related to the special character of the Homeric text) yield approximately the same probabilities for the application of postposing. Previous work suggests that the effect of stylistic/pragmatic factors on grammatical rules is constant across time as well as across contexts. Although the quantitative data considered here do not fit this hypothesis any better than one in which it is assumed that the effect of these factors does change (i.e., the difference between Homer and Herodotus is the result of an increase in the rate of postposing), the structural evidence clearly contradicts the view that there is no change in the grammar. Overall, therefore, these data support the hypothesis that pragmatic/stylistic effects are constant across time.

NOTES
1. The head-complement orders of some categories do seem to depend on others, however. Thus, whereas (NFL-medial) order seems to be compatible with either a head-initial or head-final VP, INFL-final order only coexists with verb-final order.

2. The VP in the V-raising case can be head-initial or head-final, with the same results.

3. The following abbreviations are used in the examples: PTC = particle, PV = preverb; NOM = nominative, ACC = accusative, G = genitive, D = dative; / = indicates the fine-end; ( ) encloses textual emendations.

4. In addition, some forms of eimi 'to be' are clitic, and it is possible this would have an effect on their position.

5. A small number of examples occur in Luke with the verb eimi 'to be'.

6. An extremely commonplace construction in Greek, especially in subject position, is for an NP to contain a participial clause that agrees with it in case, as in (i).

7. [prosede] Kambu6to] et a)igipton xBeBilo [a)pe] Amasin spupus

8. sending-NOM Cambyses-NOM to Egypt-ACC herald-ACC asked Amasis-ACC daughter-ACC

9. Cambyses sending a herald to Egypt asked Amasis for his daughter' (Hdt. 3.1.1)

10. I have included these in the sample, both when the noun is expressed, as in (i), and also when it is empty, as in (ii).

11. [prodÓa] [e]leisuf peri y8.e amavavn]t ou]X Christou

12. foreseeing-NOM talked about the-G resurrection-G the-G Christ-G

13. 'the foreseeing talked about the resurrection of Christ' (Acts.2.31)

14. Pintzuk (1991) used the same texts in Old English with good results.

15. Although not all full pronouns are emphatic, the danger of relying on their position in the clause to help distinguish the emphatic from nonemphatic, and thereby arguing in circles, makes their use as a diagnostic too risky.

16. In Homer, autos generally means 'self'.

17. The status of clitics as pronominal affixes goes back at least to Zwicky (1977) and is explored in great detail in Klavans (1982, 1985).

18. The details of prosodic inversion are too complicated to go into here. For further information, see Halpern (1992) and Taylor (1990, 1992, 1994, forthcoming). The term prosodic inversion is from Halpern.

19. Clitics like de are also clitic and in general behave the same as other clitics (for exceptions, see Taylor, forthcoming).

20. There is one example that runs counter to this hypothesis in that one clitic appears in the VP and one in IP.

21. [epe] =me Zeus edike =so

22. since me-ACC Zeus-NOM gave you-D

23. 'since Zeus gave me to you' (Hdt. 1.267.1)

24. It is unclear, however, whether to take this one clause as a serious impediment to the analysis. Clauses with two pronominal clitics are extremely rare in the Classical period and beyond, and it is therefore difficult if not impossible to get a feeling for the "normal" syntax of these cases. The only other clause with two object clitics that I have been able to find is ambiguous: the verb is initial in the clause, and thus the clause-initial clitic is also VP-initial.

25. [haines] =spie proostata =keowtm

26. whoever-NOM them-D owes anything-ACC

27. 'whoever owes them anything' (Hdt. 1.61.3)

28. As all the other available evidence supports the analysis given here, there seems to be no way to resolve this problem from the available data, I have decided to ignore this case.

29. This table only includes unambiguous cases: that is, those with an overt preverbal subject. When the subject is missing or is postverbal, the IP and NP boundaries fall together, and thus it is impossible to tell the attachment site of the clitic.

30. Although the word orders in question might in principle be derived by processes such as verb-fronting to C, I have kept the model of usage developed in this study as simple as possible by abstracting away from this possibility. I take the extent to which the model fits the data to indicate in part the correctness of this strategy.

31. The probabilities are estimated using an iterative maximum likelihood method.

32. The expected distribution has been rounded to the nearest integer. The x², however, is based on the difference between the actual estimates and the observed data. This is true of all the tables, except Table 14.
18. I have added to the sample all examples of clauses with the verbs didišmi 'to give' and tithēmi 'to put' with three full NP arguments from The Iliad.

19. The proper disjunction here may be "formalistic" versus "nonformalistic" subjects. However, beyond discussions of formulas involving proper names, there is a great deal of disagreement about what constitutes a formula, and thus I have not made the attempt to distinguish subjects according to this criterion.

20. The possible data set—the 2NP/1PP type, as in English the girl put the book on the table—is too rare to be used. The verb tithēmi 'to put' in Greek can take either two object NPS (one accusative, one dative, like the verb didišmi 'to give') or an NP and a PP, so many of these cases are included in the 3NP set.

21. Note that I am also assuming that all verb-final clauses have subjects. Although this is undoubtedly not the case, in both Homer and Herodotus the number of two-argument clauses with subjects far outweighs those with complements, and so the effect of this assumption, if any, will be very small.

22. The values for s and p for the verb-final grammar are those obtained when all three sets of data with NP arguments are considered together. The values of s and p for the verb-medial grammar are taken from Table 16.

23. The chi-square in Table 18 is distorted because of the estimate of zero in the X × Y type. If the cells without subjects are removed, however, the resulting chi-square of 112.89 remains significant (p < .001).

REFERENCES


Noble, Shawn. (1985). To have and have got. Paper presented at NAWAVE-XIV. Georgetown University.


