

A Verb-form Frequency Count

H. V. GEORGE

THE CENTRAL INSTITUTE OF ENGLISH, Hyderabad, India, has recently published a report on a verb-form frequency count.¹

The purpose of the count

Thirty years ago, reports were appearing which stated what words occurred, and with what frequency, in general English. These and subsequent reports popularized the idea that the words most used, and most widely used, are the most useful ones. Their authors thought that a list of words showing their frequency ranking would give objective guidance in deciding priorities and the amount of attention to give to particular words at each stage in a course. Such lists have indeed become part of the standard reference material for textbook writers and teachers.

The count of verb-forms was begun in a similar belief that the most used and most widely used items of verb-form usage are likely to be the most useful ones, and that a list showing their frequency ranking will help to decide where to place items in a teaching programme, and how much teaching time and effort they each merit.

The schedule

The first task was to make a schedule. Many accounts of tenses and tense usage were available, but no comprehensive list of other items of verb-form usage. However, most courses give a great deal of attention to presentation of a variety of tenses, and one of the things worth checking was how far this attention is justified. A statement of the relative frequency of occurrence (the 'functional load') of tenses alone would, therefore, not have been enough, and a fairly extensive survey of verb-form usage had to be made.² In the end, a schedule with 169 main items was established. The items were grouped under the four sections *plain stem*,³ *to + stem*, *stem + -ed* and *stem + -ing*, and each was presented with a description and one or more illustrations. Here are two examples:

TO + STEM 25 in free adjuncts *To take* another example, the . . .
STEM + -ED 6 participle + to + stem *We weren't allowed to leave.*

¹ Available as Monograph of the C.I.E., No. 1, from the C.I.E., Hyderabad 7, India, price 3s. (or U.S. \$0.40c.) including registered postage.

² Much use was made of E. Krusinga and P. Brades, *An English Grammar*, Vol. 1, 8th edition, Noordhoff, Groningen, Holland.

³ All technical expressions are exemplified at the end of the article.

The procedure

The Central Institute of English has from fifty to sixty teachers in residence at a time, so that each of them could be assigned a small group of items which he specialized in identifying. Several weeks were spent in preparation, that is, in getting familiar with the schedule and the counting routine, and in practising identification and attribution. After counting began, difficult attributions were discussed at weekly meetings of counters with a common area interest, one group consisting of those looking for participial stem + -ed, another of counters of occurrences of Simple Past, and so on. All the participants recorded all the occurrences of 'their' items, in the same materials, so that every occurrence of every verb-form usage listed on the schedule should have been noted. Over two hundred teachers took part in the count, which reached a total of 108,784 occurrences.

The source materials

Various types of English are represented: a 'Conversation Reader', two plays, three novels, a travel book, two issues of a newspaper; and four books of a popular, factual nature, together with sample pages from *Chambers's Encyclopaedia*. (See the list at the end of the article.)

Some key figures

Full figures for the combined materials, and for each type of material separately, are given in the report.

The following ten items occur with frequencies of 17 per 1,000 or more¹ in the combined materials.

Table 1

Rank	Item	Total occurrences	Per 1,000 occurrences
1	Simple Past Narrative	17,004	156
2	Simple Present Actual	13,084	120
3	Simple Past Actual	9,038	83
4	Simple Present Neutral	7,597	70
5	Past Participle of occurrence	6,395	59
6	Past Participle of state	3,541	33
7	verb + to + stem	2,993	27
8	stem + -ing = adjective	2,674	25
9	stem + -ed = adjective	2,530	23
10	plain stem after Imperative <i>Don't</i>	1,837	17

¹'Per thousand' figures are obtained by converting into a comparable proportion of 1,000 the occurrences of a particular item compared with the total occurrences of all items, e.g. 17,004 : 108,784 is equivalent to 156 : 1,000.

²17 per thousand represents one occurrence per page, assuming 300 words per page and 1 verb-form occurrence for every 5 running words.

Of these items, the second, Simple Present Actual, or the use of the Simple Present to refer to the present ('actual') moment, does not always find a place in school courses. Indeed, it is often taken as axiomatic that the Present Progressive (Continuous) is the form used to refer to 'now'. In these materials used in the count, the Present Progressive occurs 627 times in this context, that is, with a frequency of approximately 6 per 1,000 occurrences.

The third item too receives rare recognition. It is the Simple Past form used to give contextual information or state circumstances 'at that moment'. Usually, this function is assigned to the Past Progressive (Continuous), which, in these materials, occurs in this role 509 times—5 per 1,000 occurrences.

The distribution of items

The figures just given do not agree with current general impressions of the relative importance of verb-forms and functions. It is a pertinent question, therefore, whether the overall figures represent the language satisfactorily. The question can be answered by reference (a) to the comparative ranking of items in the various kinds of material (b) to the 'per thousand' occurrence figures for the various kinds of material.

The following table shows how the ten items of Table 1 rank in the various kinds of material:

Table 2

Rank in Table 1	Rank in: conv.	plays	novels	travel	news	factual
1	3	2	1	2	6	1
2	2	1	2	3	1	3
3	5	7	3	1	2	5
4	4	8	4	4	5	4
5	1	22	8	5	3	2
6	24	9	10	8	10	6
7	7	5	5	12	7	10
8	30	24	12	10	4	8
9	—	25	29	7	9	7
10	4	4	6	29	30	30

The four most frequent items in the overall figures are seen to be high ranking items in every kind of material. The occurrences of these four items together amount to: conversation 528 (per 1,000 verb-form occurrences in this material);¹ plays 459; novels

¹The 'per thousand' figures in this section are obtained by converting into a comparable proportion of 1,000 the number of occurrences of an item in a particular kind of material, compared with the total verb-form occurrences in that material.

440; travel 492; newspapers 328;¹ factual 409; aggregate 429.

Looking at the per thousand occurrences in detail, one observes that Simple Past Narrative finds its highest frequency in novels (218), its lowest in plays (39), where, nevertheless, it is the second most frequent item. Simple Present Actual is particularly frequent in plays with the highest number per thousand occurrences (375) of any item on the schedule. Simple Past Actual is the highest ranking item in the travel book (172), Simple Past Narrative taking second place (146). Simple Present Neutral is most frequent in the Conversation Reader (271), presumably reflecting the general nature of most conversations: it ranks high in all materials.² Verb+to+stem appears more modestly in all materials, with frequencies from 18 to 40. These items may be said to constitute the outstanding requirements for English of all kinds.

Stem+ed as a Past Participle of occurrence is most frequent in factual material (second rank, with 111 per thousand occurrences) and is conspicuous in newspapers, novels, and travel; inconspicuous in plays and conversation. This is the pattern for participial items. Stem+ed as a participle of state shows a very similar distribution, but occurs modestly in plays (9).

Stem+ing as adjective in a noun group is most frequent in newspapers (54) and is probably a feature of descriptive writing generally (factual 31, travel 27, novels 18, plays 9). As a free adjunct, stem+ing is most prominent in plays (31), and has a modest rank in most materials.

The plain stem after Imperative *Don't* appears in conversation, plays, and novels (28, 27, and 28 per thousand occurrences respectively), and, as might be expected, has no importance in other materials. The Imperative follows the same pattern, but is less frequent. Vicarious *do/did* reaches the 17 per thousand mark in conversation, 10 in plays, and is inconspicuous elsewhere.

The Present Progressive appears in conversation (10), plays (15) and newspapers (13). However, the *proportion* of the Present Progressive to Simple Present Actual occurrences is not much affected by its modest appearance in these materials; in fact, in

¹As the figures suggest, greater variety of verb-form usage is shown in newspapers than in any other kind of material. In conversation, nine items appear once per page or more often; in plays eight do so; in newspapers 15.

²It is not the writer's intention to suggest that *learners* should be made aware of the difference between Simple Present Actual and Simple Present Neutral, or between Simple Past Narrative and Simple Past Actual. For them, these four items are two. Nevertheless, such distinctions are essential. Not distinguishing Simple Past Actual from Simple Past Narrative, for instance, one might imagine that Simple Past is always Narrative, and that for contextual information the form *was/were* + stem +ing is required. If, as obviously it is, 'As we waited . . . ' is alternative to 'As we were waiting . . . ', a verb-form count must state their relative frequency of occurrence.

plays, in which the Present Progressive shows its maximum prominence, it represents 'now' once for every twenty-five occurrences of Simple Present Actual, the overall proportion being one to twenty.

The most frequent verb-forms have now been examined, those with general distribution and those with frequent occurrence only in some kinds of source material. It is not without interest to look at the other end of the scale, and list specimen items showing low frequency occurrence.

On the assumptions of even distribution, 'standard' pages of 300 words, and one verb-form for five consecutive words, we would expect to find one occurrence of Future Progressive in every 77 pages, Future Perfect would make an appearance once in 150 pages, and we would wait for 1,800 pages between successive occurrences of Future Perfect Progressive. Expressions in the form *You are to do that*, *She is not to go* might be expected once in 225 pages, and those in the form *is/was*, etc., *about to do something* once in 180 pages.

The application of results

The results of the count have their most obvious application in showing which items require persistent attention, and which can well be omitted from early courses. Below a certain level of frequency (certainly below about the 6 per thousand occurrence level) guidance is less certain and the figures could not be used for comparisons among low frequency items, whose rank, in a count of this size, is necessarily due in part to chance distribution.¹

Perhaps as important an application of the results is the effect of the ranking of items upon methods of teaching and establishing verb-form usage.

Present-day descriptions of English contain no statements of proportion. It is as though 'Descriptive Geography' presented detailed accounts and entered upon minute comparative studies of the world's countries, without any statement of their sizes.

The theoretical position is reflected in the practical teaching of the subject. Many presentations, exercises, and tests are constructed on the assumptions that each bit of usage has its context and that this context contrasts equally with those of other bits; in terms of school English, for instance, Simple Past is contrasted in turn with Simple Present, Present Perfect, Past Progressive, and Past Perfect, as though each contrast were necessary and as though all these forms had equal standing.

¹In fact, the pattern of occurrences is like that of the vocabulary counts, which give valuable guidance up to, say, the two thousand word level, but from which one could not deduce that words in the sixth thousand were necessarily more important than words in the seventh.

Teaching, however, like any other process, shows less than 100 per cent efficiency. Suppose one sets the efficiency-figure at the improbably high figure of 95 per cent for any single item: then confusion or error may be anticipated on 5 per cent of occasions for the use of this item. When one attempts to establish a second, 'contrasting' item, and the contrast itself, a much lower overall efficiency may be expected. If now, in everyday English, out of every 100 occurrences of the two items one item accounts for 95 and the other for 5 occurrences, one can foresee more confusion or error *through the teaching* than could result from the omission from the programme of the 5 per cent item.

To the writer, this kind of consideration is the ABC of 'Method'. 'If I teach this now, what associations, and what interferences, are likely to ensue? What previous teaching will I confirm, and what will I modify or upset? Is resultant learning likely to represent a sufficient return for the effort expended; is the probable efficiency too low to justify the effort at this stage in the programme; or is the effort perhaps likely to produce a negative reward? Ought I then to try to imprint this now, or should I pass over it cursorily, hoping for eventual recognition knowledge, or ought I to avoid all attention to it?' These are questions to which impressionistic answers no longer suffice: it was to provide more reliable answers in the verb-form area that the count was undertaken.

If one accepts the supposition that the most used and most widely used items are the most useful ones; one necessarily keeps apart, and teaches separately, many items which at present seem to offer 'grammatical contrast', occupy unrealistically prominent places in school courses, and are taught with minimum efficiency because mutual interferences are aggravated by inadequate separation and contrastive presentation and exercise.

Exemplification of technical expressions

plain stem:	mend	arrive	fetch	give
to+stem:	to mend	to arrive	to fetch	to give
stem+-ed:	mended	arrived	fetches	gave/given
stem+-ing:	mending	arriving	fetching	giving

Simple Past Narrative: He *got* up, *went* to the door, *opened* it, and *looked* out.

Simple Present Actual: This broadcast *comes* from the B.B.C., London.

Do you see how well Mahmud is running?
(Present Progressive 'now') He *deserves* to win.

Simple Past Actual:

The old woman *paused* (Narrative), her hand on the loaf of bread . . . The sea *sounded*. Through the wide-open window *streamed* the sun onto the yellow varnished walls and bare floor. Everything on the table *flashed* and *glittered* . . . She *smiled* (Narrative).

Simple Present Neutral: The Ganges *rises* in the Himalayas.

Past Participle of occurrence: It was *gone* in a flash.

Past Participle of state: He was *delighted* to hear it.

verb+to+stem: They *decided to leave*.

stem+-ing=adjective in a noun group: a *floating* island

stem+-ed=adjective in a noun group: the yellow *varnished* walls

Present Progressive 'now': What *are* you *doing*?

Past Progressive 'at that time': He was still *talking* when I left.

Stem+-ing as free adjunct: *Talking* about participles, have you . . .

Imperative *Don't*: *Don't go* yet.

Imperative *Stay* where you are.

Vicarious *do/did*: Did you see him? I *did*.

List of source materials

- P. A. D. MacCarthy: *English Conversation Reader* (Longmans, 1956)
 I. B. Priestley: *Labouring Grove*
 G. B. Shaw: *Arms and the Man*
 E. M. Forster: *A Passage to India* (Penguin)
 P. Frankau: *The Bridge* (Four Square Book)
 D. H. Lawrence: *Sons and Lovers* (Penguin)
 P. Fleming: *One's Company* (Penguin)
The Observer, 10 December 1959
The Daily Telegraph, 22 August 1960
 I. Dugan: *Man Explores the Sea* (Pelican)
 W. Slinckin: *Minds and Machines* (Pelican)
 G. Williams: *Economics of Everyday Life* (Pelican)
 A First College English Course (Social Sciences) (C. I. E. Hyderabad)
 A First College English Course (Sciences) (C. I. E. Hyderabad)
 Chambers's *Encyclopaedia*, Vols. I to IV, p. 1 and every 50th page to p. 451 in each volume