George (1961 c) Bulletin of the Central Institute of English (Hyderabad) (1961) issue 1

REPORT ON A VERB-FORM FREQUENCY COUNT CARRIED OUT IN THE C. I. E. HYDERABAD

H. V. GEORGE

BY the time this article appears, three groups of trainees at the Central Institute of English (preceded by a pilot group in the Second Course) will have been counting the frequency of occurrences, in various types of English, of verb-forms in all kinds of phrase, clause and sentence constructions. They will have noted some 80,000 successive occurrences, each of which will have been assigned to one of the 180 sections of a schedule. All this indicates a vast amount of industry. Here is a short account of the motivation and purpose of the work and of the chief features of the results.

1. *Purpose and application*

The verb-form frequency count stems from the idea that, speaking generally, use indicates usefulness. We are finding out the actual use English people make of the verb-forms in their language.

The count may serve a practical need therefore, to the extent a course-designer or a teacher considers that usefulness for teachers of a language coincides with usefulness for general users of the language, and to the extent he considers that teaching priorities should correspond generally with the degrees of usefulness of the teaching points.

There are two kinds of priority in teaching: one determines the sequence of teaching points; the other, the amount of work the teachers and the learners direct to each of the teaching points.

Teaching priority in the first sense depends on consideration of many factors, and usage frequency-counting may, eventually, take a place among them. The order of presentation of teaching points is, in any case, mainly decided by the Course-designers. The second type of priority-attribution is particularly relevant to the general character of the students' English after they have worked through a course. This is largely in the control of the teachers. If, for instance, all the tenses have been taught with equal thoroughness, and drilled with impartial application, then the students' English is likely to differ from native English in two ways: it is likely to show a wider and a more even distribution of usage: and it is likely to show personal usage-habits which may be "correct", in so far as they represent features which could be used by native speakers, but which

cause these features to figure disproportionately. For example, some foreigners use "would (characteristic repetition) + stem" whenever they refer to a past habit. The count shows that in our material, "would + stem" appears in such contexts only once for fourteen times the "Simple Past (Habitual)" occurs. An awareness of the relative frequency or infrequency of occurrence of specific teaching points might guide teachers in their practical work ; often, course-book treatment needs supplementing, but equally often learners need warning against over-use of what they are learning.

The count covers all verb forms, whether they appear or not in tenses : tense usage in English is complicated, and course-designers and teachers often concentrate attention on tense usage, to the comparative neglect of verb forms not appearing in tenses. The count clearly indicates what tenses are infrequently required, and what verb forms are prominent features of the language.

Sometimes the use of a particular verb form is optional ; checking what is being taught against actual frequency of use could result in forbearance from bringing into students' active English usage which is both optional and infrequent. The time gained might be spent on usage which is frequent and obligatory.

Often context allows substantially equivalent expression of meaning through a variety of language forms. Perhaps students should recognise the whole variety of forms. However, when various linguistic features show overlapping areas of usage, it is misleading to teach one of these features as distinctive, when in reality it may occupy only a minority position in this area. In our schedule, for instance, we have the following passage :—

> "The old woman paused, her hand on the loaf of bread, her gaze out of the open door into the garden. 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."

The author might have written, but did not write :

> ". . . The sea was sounding. The sun was streaming through the wide-open window. . . . Everything on the table was flashing and glittering."

Knowing the relative frequencies of occurrence of simple and progressive (continuous) forms in such contexts is useful as a control over our teaching of the uses of the simple and progressive forms.

As a corollary to the overlapping usage of some verb forms, we observe each verb form doing a number of language tasks. Teaching improves if these uses are dealt with separately, and distinctly, and it is useful to know the usage distribution within each form.

When the coverage of material becomes wide enough for additional counting to make only insignificant changes to the results, it will be possible to use these for research. For instance, a writer's individual verb form usage may be appraised by comparison with average usage, and his personal stylistic habits, in this language area, may be revealed ; and the usage of previous centuries may be compared with contemporary usage.

2. *Presentation*

Before counting could begin, a presentation had to be decided upon. There was no readily available, comprehensive account of verb form usage, and this had to be made. Some comment may be required about the schedule eventually used. It has four main sections, namely :—

1. plain stem*
2. to + stem
3. stem + -ed
4. stem + -ing

These sections represent a formal division ; except that the third person singular Simple Present tense inflexion is included in the plain stem form. Inside each main section, the counting items are arranged so that the simplest come first, and the tense usage comes last. Each item, or group of items, is syntactical, and recognizable as such in the texts. To this extent, then, the schedule is formal. Inside a group of items, however, a single syntactical feature may be broken down according to criteria of "meaning" or "function". For example, the group of items numbered 7 to 12 in the plain stem section all represent the plain stem after "can", and number 9, for instance, represents "can, indicating permission given by the speaker". The schedule may be said then to consist of a descriptive account of verb form usage inside a formal framework. Examples are given for each item, so that what is counted is clear, even though there might not be agreement about the name which has been given to describe the examples. The nomenclature is fairly traditional, and not entirely systematic and rational. For one thing, the schedule

* These, and other technical expressions are exemplified at the end of the article.

had to be used by counters accustomed to the commoner varieties of terminology in traditional grammar. Everyone recognizes that there is gross confusion within this nomenclature, some terms indicating "form" (the word "simple", for example), and other parallel terms indicating "function" (the word "progressive" for example). Nevertheless this traditional nomenclature recognizes the need for a grammar which is both notional and formal. Recognizing only formal categories would have distinguished a vestigial sibilant suffix, and equated it as a form, with, for instance, the whole plain stem form ; and this, obviously, would have been futile for significant counting. Moreover, without notional divisions within the Simple Present tense form, the behaviour of certain groups of verbs with respect to Simple Past and Past Perfect usage would be inexplicable. The schedule shows how much its composer is indebted to Kruisinga ; much of the material derives from his work, "An English Grammar", Vol. I, Accidence and Syntax, a review of which appears elsewhere in this journal.

3. *Validity*

The validity of the figures from such a count depends on the organization of the work, on the care of the counters, on the ease or otherwise of clear attribution, and on the extent and amount of the material.

Two kinds of attribution are involved, formal and notional. The validity of the formal attributions must be high, as only recognition is involved. The notional attributions must be of more doubtful validity, as personal judgment is involved, and the language itself does not always recognize the compartments into which we try to accommodate it. For instance, context, and interpretation of context, may have to decide whether an entry is made under stem + -ed 1 (predicative participle of state) or 2 (past participle of occurrence). The validity of this kind of attribution depends on the capacity of the categories themselves, on the proportion of doubtful to secure attributions in the texts, and on the possibility of the counter's being biassed in his attributions of borderline cases. Particular effort was given to getting reasonably reliable attribution in the Simple Present and Past categories.

All the counters worked on all the material, so that every occurrence of a verb form in that material should have been noted and ascribed a place in the schedule. The material itself consisted of complete works except for *Chamber's Encyclopædia* of which four volumes were sampled, all occurrences being counted on every fiftieth page, as far as page 451, in each of the first four volumes.

The other materials have been: two novels, two plays, three books of a popular, factual nature, an issue of an English newspaper, and the conversational section of MacCarthy's "English Conversational Reader".

For convenience of comparison, each major section of the schedule (plain stem, etc.) has been analysed separately.

For interest, the figures for actual occurrences counted over the first two courses are given herewith:—

plain stem	18,707	or 36.7%
to + stem	4,875	9.6%
stem + -ed	21,120	41.5%
stem + -ing	6,199	12.2%
Total	50,901	100.0%

4. *Main Features*

The figures for the four main verb form categories show that the plain stem and stem + -ed forms together outnumber the to + stem and stem + -ing forms by almost four to one.

The plain stem

The most notable feature of plain stem usage is the relatively even distribution of low-frequency items, except for Simple Present Actual and Simple Present Neutral tenses.

The plain stem occurs after "can", "may", "shall", "will", "must", "could", "might", "should" and "would" to the extent of 18.4% of the total occurrences of the plain stem; but this percentage is the sum of over sixty distinct items, only five of which account for more than 1% each of plain stem occurrences. It may be worth noting that in this group, the preterite occurs more frequently than the present form.

The Simple Present tense form breakdown is interesting. This form is commonly taught in such a way that students believe that its characteristic use is to indicate habitual happenings, though with various groups of verbs forming exceptions. The count reveals that the use of this tense form to represent recurring (iterative) happenings (a slightly wider term than "habitual") accounts for only 5.3% of its occurrences. 56.9% of its occurrences apply to the "present" or "actual" moment, and 35.5% of its occurrences are in statements devoid of time-connotation, or "neutral".

Students who have been taught that the Simple Present form is characteristically associated with indication of habitual happenings probably associate indication of "present" time with the use of the Present Progressive form. Indeed, the opposition is frequently made :

Present Progressive — now
Simple Present — every day

It is interesting to note, therefore, that of the occasions in our material, on which "now" is the time-reference, the Simple Present Actual accounts for 95.1% of these occurrences, and the Present Progressive for 4.9%. It may be worth adding that the Simple Present form is used in future time contexts far more frequently than is the Present Progressive.

It is not to be concluded, necessarily, that our first teaching of Simple Present and Present Progressive ought to be different from what is current practice in many parts of the world. It cannot be denied, however, that those countries in which the Simple Present is the first tense to be taught and in which only a minor position is given to the Present Progressive (the U.S.S.R. and, I believe, Scandinavia) have a strong case on statistical grounds. Actual usage in this area might well be due for reconsideration : our impressions of what is "idiomatic" we have inherited from early practitioners of "oral methods" who were strongly influenced by their desire for colloquial and informal English to oust the written and formal variety. The Progressive tense forms may well be considered less formal than the Simple, but may not be, for that reason, more idiomatic.

to + stem

Over half of the occurrences of to + stem fall into the patterns : verb + to + stem, verb + noun/pronoun + to + stem, and verb + noun/pronoun + preposition group + to + stem.

Otherwise, there is a great deal of statistical suggestion with a negative import. Sentences of the pattern "I am to go", "You are not to . . ." account for 0.4% of occurrences. The pattern ". . . kind of you to . . ." accounts for 0.16%, and ". . . seems to have done . . ." for 0.04%.

The fairly strong representation of implication of meaning through form may be noted : non-fulfilment of an arrangement 3.8%, obligation 7.8%, and so on.

Stem + -ed

The predicative participate uses of the stem + -ed form, i.e., as participles of state and of occurrence account for 23.4% of total occurrences. It seems that these might be taught as adjective equivalents, to begin with, without much fear of misconceptions being established.

The Simple Past tense form presents as interesting a breakdown as the Simple Present form. The Narrative accounts for 53.1% of occurrences, the Actual for 29.0%, the Habitual for 3.0%, the neutral for 12.1%, and the Modal uses for 2.8%. In some books, the use of the Simple Past to give context-information exceeds its use as a narrative past; and its percentage of simple past occurrences is remarkably high. Frequently, teachers assign the task of giving circumstantial detail "at that time" to the Past Progressive. However, we have in English "While we waited . . ." as well as "While we were waiting. . . ." As has already been observed, in the schedule examples of Simple Past Actual quoted earlier in this article, it would have been possible for the author to have used the Past Progressive: "The sea was sounding . . .", etc. The statistical position is that the Simple Past was used on 91.7% of such occasions, and the Past Progressive on 8.3% of such occasions.

The Simple Past Neutral comes out as a very frequent Simple Past tense, surely an indication of the strong feeling, in English, for concord of tense form; this feature of English is not always appreciated by foreigners learning the language.

To this use of the preterite form on occasions which are neutral for time may be juxtaposed the modal occurrences which account for 2.8% of Past Simple occurrences, to give an indication of the extent to which the preterite appears in English in a formal way, that is, independently of any past time context.

The last feature of stem + -ed form usage to require comment is the remarkably low frequency of occurrence of all "Perfect" tenses. The "Perfect Present", most frequently the learners' introduction to this tense-form, accounts for 1.0% of stem + -ed occurrences (or 8.3% of total Perfect tenses). It will be seen that many of the tenses which occupy considerable teaching time are barely represented. (The Future Perfect Progressive appeared on *one* occasion!)

Stem + -ing

Whereas the "tense" sections under the plain stem form account for 67.1% of plain stem occurrences, and the tense sections under

the stem + -ed form account for 67.5% of stem + -ed occurrences, the tense sections under the stem + -ing form account for only 15.0% of stem + -ing occurrences.

A very frequent use of stem + -ing is as a free adjunct: "Returning from Hyderabad . . .". This use amounts to 18.2% of the total.

As an adjective in a noun group, stem + -ing has a total frequency of 18.6%, divided evenly between the two kinds of example (adjectives preceding and adjectives following the noun).

Stem + -ing appeared as a noun on 6.7% of occasions.

Comparison has already been made of Present and Past Simple and Progressive usage. A sampling frequency count, for these features only, produced altogether comparable figures for the various types of material. Those which may be of additional interest are of an Enid Blyton school tale, "Five on a hike together", and an English Primary School reader, "Janet and John", Book IV (all occurrences in this book were counted). In the school tale, occurrences referring to "now" were distributed 96.5% to Simple Present Actual, and 4.5% to Present Progressive. In the Reader, the percentages were 92.8% to Simple Present Actual, and 7.2% to Present Progressive. An examination of two English school books, namely "English with a purpose", Books 1 and 2, by R. Patterson gives for Simple Past Actual and Past Progressive "at that moment" 94.2% and 5.8% of occurrences respectively.

Mr. Peter MacCarthy's "English Conversational Reader" (the conversations only), was used in an attempt to see whether conversational English was markedly different from written English. Reference to "now" was made through Present Simple Actual to the extent of 91.66% and through Present Progressive to the extent of 8.34%.

Exemplification of technical terms

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

Simple Present Actual

This Broadcast *comes* from the B.B.C., London.
I *don't know*, and I *don't care*.
Do you *see* how well Mahmud is running (Present Progressive "now") ?
He *deserves* to win.

Simple Present Neutral	The Ganges *rises* in the Himalayas, *flows* across the Indian Plain, and *enters* the sea in the Bay of Bengal.
Simple Present Iterative	We *see* each other fairly often.
Predicative Participle of State	He was *delighted* to hear it.
Predicative Participle of Occurrence	It was *gone* in a flash.
Present Progressive "now"	What *are* you *doing* with that piece of paper ?
Past Progressive "at that time"	They *were* still *talking* when I left.
Simple Past Narrative	He *got* up, *went* to the door, *opened* it and *looked* out.
Simple Past Actual	The old woman paused (Narrative), her hand on the loaf of bread, her gaze out of the open door into the garden. 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*. In the middle there *was* an old salad bowl filled with yellow and red nasturtiums. She smiled. (Narrative.)
Simple Past Neutral	He told (Narrative) us all about the Heaviside layer, that it *reflected* radio waves. . . .
Simple Past Modal	It is time the question of speech, its standards and its dialects *was* raised above the level of mere prejudice.
Perfect Present	"Is (Actual) Tom in ?" "No, he's just *gone* out."