A BASIC MALAY VOCABULARY

by

JOHN LE PREVOST, B.A.

Consisting of the 1,000 most frequently used words in the written Malay language and a list of those words which occur at least once in every 1,000,000
Malay Word Count
INCLUDING 1,000-WORD LIST

A Report on a Word Frequency Count of the Written Malay Language

1952

Bilangan Perkataan, Bahasa Melayu

Penyata berkenaan dengan bilangan perkataan, Melayu dalam bahasa tulisan

by J. LE PREVOST, B.A. (Cantab.)
Colonial Education Service

UNIVERSITY OF LONDON PRESS LTD
WARWICK SQUARE, LONDON, E.C.4
FOREWORD

I commend this Word Frequency Count to all friends of the Malay Language, and especially to those engaged in studying it or teaching it. Mr. Le Prevost and his team of teachers are to be congratulated on their careful and patient work.

The making of a word frequency count is a highly technical matter and it would not be proper for me to make any comment on the machinery the team employed for the purpose; but we can rest assured that the spread of material was wide, and that the result is useful and appropriately accurate.

Mr. Le Prevost wisely limits his comments on the results to very broad principles. This investigation, as he suggests, opens up a field of fascinating research. Two of his suggestions I regard as of great importance. One of the next steps should be the illustration of the use of the higher frequency words in appropriate contexts. This is obviously a fruitful field of study for those Malay gentlemen who are at present interesting themselves in the fate of Malay writing. He also stresses the richness of the spoken language as against the apparent conservatism of the written language. A word count of the spoken language with the mechanical devices now available should not be too difficult. It would then be possible to find ways of enriching the written language with the virile material and idioms of the spoken language.

As Mr. Le Prevost says, language is a living thing. It is one of the social habits of its users. Those of us who are daily concerned with the Malay language realise how its character is changing with the impact of other languages and external ideas upon it. It is obvious that another count will be required in the course of time. A comparison of the second and first counts might reveal some startling changes and indicate the nature of development, and help to diagnose where help was needed in preserving some of the natural characteristics of the language.

I am proud that this first step has been taken in the technical study of the Malay language by members of the staff of the Department of Education, Singapore. They are blazing a trail parallel to that upon which have journeyed such figures in the linguistic world as Thorne, Dike, Faucett, Itsu Maki, Palmer, West, Hornby, Eaton and Bongers. To me it is significant that the burden of this work was carried by twenty-eight teachers in Malay schools. The future of Malay should be safe if we can find such patience and detailed accuracy among its teachers.

A. W. FRISBY
Director of Education,
Singapore
PENDAHULUAN

Saya aku'i kebaikan daftar bilangan perkataan Melayu ini kapada sakalian tuan yang kasehkan bahasa Melayu, terutama kapada mereka yang mempelajari atau mengajarkan-nya. Tuan Le Prevost dengan kumpulan guru-nya boleh benar di-puji kerana pekerjaan-nya yang chermat dan sabar itu.

Membuat daftar bilangan perkataan ini ada-lah satu pekerjaan yang tinggi technique-nya; kerana itu tidak-lah sayoga-nya saya mesti memberi apa-ulausan berkenaan dengan chara-nya yang telah di-gunakan oleh kumpulan itu; tetapi satu perkara yang boleh kita aku'i ia-itu jenis bahan yang di-ambil dan di-jadikan kandongan daftar itu ada-lah luas; yang demikian di-perchaya' di daftar itu berguna dan betul.


Sa-bagaimana kata Tuan Le Prevost dalam penyata-nya, bahasa itu ada-lah satu perkara yang hidup. Sa-benar-nya, bahasa itu memang terhitong satu daripada beberapa alat mashrakat bagi orang yang menggunakan-nya. Kami sakalian yang berhubong
sa-tiap hari dengan Bahasa Melayu, terasa akan perubahan bahasa itu, berubah sedikit demi sedikit di-sebabkan oleh pengaruh bahasa asing dan fikiran luar yang mengena’i bahasa itu. Yang demikian nyata-lah bahawa di-kehendaki satu lagi daftar bilangan perkataan pada suatu masa yang akan datang. Dengan jalan membandingkan kumpulan perkataan yang pertama dan yang kedua dalam daftar itu, akan terbokti beberapa perubahan yang mengherankan; bahkan akan ternyata juga bagaimana chara-nya kerja itu telah di-buat dan daripada-nya juga akan dapat kita memereksa apa yang kurang yang patut di-baiki supaya terpelihara sa-tengah daripada sa-tengah lenggok bahasa yang asli itu.


A. W. FRISBY
Penguasa Pelajaran
Singapura
ACKNOWLEDGMENTS

I wish to record the patience and the cheerful manner in which the twenty-eight teachers toiled to assist me in the preparation of this report.

I wish to thank Che Mohd. Arif b. Ahmad through whose generosity we were able to count the necessary words occurring in his personal correspondence files.

In particular I would like to pay tribute to the care and interest with which my colleague Che Abdul Karim b. Mohd. Shariff undertook the translation of Part I.

Lastly I wish to thank the Hon. Mr. A. W. Frisby, Director of Education, Singapore, who has encouraged me in this work and has consented to write a Foreword to this report.

J. Le P.

Singapore

February 1953
PENGAKUAN TERIMA KASEH

Saya suka merakamkan perasaan saya di-sini terhadap kapada sabar dan gembira-nya 28 orang guru Melayu yang telah bekerja dengan bersusah payah menolong saya menyediakan penyata ini.

Saya uchapkan terima kaseh kapada Che' Md. Arif bin Ahmad, kerana dengan kemurahan hati-nya telah menyebabkan kita dapat membilang perkataan mustahak yang di-ambil daripada surat kiriman-nya sendiri.

Khas-nya, saya suka menyatakan pujian saya di-sini kapada rakan sa-jawat saya Che’ Abdul Karim bin Mohd. Shariff yang telah mengambil berat dan berhati menterjemahkan Bahagian I itu.

Penghabisan sakali saya sukachita menguchapkan terima kaseh kapada Yang Berhormat Tuan A. W. Frisby, Penguasa Pelajaran, Singapura, yang telah memberangsangkan saya dalam kerja mengumpulkan perkataan ini dan telah sudi menuliskan penda­huluan penyata ini.

J. Le P.

Singapore

February 1953
CONTENTS

Foreword 5
Pendahuluan 7

Acknowledgments 9
Pengakuan terima kaseh 10

PART I
Report on motives, methods and results 14
Penyata tentang sebab, chara dan faedah-nya 22

PART II
Words occurring at least once per 1,000,000 words
Bilangan kali sa-suatu perkataan itu di-jumpai dalam 1 juta perkataan 36

PART III
The 500 words occurring most frequently and the 500 words occurring next most frequently
500 perkataan yang pertama dan 500 yang kedua 64

APPENDICES
A List of materials used
Daftar segala bahan yang di-gunakan 72
B List of names of teachers who took part
Daftar nama guru yang mengambil bahagian 77
C Report pro forma used during the Count
Penyata berkenaan dengan kemajuan bilangan itu 78
REPORT ON MOTIVES, METHODS AND RESULTS

This is, I believe, the first comprehensive word frequency count made in the Malay language. It is a frequency count of the written language as represented broadly over different forms of written material.

To enable students to assess the value of its findings, this section of the report will contain a detailed account of the manner in which the project was conducted.

The section will also indicate the motives which inspired the count and a provisional assessment of its success and usefulness.

Motives

Language is a living part of the people who use it. As they themselves change, so their language must change. To understand a people in their fullness, an outsider must understand their language. To understand each other within a community, each person must have a thorough grasp of the common language. The administrator, the teacher, the leader, these three who seek to achieve good government, wisdom and unity of purpose know full well the potency of a rich and adequate language, a language which will satisfy the social-economic, the political and the cultural needs of the community, both in writing and in speech.

These problems express themselves in commonplaces such as "the official language", "the language of citizenship", "a common language", "the materials and techniques of language teaching". They are uppermost in the minds of most responsible people in this country when they consider the problems of intercommunication and administration of a polyglot community, child education, adult education, a national language, a lingua franca and other aspects of these vital problems within such a community.

It was as an initial step, in the study of the Malay language and in the satisfaction of the four immediate needs of our community, that this task of counting more than 22 million Malay words was undertaken. The needs as I saw them were:

(a) the social and political need to teach Malay quickly and efficiently to the other races that live and work in Malaya and Singapore.

(b) the need for the production of a large amount of adequate, locally written and graded reading material for child and adult schools.

(c) the need for research into the vocabulary of the Malay language, to determine and develop its capacity to express the needs of a nation which is rapidly growing more literate and more conscious of the demands of modern technological development.
the need to improve the techniques of language teaching in all Malayan schools.

I know it is the heart-felt desire of the Malays themselves to preserve their culture and their language. I believe that this cannot be done by refusing to face actuality. I believe too that they may achieve the very opposite by so doing. Unwittingly they may kill what they seek to preserve. I have seen educated Malays gradually lapse into less pure Malay forms of language under the continuous pressure of communication with other races whose knowledge of Malay is imperfect. In matters of language, the tendencies which lead towards an acceptance of less pure forms are always very strong. The only way to resist them is to study the language and perfect the means of teaching that language quickly and well to other races, so that the language itself grows more vital and flexible in its use. Malay will grow in strength as more people learn and study it. History clearly shows us that languages reserved to the few and the select die out. It is not my purpose here to forecast the future of the Malay language; but, having completed this word count, I am more than ever convinced that the future magnificence of that language rests in the hands of the Malay scholars and teachers and in their ability to toil at the detailed study and research of every available source of material, written and spoken, in their own language. Then they have a large field to sow, for the Malay language is already partially, though imperfectly, known to very many outside their own race.

It was this sense of urgent mission, a love for their own language and a desire to make it stronger and more widely known, that sustained the volunteer Malay "teacher-counters" through the long arduous months of counting which this report records. It was all carried out in their own time and they kept rigidly to the timetable set. That their reward will be the enrichment and enhanced importance of their own language is their earnest hope.

**METHOD**

On July 21st, 1951, the twenty-two teachers, whose names are recorded in Appendix B, met me to discuss ways and means. They had been selected from three big schools and were arranged in five groups, each group with a teacher of long experience to act as a supervisor. After the task had been outlined to them, they agreed to carry it out. This agreement was made with a readiness which continued undiminished throughout the holidays and working days alike of the next nine months.

Each counter had to read 1,000,000 words. This total was to be divided up proportionately as set out below, to cover the whole spread of Malay written material, both in Jawi and Rumi. The counter was given a section of the alphabet and adopted these letters as the initial letters of the words for which he was to search. He was issued with a large ledger and asked to record all words found, the total of words read under their respective types and the materials used.