



COMPENDIUM  
OF THE  
WORLD'S  
LANGUAGES

VOLUME I  
ABAZA TO LUSATIAN

GEORGE L CAMPBELL

2  
Francis & Taylor  
C1522

George Campbell  
1963

CONTENTS

# COMPENDIUM OF THE WORLD'S LANGUAGES

---

BY  
**GEORGE L. CAMPBELL**



LONDON AND NEW YORK

410.202  
C152C

General Slavic Ref.  
(Slavic)

V. 1

copy 2

# CONTENTS

## VOLUME I

Introduction  
vii

Abbreviations  
xiii

List of languages  
xv

COMPENDIUM OF THE WORLD'S LANGUAGES  
ABAZA TO LUSATIAN  
1

## VOLUME II

COMPENDIUM OF THE WORLD'S LANGUAGES  
MAASAI TO ZUNI  
829

Scripts  
1509

Bibliography  
1561

Ref.

---

## INTRODUCTION

'Les langues imparfaites,' complained Mallarmé, 'en cela que plusieurs, manque la suprême.' ('Languages [are] imperfect, in that there are many [of them], [but] the supreme [one] is lacking.')

(*Variations sur un sujet: Crise de Vers.*) In some ways, this is a puzzling remark. Exactly what does *plusieurs* refer to? If it is to the multiplicity of natural languages, there need be no disagreement, on this point at least. Languages have indeed proliferated on the planet; the most recent and most comprehensive count (Merritt Ruhlen, 1988) records around five thousand languages which are or have been spoken, some by many millions, most by several thousands, and some by only a few hundred people. We have no way of gauging how many languages, once existent, have now vanished without trace. It seems unlikely, however, that Mallarmé had made first-hand acquaintance with all or even many of the natural languages, or had been able to assess in what degree they fell short of perfection. One might have expected him to find much to his taste in Classical Chinese or Classical Arabic poetry. Moreover, how and why could or should one natural language be 'supreme'.

The sentence begins perhaps to make more sense if we take *plusieurs* to refer to *levels* of language, particularly since Mallarmé goes on to deplore the *numéraire facile* – the 'easy currency' – of everyday speech, which is non-convertible into *la notion pure*. Given this non-convertibility, it is the poet's job to transcend natural language and communication by providing what Mallarmé calls *un complément supérieur* – a 'higher complement'. In Mallarmé's hands, this higher integration functions in a realm of autonomous verbal creativity, yielding a kind of cerebral algebra of images, hardly if at all anchored in the world of everyday experience and communication. It is a construct which we acclaim as arcane and beautiful poetry.

But the *numéraire facile* rejected by Mallarmé as 'imperfect' represents something even more mysterious and far more valuable: the original mapping of the shared human environment and shared human experience onto the plane of speech. This construct is initiated and sustained by an innate human competence, whose nature we can hardly even guess at. Key factors in the mapping are the remembered use of words, and their division into constants which can be transposed, and operators and operational codes which cannot.

The 'easy currency' has enabled man *qua* Robinson Crusoe to plot his position, catalogue his possessions, and tabulate his actions. Far from being divorced from natural experience, natural language might be defined as the audible inventory of that experience, and words are its bench-marks. In contrast, the conscious creation of alternative and idiosyncratic worlds has to do

rather with a secondary mimetic mapping from the spoken to the written plane. This was perfectly clear to Mallarmé himself: 'Tout, au monde, existe pour aboutir à un livre.' (Everything, in the world, exists to end up in/as a book.) (*Variations sur un sujet: Quant au Livre.*)

The natural inventory is not, of course, entirely homogeneous. There are local nuances of modality and circumstance. This or that facet of a nominal or verbal event is stressed in one language, ignored in another. Thus, where Indo-European languages tend to use adjectival and adverbial modifiers, in many American Indian languages such information inheres in the verb stem itself. In verbal systems, some languages adopt the standpoint of temporal sequence – tense, that is – while others prefer to make an aspectual distinction between completed and uncompleted action. But these are surface variations. From the Egyptian, Sumerian, and Archaic Chinese of 3000 BC to modern Russian, German, and Arabic, from Old Norse and Sanskrit to Navajo and Quechua, there is a general consensus on the fundamental parameters of existence: the spatio-temporal matrix, and the kinetic operational paradigms which underlie the surface diversity. *Vis-à-vis* Mallarmé's *notion pure* and *complément supérieur*, the world's natural languages may indeed be askew; but they are orthogonalities for their users, and none is superfluous. As Sir Thomas Browne said: 'There are no Grotesques in Nature, not anything framed to fill up empty Cantons and unnecessary spaces.'

The construct yielded by the primary mapping is not immutable. Over long periods of time, phonological and morphological systems are subject to superficial erosion and replacement. Declensions are reduced, class and gender taxonomies are amended, verbal structures simplified. Essential linguistic profiles do not seem to be affected. The relative brevity of the written record, with a cut-off point at around the third millennium BC, limits our perspectives here; extrapolation is risky. But it is difficult to imagine that Chinese was, or ever could be, 'like' Sanskrit, or Hawaiian like Ket, though the same 'generative template' underlies all four.

Innovation where it occurs is largely a surface phenomenon, involving transposition of constants, and local differentiation. Like Mallarmé, the great Tang poet Tu Fu can score brilliant aesthetic effects by transposing semantic nuclei – in Chinese, a far more radical step than it is in French (see, for example, lines 5–6 of *Wàng yuè*, 'On a Prospect of T'ai-Shan', in David Hawkes (1967) *A Little Primer of Tu Fu*, Oxford University Press). Arabic, on the other hand, prefers to make novel, not to say unheard-of, utterances via a cerebral play upon words, and a kind of phonological mirror-imagery (see, for example, the *Maqāmāt*, the 'Assemblies', of the eleventh-century poet al-Ḥarīrī, especially Assembly 32). But essential linguistic profiles are not affected by such innovations; al-Ḥarīrī could no more insert the Arabic adjective between the article and the noun than Tu Fu could mark his verbs for gender, number, and person. If the surface diversity of languages is the main subject of this book, the underlying consensus emerges, it is hoped, as a corollary.

For reasons of space, certain other properties of language receive less than the attention due to them. For example, in a sense (not that of Mallarmé), words

carry their own *complément supérieur* in the form of a mythopoeic potential or charge, which seems to be latent in them from the outset, and which is activated in literary creativity, and, most particularly, in any language which is the vehicle of a religion or taboo system. Words resonate beyond themselves into more or less adjacent semantic fields. Hugo van Hofmannsthal in his wisdom put it beautifully: 'und dennoch sagt der viel, der "Abend" sagt' ('and yet he says much, who says "evening"'); *Ballade des äusseren Lebens*).

The first eight verses of Chapter 1 of St John's Gospel, which, where available, follow each article, serve not only to exemplify the language in question, but also to show how readily co-ordinates can be tacitly transformed, so that spatio-temporal markers that are meaningful in one frame of reference appear to remain semantically invariant, that is, meaningful, in another: 'in the beginning', 'was the Word', 'with God'.

The late twentieth century has witnessed an increasing tendency for both primary and secondary mappings to cede status at least to a derivatory mapping – their instantaneous and infinite electronic multiplications, with or without visual display: 'I heard it on TV'. It is too early to assess the full implications of our reliance on images but we all know what happened to the Lady of Shalott when she was suddenly deprived of them.

The book as it now appears differs considerably from that originally planned. To begin with, publisher and compiler envisaged an alphabetical list of between five and six hundred languages, each entry to contain the relevant statistical information along with a note on genetic status and a phonological and morphological profile. It soon became apparent, however, that while this format was adequate for geographical, demographic and genetic detail, little of interest could be said, within such constraints, on precisely what it is that differentiates one language from another. Thereupon, phonological and morphological sections began to expand, full or partial paradigms were introduced; and, as the articles lengthened from two or three paragraphs to several pages, the list of entries shrank accordingly. Thus, for example, *Chontal*, *Nukuoro* and *Osage*, present on the original list, are now absent, while the Procrustean paragraph originally allotted to *Abkhaz* has grown to four, still inadequate, pages, and *Chinese* has spread from one page to twelve.

Herein, however, or such is my hope, lies any real usefulness the book may have. For, while it is not difficult for someone interested in a given language to find the relevant geographical and demographic facts concerning it, it is not always so easy to find a simple account in English of how it actually works – of its nuts and bolts. There is, of course, no shortage of books on the major languages of the West, or on the more important Oriental languages such as *Chinese*, *Japanese* and *Arabic*; and today's students are fortunate in having ready to hand such invaluable exotica as Hewitt's *Abkhaz*, Saltarelli's *Basque* and Aronson's *Georgian*, to name only three of the best examples. But the quixotically curious – those who might turn to a single reference work in the hope of finding out at one and the same time how the *Navajo* verb and the *Andi* nominal system work, why vowel harmony has been eroded in *Uzbek*, and what sort of script the

*Cambodians* use – are not so well catered for. It has been my hope and intention that this book should provide basic guidance in simple language, and point the reader, via the bibliography, towards the more competent and more detailed accounts, to which I gratefully acknowledge my indebtedness.

In general, the articles follow the same lay-out, sub-divided as follows:

1. *Head-word*: background note on the language, its affiliation, location and number of speakers; dialects; where relevant, some remarks on the literature of which it is the vehicle;
2. *Script*: if other than Roman, the script used by a language is described and displayed in a chart at the end of the book;
3. *Phonology*: for the most part, the phonological inventories are set out in terms of mode of utterance; that is, consonants are sub-divided into stops, affricates, fricatives, nasals, laterals, semi-vowels; vowels into oral and nasal series. In a few special cases, for example *Sanskrit*, a positional grid is used, showing the phonemes of the language in terms of labial, palatal, dental, retroflex, and velar series. This positional grid is also retained in the case of certain *Caucasian* languages, for example, which make a phonemic distinction between uvular, pharyngeal, and glottal series. Tone, if relevant, is also described here.

Wherever possible, phonological inventories are in IPA symbols. Exceptions to this general rule are provided by languages, long dead, whose phonological values are conjectural, and by languages such as *Andamanese* or *Chimu* for which no IPA values seem to be available. Transcription is broad.

4. *Morphology and syntax*: the main sub-headings are: article; noun; adjective; pronoun; numerals; verb; pre/postpositions; word order. For reasons of space, only very general questions of syntax are taken up.
5. *Illustrative text*: verses 1 to 8 of the first chapter of St John's Gospel have been chosen as a suitable example, available in most of the languages described. Alternative biblical specimens are substituted, and identified, in cases where a translation of St John's Gospel is not available.

### Selection and classification

My aim was to include all of the world's literary languages, along with certain other languages which, though lacking a written literature, were nevertheless felt to be of sufficient interest and importance to warrant inclusion. The list of 1,000 languages, given by David Crystal (1988) on pp. 436–44 of his *Encyclopaedia of Language*, served as a general base, extended where necessary by specialist works such as Bernard Comrie (1981), *The Languages of the Soviet Union*. In Crystal's list, the cut-off point for number of speakers is 10,000. I have included several languages with much smaller, even minimal tallies, e.g. certain members of the *Caucasian*, *Palaeo-Siberian* and *North American Indian* groupings.

### Classification

Here again, I follow David Crystal's middle course between 'lumpers' – linguists who seek to identify genetic or at least typological relationships connecting ever wider and more numerous groups of languages – and 'splitters' – linguists who, on the contrary, sub-divide large putative groupings into smaller well-defined and demonstrably coherent units. Thus, the *Semito-Hamitic* languages are described as forming part of the *Afro-Asiatic* family, and *Bantu* as *Niger-Congo*; but *Indo-European* and *Japanese* are not classified as '*Euroasiatic*', nor are *Quechua* and *Cree* lumped together as '*Amerind*'. Nor is there, perhaps mistakenly, any mention of '*Nostratic*' or '*Austriac*'.

### Sources

The books which have been used in the preparation of this work are listed in the Bibliography. Examples in the text, mainly in the section on Morphology and Syntax, fall into two categories. In the case of languages with which I have worked over the years, examples are drawn partly from the standard works listed in the Bibliography, partly from other sources. In the case of languages with which I am not actively familiar, all examples are taken from the works listed. Thus, for example, all *Ainu* examples are from Refsing (1986); the *Nama* examples are from Hagman (1977); *Bambara* from Brauner (1974); *Samaritan* from Vil'sker (1974), and so on. I am especially indebted to two Russian collective works: *Jazyki Narodov SSSR* and *Jazyki Azii i Afriki*.

### Permissions

The Mayan glyphs on p. 911 are reproduced from *Maya Treasures of an Ancient Civilization*, Harry N. Abrams and Albuquerque Museum, New York, 1985, redrawn by Dolona Roberts from *The Ancient Maya*, 4th edn, Sylvanus G. Morley and George W. Brainerd, revised by Robert J. Sharer, with the permission of the publishers, Stanford University Press, © 1946, 1947, 1956, 1983, by the Board of Trustees of the Leland Stanford Junior University.

Bishop Landau's Mayan alphabet in the Appendix of Scripts is reproduced from Figure 60 of *Deciphering the Maya Script*, David H. Kelley, University of Texas Press, © 1976, by permission of the publisher.

In conclusion, I wish to express my gratitude to all those who have helped to make this book possible: to Wendy Morris for initiating the project, and to Jonathan Price and his colleagues at Routledge – Mina Patria, Shan Millie, Emma Waghorn, Nigel Marsh, Susie Hilsdon, Judith Watts – for seeing it through its various stages; to Jenny Potts, for her meticulous copy-editing; to Richard Caesar, Simone Osborn, and the skilled typesetters of Computype for coping so beautifully with a most forbidding manuscript; to Monotype for their



## INTRODUCTION

---

co-operation in providing characters for the Appendix of Scripts; my wife for providing the bibliography; and to Alan Jesson and Gerry Bye of Cambridge University Library, who supplied the New Testament specimen passages appended to most of the articles.

G.L.C.