

PACIFIC LINGUISTICS

Series D - No. 37.

A GRAMMAR OF TOBA-BATAK

by

P.W.J. Nababan

(MATERIALS IN LANGUAGES OF INDONESIA, No.6.)

W.A.L. Stokhof, Series Editor



Department of Linguistics
Research School of Pacific Studies
THE AUSTRALIAN NATIONAL UNIVERSITY

PACIFIC LINGUISTICS is issued through the Linguistic
Circle of Canberra and consists of four series:

SERIES A - Occasional Papers
SERIES B - Monographs
SERIES C - Books
SERIES D - Special Publications

EDITOR: S.A. Wurm

ASSOCIATE EDITORS: D.C. Laycock, C.L. Voorhoeve, D.T. Tryon, T.E. Dutton

EDITORIAL ADVISERS:

B.W. Bender
University of Hawaii
David Bradley
University of Melbourne
A. Capell
University of Sydney
S.H. Elbert
University of Hawaii
K.J. Franklin
Summer Institute of Linguistics
W.W. Glover
Summer Institute of Linguistics
G.W. Grace
University of Hawaii
M.A.K. Halliday
University of Sydney
A. Healey
Summer Institute of Linguistics
L.A. Hercus
Australian National University
Nguyễn Đăng Liêm
University of Hawaii

John Lynch
University of Papua New Guinea
K.A. McElhanon
University of Texas
H.P. McKaughan
University of Hawaii
P. Mühlhäusler
Linacre College, Oxford
G.N. O'Grady
University of Victoria, B.C.
A.K. Pawley
University of Auckland
K.L. Pike University of Michigan;
Summer Institute of Linguistics
E.C. Polomé
University of Texas
Gillian Sankoff
University of Pennsylvania
W.A.L. Stokhof National Center for
Language Development, Jakarta;
University of Leiden
J.W.M. Verhaar
Gonzaga University, Spokane

All correspondence concerning PACIFIC LINGUISTICS, including
orders and subscriptions, should be addressed to:

The Secretary
PACIFIC LINGUISTICS
Department of Linguistics
Research School of Pacific Studies
The Australian National University
Canberra, A.C.T. 2600
Australia.

Copyright © The Author.

First Published 1981.

The editors are indebted to the Australian National University for assistance in
the production of this series.

This publication was made possible by an initial grant from the Hunter Douglas
Fund.

National Library of Australia Card Number and ISBN 0 85883 254 2

TABLE OF CONTENTS

| | |
|--|---------------------|
| Introduction | <i>Page</i> viii |
| Introduction to the Thesis (1958): Chapters on Phonology | xvi |
| Introduction to the Dissertation (1966): From Morphophonemics to Syntax | xx |
| Maps | xxiii |
| CHAPTER I - THE BATAK PHONEMES | 1 |
| 1. Consonants | 1 |
| 1.1. The Stops | 2 |
| 1.2. The Fricatives | 3 |
| 1.3. The Affricates | 5 |
| 1.4. The Nasals | 6 |
| 1.5. The Linguals | 7 |
| 1.6. The Semivowels | 8 |
| 1.7. The Glottal Stop | 9 |
| 1.8. Concluding Remarks | 11 |
| 2. Vowels | 13 |
| 2.1. The Suprasegmental Phonemes | 21 |
| 3. Stress | 21 |
| 4. Juncture | 29 |
| 5. Pitch | 33 |
| CHAPTER II - PHONOTACTICS | 41 |
| 2.1. Word Patterns | 41 |
| 2.2. Consonant Clusters | 43 |
| 2.3. Vowel Clusters | 50 |
| CHAPTER III - MORPHOPHONEMICS | 57 |
| 3.1. Words | 57 |

| | <i>Page</i> |
|--|-------------|
| 3.2. External Sandhi | 58 |
| 3.2.1. Examples of External Sandhi | 59 |
| 3.3. Internal Morphophonemics | 60 |
| 3.3.1. Regular Internal Alternations | 60 |
| 3.3.2. Irregular Internal Alternations | 61 |
| 3.3.3. Stress Alternation | 63 |
| 3.3.4. Examples | 63 |
| 3.4. Conclusion | 64 |
| CHAPTER IV - MORPHOLOGY | 65 |
| 4.0. Introduction | 65 |
| 4.1. Parts of Speech | 65 |
| 4.2. Inflectional Categories | 66 |
| 4.2.1. Substantive Class | 66 |
| 4.2.2. Substantive Reference | 66 |
| 4.2.2.1. Allocation | 67 |
| 4.2.2.2. Actor | 67 |
| 4.2.2.3. Object | 68 |
| 4.2.3. ' Voice and Resolution | 68 |
| 4.2.3.1. Voice | 68 |
| 4.2.4. Mode and Aspect | 68 |
| 4.2.4.1. Mode: Active Transitive | 68 |
| 4.2.4.2. Passive Transitive | 69 |
| 4.2.4.3. Aspect | 70 |
| 4.2.5. Degree | 71 |
| 4.2.6. Ordinal | 72 |
| 4.3. Contentives | 72 |
| 4.3.1. Nouns | 72 |
| 4.3.2. Verbs | 73 |
| 4.3.3. Adjectives | 75 |
| 4.3.4. Adjuncts | 75 |
| 4.3.5. Interjections | 76 |
| 4.4. Functors | 76 |
| 4.4.1. Substitutes | 76 |
| 4.4.1.1. Pronouns | 77 |
| 4.4.1.2. Proverbs | 79 |
| 4.4.1.3. Pro-adjective | 79 |
| 4.4.1.4. Pro-adjuncts | 79 |
| 4.4.1.5. Pro-clauses | 80 |
| 4.4.2. Directive Particles | 80 |
| 4.4.2.1. Prepositions | 80 |

| | <i>Page</i> |
|--|-------------|
| 4.4.2.1.1. Locative Prepositions | 80 |
| 4.4.2.1.2. Instrumental Prepositions | 81 |
| 4.4.2.1.3. Attributive Preposition | 81 |
| 4.4.2.1.4. Possessive-Agentive Preposition | 81 |
| 4.4.2.1.5. Objective Preposition | 81 |
| 4.4.2.1.6. Comparative Prepositions | 82 |
| 4.4.2.1.7. Resultative Prepositions | 82 |
| 4.4.2.1.8. Some Phrases Function as Prepositions | 82 |
| 4.4.2.2. Subordinating Conjunctions | 82 |
| 4.4.3. Modal Particles | 84 |
| 4.4.3.1. Predicative Particles | 84 |
| 4.4.3.2. Imperative Particles | 85 |
| 4.4.4. Connectives | 86 |
| 4.4.5. Auxiliaries | 86 |
| 4.4.5.1. Predicator Auxiliaries | 86 |
| 4.4.5.2. Noun Auxiliaries | 87 |
| 4.4.5.3. Adjective Auxiliaries | 88 |
| 4.4.5.4. Adjunct Auxiliary | 88 |
| 4.4.5.5. Substitute Auxiliary | 89 |
| 4.4.5.5.1. Numeral Auxiliaries | 89 |
| 4.5. Structure of Words | 90 |
| 4.5.1. Composition | 90 |
| 4.6. Structure of Stems | 91 |
| 4.6.1. Secondary Derivation | 91 |
| 4.6.2. Primary Derivation | 91 |
| 4.6.2.1. Root Compounds | 92 |
| 4.6.3. Noun Stems | 93 |
| 4.6.3.1. Simple Noun Stems | 93 |
| 4.6.3.2. Primary Derived Noun Stems | 93 |
| 4.6.3.3. Secondary Derived Noun Stems | 95 |
| 4.6.4. Verb Stems | 97 |
| 4.6.4.1. Intransitive Verb Stems | 97 |
| 4.6.4.1.1. Simple Intransitive Verbs | 97 |
| 4.6.4.1.2. Primary Derived Intransitive Verbs | 97 |
| 4.6.4.1.3. Secondary Intransitive Verbs | 98 |
| 4.6.4.2. Transitive Verbs | 101 |
| 4.6.4.2.1. Simple Transitive Verbs | 102 |
| 4.6.4.2.2. Primary Transitive Verbs | 102 |
| 4.6.4.2.3. Secondary Transitive Verbs | 102 |
| 4.6.5. Stems of other Word Classes | 104 |
| 4.7. Roots | 104 |

| | <i>Page</i> |
|--|-------------|
| CHAPTER V - SYNTAX | 107 |
| 5.0. Introduction | 107 |
| 5.1. Constructions | 107 |
| 5.1.1. Endocentric Constructions | 107 |
| 5.1.1.1. Attributive Constructions | 107 |
| 5.1.1.1.1. Noun Phrase | 108 |
| 5.1.1.1.2. Verb Phrases | 109 |
| 5.1.1.1.3. Adjective Phrases | 110 |
| 5.1.1.1.4. Adjunct Phrases | 110 |
| 5.1.1.1.5. Substitute Phrases | 111 |
| 5.1.1.2. Co-ordinate Constructions | 111 |
| 5.1.1.2.1. Additive Constructions | 111 |
| 5.1.1.2.2. Alternative Constructions | 111 |
| 5.1.1.2.3. Appositive Constructions | 112 |
| 5.1.2. Exocentric Constructions | 112 |
| 5.1.2.1. Directive Constructions | 112 |
| 5.1.2.1.1. Marked Directive Constructions | 112 |
| 5.1.2.1.2. Verbal Directive Constructions | 112 |
| 5.1.2.2. Predicative Constructions | 113 |
| 5.1.2.2.1. Verbal Predicative Constructions | 113 |
| 5.1.2.2.2. Equational Constructions | 114 |
| 5.2. Clauses | 114 |
| 5.2.1. Word Order | 115 |
| 5.2.1.1. Direct Order | 115 |
| 5.2.1.2. Inverted Order | 116 |
| 5.2.2. Clause Types | 117 |
| 5.2.2.1. Favourite Types | 117 |
| 5.2.2.2. Minor Types | 118 |
| 5.3. Sentences | 119 |
| Appendix I - SYNTACTIC ANALYSIS OF A TEXT | 121 |
| 1. Introduction | 121 |
| 2. Text | 121 |
| 3. Free Translation | 122 |
| 4. Analysis | 123 |
| Appendix II - SPECTROGRAPHIC ANALYSIS OF SOME BATAK PHONEMES | 133 |
| 1. Vowel Length | 133 |
| 2. Vowel Formants | 135 |
| 3. Junctures | 136 |

| | <i>Page</i> |
|--|-------------|
| 4. Pitch Contours | 137 |
| 5. Phonotactics | 138 |
| <i>Appendix III - THE BATAK ALPHABET</i> | 141 |
| <i>Bibliography</i> | 145 |

INTRODUCTION

1. It presents something of a problem to write an introduction to one's thesis and dissertation after a lapse of 22 and 14 years respectively. However, in addition to the aspect of nostalgia not much different from revisiting the town of one's childhood, it also has an evaluative aspect which gives one a new perspective of things remembered as being different from their present appearance.

The thesis was researched and written in 1957 and 1958, the culminating period of structuralism in the United States, on the eve of the linguistic upheaval caused by the transformational-generative approach; Chomsky presented his model on a wider scale at the Texas Conference on Linguistics, in Austin in 1958. Naturally, the linguistic theory underlying the thesis was structuralism.

The dissertation was researched in 1964-1965 and written up in the latter half of 1965. The method and emphasis in the analysis naturally reflect the (prevailing) spirit of the times, namely (1) the view of language as a system governed by rules that can be presented as an ordered set, and (2) the ground-gaining view that the sentence is not the upper limit of linguistic description. The first view was intimately related to the interest of linguists in the exploration of the use of computers in the processing of linguistic data and the application of algorithmic computer techniques to the understanding of thinking. The second view was related to the theory that the study of phonology and morphology had reached a stable stage and the paradigms used had attained a degree of ease in their application. It was realised that the broader study of syntax would necessarily take one to the study of meaning and the paraphrase. This involved not only the hierarchical and constructional relations between the elements of a sentence, i.e. the constituent structure of the sentence, but also the horizontal relations of sentences that represent or realise the same

semantic network, in other words, express the same meaning. This led the generativists in the 1960s to posit a 'base component' of the syntactic structure from which the surface or actual form of the sentence could be 'derived'; among the other terms used (by different people at different times) for a similar (but different) concept were 'deep structure', 'semantic network', 'conceptual structure', 'underlying structure'.

2. The chapter on Phonology, Chapter 1, is a reproduction of some chapters of my M.A. thesis, entitled "A phonemic analysis of Batak and its bearing on the teaching of English pronunciation to a speaker of Batak", completed at the University of Texas, May 1958. The chapters dealing with the contrastive analysis of Batak and English and with its pedagogical implications are left out; however, the appendix describing the results of the spectrographic analysis of the Batak vowels and of some junctural features are included.

The technique used in identifying the phonemes is to find out whether two minimal sound segments are in contrast in identical or similar environments. Whenever only one or no identical environments are found, the process becomes somewhat cumbersome at times. For example, in the case of /ɔ/ and /o/ and that of /e/ and /ɛ/, presented on pp.13 ff, and in diagrammatic form on p.16, the argument appears overly long but does look fairly conclusive.

The case of the glottal stop is also quite complex and merits a lengthy treatment (pp.9-11). The conclusion reached contradicts the finding of other researchers (e.g. Percival 1964) who say that the glottal stop is a separate phoneme. In my analysis, the glottal stop is an allophone of /k/ in its obligatory occurrence preceding a voiced consonant; it is not recognised as an allophone in its optional occurrence as a vowel onset at the beginning of an utterance and a vowel coda at the end of an utterance.

The internal juncture (p.31-33) has also been given a good deal of space as this linguistic phenomenon is difficult to analyse in Batak (as in any other language). It is partly a tempo phenomenon but the other features of sound modification (onset, length, stress, etc.) seem to be more important, resulting in a complicated network of interrelations. Therefore, the discussion on the plus juncture ends in a somewhat inconclusive tone. At present, I would more definitely give a much smaller role to the pausal or break aspect of the plus juncture, because it is in fact a 'potential' rather than an 'actual' break in between two phonemes. Nevertheless, it still remains a stubborn feature in the description of Batak phonology.

The phonotactic description of the Batak sound system in which several sequences are represented by only one example (see Diagram on p.49) indicates the changing character of the Batak language through its contact with bahasa Indonesia, the national language of Indonesia. Since the completion of the thesis (1958), I have heard other novel sequences of phonemes which now seem to be established too and also more examples of the sequences described as unique (i.e. single occurrence).

3. In the 1950s and 1960s many scientists took the analogy of the computer with the mind seriously. One can see some influence of this thinking in the dissertation in the description of the morphophonemics of Batak as an ordered set of rules. If at the time of writing of the rules, I considered the exercise as an experiment in writing an item-and-process (IP) description, after 14 years, it still does not look so bad that I would like to rewrite it in the more traditional 'item-and-arrangement' (IA) model. This format is more difficult to understand and of course to write too, but it is much more concise than the regular IA format.

One way of representing the morphophonemics of Batak is in tabular form. This can be easily done for the external sandhi phenomena as in Table 2, p.59, where the row headings show the base form, i.e. phoneme, of the final segment of the first word and the column headings are the base forms of the initial phoneme of the second word. However, the irregularities of the internal sandhi make it less easy, or rather impossible, to present the alternations in tabular form. This was, as I recall, part of the attraction of experimenting with a rewrite rule model. It did prove feasible although it took many precious hours and days to make a working set of rules.

Any merit, or interesting aspect, of the model using an ordered set of rules is brought out by the ability of the rules to handle the complexity of the internal sandhi, both regular and irregular. Secondly, there is also support for this from the fact that there also seems to be a fixed order in the application of the three types of rules: Irregular Internal Sandhi (IIS), Regular Internal Sandhi (RIS), and External Sandhi (ES), in that order. Thirdly, within each of these subtypes there also seems to be a fixed order of application of the rules.

4. The chapter on Morphology follows the mainstream of structuralist grammar. The model used for it was Bloomfield 1961, with modifications to fit the structure of the Batak language. The modifications consist

in the re-organisation of the model for 'inflectional categories' and in the addition of certain terms like 'promissory' (p.71), 'explicit plural object' (p.79), 'intensive-instrumental' (p.70), for concepts that are specific to the Batak language.

The two charts on p.74 are a summary representation of the verbal inflections of Batak. The upper table does not only give the morpheme order structure of the verbal inflections but also the correlative components of discontinuous morphemes and the mutual exclusion of some of those inflections. The lower chart represents the aspectual forms of the verb, indicating the gaps in the patterns, by means of dashes.

In the choice of terms, the semantic or functional aspects of the morphemes have had some influence. However, the analysis is a formal, not a semantic, one.

The structure, though not necessarily the form, of personal pronouns is similar to the other members of the Indonesian branch of the Austronesian family.

The sub-classifications of the prepositions is grammatical-functional, and so is that of the conjunctions and the particles. This approach was taken after trying out other bases of classification, e.g. a purely semantic basis. The (grammatical) functional approach turned out to be a more satisfactory basis.

The last section of the Morphology deals with the structure of stems at some length. There are two morphophonemic symbols (p.92 ff) used in non-productive primary derivation, i.e. [l] and [t]. They follow a pattern of harmonious dissimilarity realisation. Thus [l] is /r/ before a root, i.e. bound base, containing /r/; otherwise it is represented by /l/. The morphophoneme [t] is /u/ before roots with the dissimilar non-back vowels /i e E a/ in first syllable; otherwise it is realised as /i/. This is a neat morphophonemic pattern of phoneme harmony that does not exist in the productive morphophonemics of Batak. An example is given on p.94, item (5); others are on p.97. Actually, there is another example on p.101, (12b) and (12c), that could be included in this rule. In this way, rules (12b) and (12c) could be conflated into $tar + st + \text{Root}$, with the additional 'coalescence rule of /i/ + /u/ into /i/' or 'the loss of /u/ if immediately followed by /i/'. I opted for the separation of the two instances as there were no other examples (that I could find) to support such an analysis. Nevertheless, it still looks like an attractive alternative to me, even at present.

Another interesting morphological phenomenon in Batak is reduplication; full reduplication is called 'doubling' in the dissertation. The description of both types of reduplication, full and partial, can be found from page 92 onwards. The interesting thing to mention here

also is the vowel harmony (i.e. dissimilarity) pattern in some reduplicated bases with a monosyllabic root. It is of course a moot point whether it would not be better to analyse them as two-syllable roots (and thus follow the predominant structure of word bases) that manifest a pattern of vowel harmony rather than a case of reduplication. However one may analyse the phenomenon, the fact remains that vowel harmony, with or without consonant changes, is common among the Indonesian languages; in bahasa Indonesia, for example, we have *lenggang-lenggang* 'swaying movement', *bolak-balik* 'back and forth', *desas-desus* 'rumours', *tetamu* 'guests', *leluhur* 'ancestors'; in Javanese, we find *wira-wiri* 'go back and forth', *bengak-bengok* 'shout incessantly', *nongas-nangis* 'cry intermittently', etc. Other languages have this phenomenon too; for example in English: 'crisscross, zigzag, ticktock' etc. As this vowel harmony does not seem to be productive at present, its ubiquitousness indicates that it might have been a universal feature in the past, perhaps in the times when language was exclusively oral. The interesting question would then be whether the present-day unwritten languages like those in the interior of Irian would have vowel harmony to a greater extent than the more familiar (written) languages, and especially whether it can be found as a productive grammatical feature.

5. Chapter IV of the dissertation, Syntax, here Chapter V, consists of two main divisions: (a) the structure of phrases (Section 5.1.) and (b) the structure of clauses (Section 5.2.), and a short one (c) the structure of sentences (Section 5.3.). In Section 5.1. the verbal directive construction (Section 5.1.2.1.2.) is of particular interest as the constituent formed by this construction are the main predicators of the favourite clause types containing a transitive verb, viz. (a) the 'Verb + Agent + Object', (VAO) type and (b) 'Verb + Object + Agent' (VOA) type. Examples of these two types can be found in Section 5.1.2.1.2., (1) and (2).

An example of the VAO type is:

| | | | | | |
|---|--------|----------|--------|------|-----|
| ditongos | nasida | ma | hepeng | i | [1] |
| sent (passive) | they | particle | money | that | |
| 'they sent the money or the money was sent by them' | | | | | |

An example of the VOA type is:

| | | | | |
|---------------------------|------|----------|------------|-----|
| manuan | eme | do | inatta | [2] |
| to plant | rice | particle | our mother | |
| 'Mother is planting rice' | | | | |

These two examples will also exemplify the main functional structure of the Batak sentence, namely comment + topic, i.e. what is said of something (= new information) + thing about which something is said (= assumed knowledge or old information).

In the analysis of the syntax, the influence is seen of the incipient view of the theoretical necessity of keeping apart the two upper levels of linguistic structure of the clause, namely the levels that later came to be called the **conceptual structure** level and the **syntactic structure** level. The former refers to the semantic relations of the units of information in an utterance, while the latter refers to the formal relations of the units in a clause. Sydney Lamb (1964) calls the former the **sememic stratum** and the latter the **lexemic stratum**. In Lamb's stratificational model, the sememic structure of the information carried by a clause like *'the man killed the tiger'*, or conceptually represented:

the → man ←→ killed ←→ tiger ← the

consists of three units: *'man, kill'* and *'tiger'*, related thus: *'man'* is the *'actor'* of the *'action'* expressed by *'kill'* and directed towards *'tiger'* the *'goal'* of that action. The same sememic relation is represented by the clause:

the tiger + was killed + by the man

The difference between the two clauses lies in syntactic structure and not in conceptual or semantic content.

In the syntactic section of the dissertation one can see the influence of this stirring model in the use of the notions of *'agentive'* from agent and *'objective'* from object (because *'goal'* was thought to be too outlandish at the time, and the term *'patient'* was considered not so suitable) in the description of the verbal directive constitutes that make up the comment component of the clause (Section 5.1.2.1.2.). In this way, I did not follow the conventional model of analysing sentence [1] as VSO and sentence [2] as VOS, because S in [1] is not really the syntactic subject of the sentence but the agentive component of the comment of the sentence; likewise in [2], S is not the subject of the sentence, but rather the topic. It was hoped that the retention of the term *'object'* would not be too confusing, as the term is not used for any other purpose but the *'goal'* of the conceptual level.

In the analysis made in the dissertation, the clauses:

manurat buku do ibana [3]
to write book particle he

'he is writing a book', will be VOA;

and

disurat ibana do buku [4]
to be written he particle book

'he has written a book', will be VAO.

The two clauses carry the same information as to the relationships between *'writing, book'* and *'he'*; the only difference lies only in the *'aspect'* of the action: in [3] the action of writing is emphasised, in

[4] the completion of the action is foregrounded. There is no implication of, or contrast between, active and passive here as in the corresponding English translation, although I use the terms as handy components of complex labels like 'passive transitive' and 'active transitive' to distinguish those two types of clauses.

In the discussion on the different sub-types of the major clause types, Section 5.2., I use the terms 'active transitive' and 'passive transitive'. These terms are based only on the morphological shapes of the verbs, and should not be understood in the same sense as in English to refer to the whole clause or sentence, namely that in the active voice the actor of the verb is the subject of the sentence, and in the passive voice the 'patient' of the verb is the subject of the sentence. As mentioned above there are no subjects and objects in the grammar of Batak in the same sense as in English. (In the Appendix, however, the term is used instead of 'topic').

The section on word order, Section 5.2.1., is also of some interest. Two major types of word order are recognised: (1) the direct word order and (2) the inverted word order (which is further sub-divided into normal emphatic and secondary emphatic). The main function of the inverted order mentioned here is for emphasis; the element moved to initial position is emphasised. At present, we would perhaps prefer to call this moving of an element to a prominent (here initial) position 'foregrounding'. The conceptual effect is, however, the same, namely to put some element in sharper focus or to bring out that element. Syntactically in Batak the foregrounded element becomes the comment of the clause; so we could perhaps also call this process 'commentisation', or 'commentalisation' in analogy to topicalisation.

6. The last part of the dissertation is an Appendix, in which I have tried to apply the grammar, or rather syntactic analysis, to the regular text of a story. This try-out of the system does work in terms of its objectives, namely to analyse the sentences into their constituent parts and to show the grammatical relations of the parts to each other. Naturally, no attempt has been made to make a conceptual structure analysis, as, in spite of the 'deep structure' - 'surface structure' dichotomy of the transformationalists, the notions relating to an underlying conceptual structure (or sememic or deep structure) and the actual linguistic (or realisable or surface) structure were still very vague and tentative.

It may be interesting to attempt a discourse analysis according to Pike's model (1977). It would be informative to attempt such an analysis on the referential level (or pragmatic level as some would

call it), and compare the results with the discourse structure in some other languages. A similar comparison of the grammatical structures of different languages would also yield interesting results.

7. In conclusion, I would like to repeat that the writing style of this dissertation is purely descriptive. It does not suit a pedagogical grammar of which the objective is to instruct the learner to use the grammar to make acceptable language (i.e. sentences) for the purpose of interaction and communication. Therefore, the presentation in the dissertation is very succinct to the extent that in places it gives a jerky impression. No statement is repeated or paraphrased; when necessary, reference is made to the particular sections at appropriate places. However, with some patience in reading, the structure of the Batak language will, I think, readily be apparent.

INTRODUCTION TO THE THESIS (1958):
CHAPTERS ON PHONOLOGY

As Indonesia and particularly Batak is rather unfamiliar in this part of the world, it seems to be proper to begin with a brief description of the two names.

Indonesia is the island republic that comprises some 3,000 islands lying between the continents of Asia and Australia. Up to 1949 it was politically more widely known to the other parts of the world as The East Indies or The Dutch East Indies, due to the Dutch colonisation of the area. In 1942 the Japanese captured it from the Dutch and when the Japanese capitulated in August 1945, the people of Indonesia seized the opportunity to proclaim their independence and fought for it till this new nation obtained *de jure* recognition of its independence and nationhood in December 1949.

Indonesia is made up of a polyglot population speaking some 200 languages. To facilitate intercommunication and exert a unifying influence, a common language, *Indonesian*, was adopted as the official national language. The nucleus of it is Malay, the language spoken natively in central East Sumatra, the most western and one of the biggest islands of Indonesia, and in the Malay Peninsula. This *bahasa Indonesia*, as the language is called, is still in a rapid stage of vocabulary growth to meet the needs of the nation in the various fields of its activities.

One of the languages spoken in Indonesia is *Batak*, spoken by a people designated by the same name, Batak, who inhabit the area around and south of the Toba Lake in central North Sumatra.¹ Collier's Encyclopedia (1952:10-499) gives the number of Bataks in 1949 as approximately one million, though two million would be a closer estimate.² They are classified as pure Indonesians.

¹See Map 1, p.

²Figures published in 1981 indicate a total Batak population of approximately 3½ million with the Toba Batak population being around 1½ million.

The Batak language has five major dialects, viz. Karo, Simelungan, Dairi, Toba, Mandailing.¹ Of these Toba-Batak has the greatest number of speakers and when people speak of Batak they usually refer to Toba-Batak. This study is an attempt at describing the phonemic structure of Toba-Batak, and when the Batak language is mentioned in the body of this thesis, the Toba-Batak dialect will be meant, if there is no statement to the contrary.

The Bataks possess a script of their own, which is a form of Devanagari alphabet.² The written literature of the Bataks consists mainly of works on magic, astrology and mythology. Nowadays this alphabet, though it has no practical value any more, is still taught in Junior High schools as a cultural feature. The Batak language is now written only in Latin characters.

In pre-Independence days English already occupied the most important place among the foreign languages taught in Indonesia, that is after Dutch of course, which was then the medium of instruction in the so-called Dutch schools. During the Japanese occupation both English and Dutch were struck off the curriculum. After the Proclamation of Independence in 1945 English was again taught, in all schools from Junior High on. As Dutch was no longer taught, English became the first foreign language in the curriculum. The teaching of English was stepped up due to the need of a widely understood foreign language so necessary for the international communication of the new-created nation, as in politics and commerce, and for a means to the sources of information and education.

When it was said above that English was the first foreign language taught in schools, it was not quite correct in the strictest sense for most parts of Indonesia where Indonesian is not spoken natively, for in those areas bahasa Indonesia is taught from the second grade on and is the medium of instruction from the fourth grade on. For the people in these areas Indonesian is to be regarded linguistically as a 'foreign' language, as the difference between the various vernaculars and Indonesian is about as great as the difference between - say - English and German. This does not apply to the major cities where all the children are bilingual, speaking the mother tongue at home and Indonesian outside, at play or elsewhere. But in the countryside and most small towns the vernacular is spoken in and outside the home and Indonesian is heard only occasionally as in schools and government offices. However, due to the close similarity of the vernacular and

¹See Map 2, p.

²See Appendix I.

Indonesian and the fact that Indonesian is the medium of instruction in schools from at least the fourth grade on, the children acquire a maximum of fluency in it in a relatively short time and with a minimum of effort. Yet for most such people Indonesian is a recognisably accented and acquired language. It is for the teaching of English to this category of people that a comparison of the structure of English and the vernacular is important and necessary. This thesis is an attempt at making such a comparison on the phonemic level, of English and Batak.

As language is, as all linguists have been stressing especially the last few decades, primarily speech, and as the English instruction in Indonesian schools has a double objective - i.e. mastery of the written and spoken language, though the written part of it is still given priority in practice - this study will also attempt to arrive at suggestions for the teaching of oral English at secondary schools in Batak-speaking areas, though they can easily be modified to meet the requirements in other similar areas. This study presupposes a knowledge of the phonemic system of English; therefore a description of only that of the Batak language is given.

There has not been much done yet in the study of the sound system of Batak. Due to this fact and the scarcity of books dealing with the subject and the lack of informants, this study must needs have a more or less independent character and the conclusions reached be tentative. The writer has analysed his own speech which fortunately cannot be said to be too dialectal, as he hails from Humbang¹, the central district of the region where Toba-Batak is spoken, so that his speech may be said to be what could be termed 'Standard Batak'. Moreover he has lived for many years in the other main districts of the Toba-Batak speaking area, namely in Silindung (where he attended elementary and secondary school) and in the Toba Valley (where his parents lived during his school years and where he thus spent his vacations and also three years of his adult life, part of the time as a teacher) so that his speech has been influenced by the three key - though not much differing - areas. Also it should be kept in mind that it is phonemics which is dealt with here rather than phonetics, in which it would be necessary to describe in detail the dialectal variations, both social and geographical.

Two appendices are inserted at the end. Appendix III is a concise description of the Batak alphabet. It is included here merely as an item of interest.

¹See Map 2 for the geographical names.

Appendix II gives a brief discussion of some of the spectrograms the writer has made in connection with this study.

It is hoped that this analysis of the phonemic system of Batak may at least show that much needs to be done still in the field of the descriptive analysis of Batak in particular, and the Indonesian languages in general.

INTRODUCTION TO THE DISSERTATION (1966):
FROM MORPHOPHONEMICS TO SYNTAX

1. AIM AND SCOPE

This study is a description of the Toba Batak dialect. It covers the phonology, the morphophonemics and the grammar of the language. In the grammar, morphology is covered more fully than syntax. A sample analysis is given in Appendix I.

2. THE LANGUAGE

Toba Batak is a dialect of the Batak language. Batak is usually divided into four dialects: Karo (+ Dairi and Pakpak); Simalungun; Toba; and Angkola (+ Mandailing); see e.g. the reproduction of Esser's linguistic map in Voorhoeve 1955. Toba and Angkola-Mandailing have a high degree of mutual intelligibility; on the basis of a comparison of 196 items, Dyen lists them as a single 'language' (Dyen 1965:27), distinct from Karo and Simalungun.

Toba Batak is spoken by about one million people living on Samosir Island and to the east, south and south-west of Lake Toba and henceforth in North Sumatra (see map). It is customary for the speakers of the language to distinguish five major subdialects: Samosir, on the island of Samosir; Toba Holbung, in the lowlands east and south of Lake Toba; Humbang or Toba na Sae, in the highlands south of Toba Holbung; Silindung, in and around the valley south of Humbang; and Hullang, in the mountainous regions southwest of Humbang down to Upper Barus.

This study is based on the speech of the writer, who is from Humbang but has had extensive contact with Toba Holbung and Silindung. To refresh his memory the following books were used: Loembantobing 1920; Sihombing 1959 and Van der Tuuk 1861.

3. EARLIER STUDIES

Although many references to the language can be found in ethnographic articles and travelogues, not many books and articles have been written on Toba Batak.

The first and the most important of the older books is Van der Tuuk's grammar (Van der Tuuk 1864, 1867), which deals with the sound system and parts of speech (+ morphology) respectively. The grammar is based on Toba Batak as it was spoken in the area of Tapus (Sibolga) and Upper Barus. The present writer has found, not surprisingly, that many of the forms cited there do not accord with, or are not found in, his speech. To date, Van der Tuuk's grammar is still the most exhaustive one written on those aspects of the language. It also furnishes much comparative material of Toba with Karo, Dairi, and Angkola-Mandailing, as well as some with Malay and Javanese. Van der Tuuk's interpretation differs from this study on the topic of verbs, nouns, and some functors. His discussion of the sound system is presented with the Batak script as the starting point; it consists mostly of rules of pronunciation of those symbols. Here, too, the description does not always accord with the present writer's pronunciation.

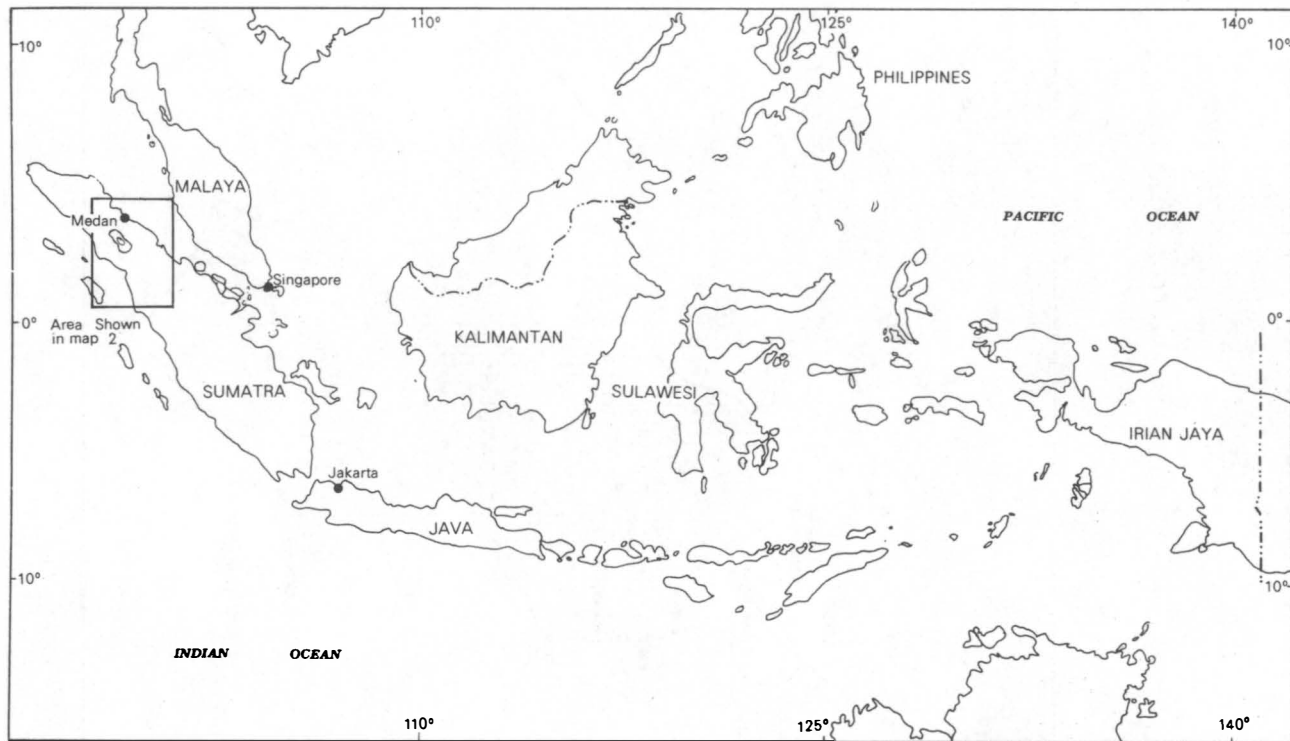
The next book to appear was Meerwaldt 1904. The language it describes is more in agreement with the present writer's speech than Van der Tuuk's corpus. It is broader in scope but much less exhaustive in treatment than Van der Tuuk's grammar. The Latin orthography of the language is now taken as the starting point of the discussion of the sound system. It significantly differs from Van der Tuuk's grammar in recognising most forms of the passive verb as distinct from a noun. However, it does not have the scholarly orientation of Van der Tuuk's grammar since it was intended for the use of missionaries who wanted to learn the language.

Marcks 1912, contains some descriptive material. It is a manual for the learning of the language; the greatest part of the book consists of exercises. The scant descriptive notes between the exercises contain some points not brought out in the earlier books.

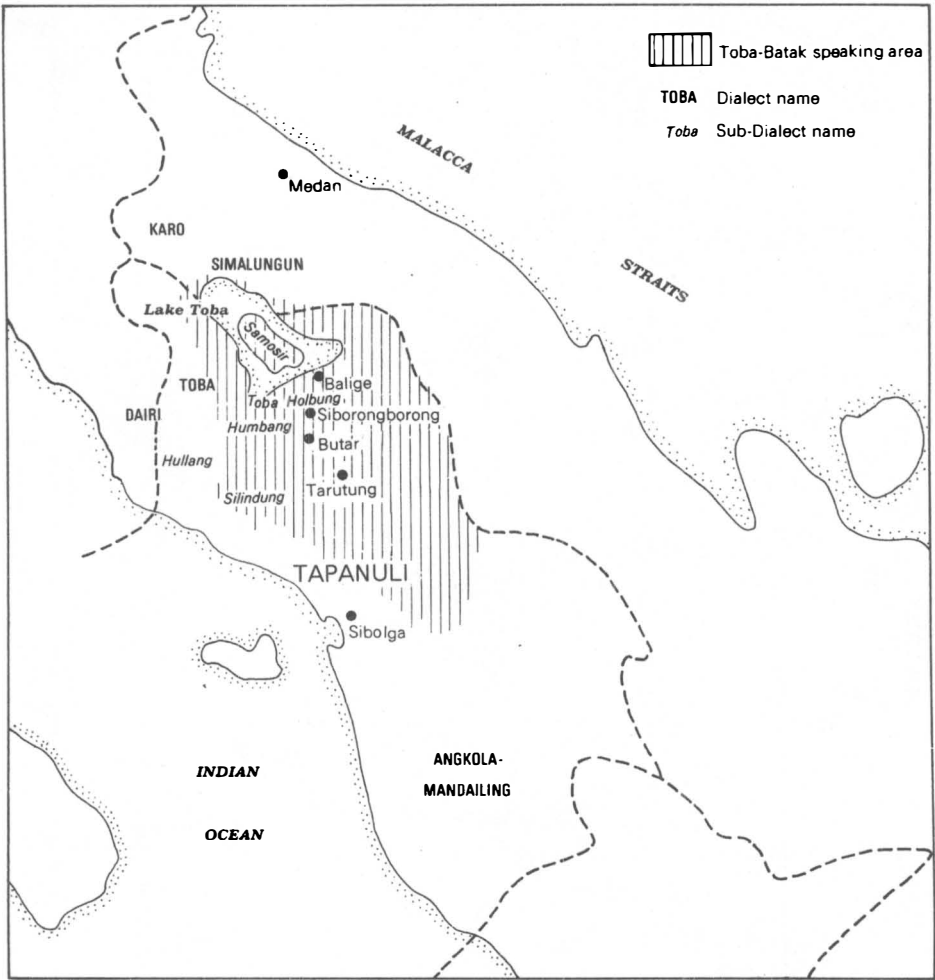
Of the more recent studies, there are two unpublished theses: the present writer's (Nababan 1966) which is the bases of the chapters on phonology, and Percival 1964 (this did not come to the writer's attention before the dissertation was nearing completion), which covers the sound system, morphophonemics, morphology, and syntax of Toba Batak. This is the most recent and the only description known to this writer which is based on modern methods covering the central subsystems of the language. It is based on a speech that is practically the same as

the one described in this study. This study differs on several points from the conclusions reached in Percival's dissertation, especially in phonology and morphology; the difference lies also in organisation and orientation of the syntax.

MAP 1
INDONESIA, SHOWING BATAK-SPEAKING AREA



MAP 2
THE BATAK DIALECT AREAS



CHAPTER I

THE BATAK PHONEMES

1. CONSONANTS

Phonemes are found by comparing two different utterances in which the difference is signalled by only one significant feature. Two such utterances are called a minimal pair. Thus in the pair *pála* 'nutmeg' - *bála* 'plague', the difference in content is recognised by any native speaker of Batak as caused by the initial sounds, or more specifically consonants, /p/ and /b/. The phoneme status of /p/ and /b/ can further be established by comparing other pairs, as:

pége 'ginger' - *bége* 'hear';
hápit 'clamp, vb.' - *hábit* 'sympathise';
sípak 'kick' - *síbak* 'tear, vb.'; etc.

By examining other utterances which pair minimally with /pála/ and which differ from it only in the first component, other phonemes can be established. Thus:

/t/ is found from the pair: *pála* - *tála* 'pond'
/h/ is found from the pair: *pála* - *hála* 'scorpion'
/s/ is found from the pair: *pála* - *sála* 'mistake'
/j/ is found from the pair: *pála* - *jála* 'net'
/l/ is found from the pair: *pála* - *lála* 'dissolve'
/n/ is found from the pair: *pála* - *nála* 'flame'
/m/ is found from the pair: *pála* - *malá* 'give',

though this is not exactly a minimal pair due to the differing stresses.

/g/ is found from the pair: *pála* - *gála* 'a child's game'
/d/ from the pair: *pátu* 'defeated' - *dátu* 'medicine man'
/c/ from the pair: *pátu* - *cátu* 'ration'
/r/ from the pair: *párik* 'earthen wall' - *rárik* 'scratch, vb.'
/k/ from the pair: *páti* 'essence' - *káti* 'a measure of weight'

/ŋ/ from the pair: pápan 'board' - páŋan 'eat'

/w/ from the pair: pápan - páwan 'foreman'

/y/ from the pair: hápur 'chalk' - sáyur 'vegetable'

The last pair /hápur/ and /sáyur/ is not a minimal pair. As no minimal pair can be found to contrast /p/ and /y/, this near-minimal pair, in which two differing features are found, is used. /h/ and /s/ can be shown by many minimal pairs to be different phonemes, as /hála/ 'scorpion' - /sála/ 'mistake', /háli/ 'dig' - /sáli/ 'borrow', /hódok/ 'sweat' - /sódok/ 'spade'. If the difference in content was caused by the difference in the initial /h/ and /s/, then /hápur/ and /háyur/ would be identical in content, which they are not, as /háyur/ does not exist in Batak. If the sequence /hVp../ were always contrasted by the sequence /sVy../ then the contrast /hípu/ 'hurry' - /sípu/ 'ember' would be impossible, which they are not, as those two forms are native Batak words, and many pairs like them can be found.

/w/ and /y/ are found in Batak only in their consonantal functions, as there are no diphthongs or complex nuclei in the Batak language.

The above contrasting pairs of utterances establish the Batak consonant phonemes, seventeen in all. A more detailed description of these consonants follows below.

1.1. THE STOPS

/p/ A voiceless bilabial stop. It occurs initially as in /pat/ 'foot', medially as in /dápur/ 'kitchen', and finally as in /álap/ 'fetch'. When final the /p/ is never released except when the following word begins with a vowel. A released final /p/ sounds foreign to a Batak ear, as often occurs when a foreigner learns Batak. This language habit of not releasing a final /p/ may be due to the fact that the corresponding voiced stop /b/ never occurs finally, so that no mishearing can result from not releasing the final bilabial stop.

Nor is the /p/ ever aspirated in Batak. Because of this lack of aspiration the Batak /p/ very often sounds as a /b/ to an American ear. This language habit makes it hard for a Batak to learn the non-phonemic but characteristic aspirated /p^h/.

/b/ A voiced bilabial stop. It occurs in initial and medial positions; initially in /bása/ 'generous', medially in /ábit/ 'clothes'. It never occurs in final position. When a /b/ occurs in final position originally in borrowed words, the tendency is to preserve a b in spelling, but it is invariable pronounced as /p/, e.g., bab 'chapter', an Arabic loanword, pronounced /bap/.

/t/ A voiced alveo-dental stop. The tip of the tongue touches the back of the upper teeth and the gum. It does not exactly correspond to /d/ as to place of articulation as the /d/ is articulated more to the back. It occurs in all three positions; in initial position as in /táli/ 'rope', in medial position as in /láta/ 'seed', and in final position as in /hot/ 'fixed'. Like the /p/, /t/ in final position is as a rule not released.

It is never aspirated. The observation made about this habit of non-aspiration for the /p/ also goes for the /t/ in learning the aspirated English [t^h]. However, this non-aspirated /t/ is never mistaken for a /d/.

/d/ A voiced alveolar stop. It occurs in initial and medial positions only, as in /dátu/ 'medicine man' and /ída/ 'see'. In loanwords where the /d/ originally occurs in final position, though still spelled with a d, it is pronounced /t/ in Batak, as in abad 'century', an Arabic loanword, which is pronounced /ábat/, even though thereby a homophone is created with the native word /ábat/ 'obstacle'.

/k/ A voiceless velar stop. It occurs in all three positions, initially as in /kátí/ 'a measure of weight', medially as in /tókɔ/ 'shop', and finally as in /roták/ 'dirty'. Its occurrence in initial and medial position is a recent development of Batak through borrowings.

For its relationship to /h/, see under /h/. When a final /k/ is followed by a vowel, either in a suffix or a following word, it is changed into /h/. For example:

/hódok/ 'perspiration'
 /hódohí/ 'the perspiration'
 /ának/ 'son'
 /ânahón/ 'this son'

But in loanwords the /k/ may be preserved or not, thus for example /kôtakí/ is found beside /kôtahí/ 'the box'.

The /k/ in final position, in contrast with the other two voiceless stops /p/ and /t/, is always released.

/g/ A voiced velar stop. It occurs only in initial and medial position, as in /gádɔŋ/ 'sweet potato' and /lágu/ 'favour'.

1.2. THE FRICATIVES

/s/ A voiceless alveolar fricative. It occurs in all three positions, initially in /sáda/ 'one', medially in /hósa/ 'breath' and finally in /hipás/ 'healthy'. The corresponding voiced alveolar fricative /z/ does not ordinarily occur in the speech of those who have the /j/ phoneme in their speech. Such people, amongst them the writer, usually

substitute /j/ for the /z/ in loanwords such as in *zaman* 'era' which they pronounce /jáman/.

/h/ A voiceless glottal fricative. It occurs with a single exception only in initial and medial positions, as in /hɔ/ 'you', in /fɪhur/ 'tail', and /hɔ́hɔs/ 'belt'. The only instance in the writer's idiolect in which /h/ is found in final position is /ah/ 'an interjection of reluctance or refusal', meaning something like 'Oh, no, I won't'.

Some words, in which /h/ occurs in medial position between a preceding stressed vowel and a following unlike and unstressed one, are in stylistic free variation with forms without a medial /h/, for example /áhu/ which varies with /áu/. The first form, /áhu/, is used in very formal situations, whereas /áu/ is used in ordinary conversation. Other similar words are:

/láho/ ~ /láɔ/ 'go'

/báhen/ ~ /báɛn/ 'do'

The writer himself invariably uses the form without /h/. This variant is being used more and more extensively now also in formal situations.

In older Batak the /h/ and /k/ seem to have been in complementary distribution, thus making up one phoneme: /h/ occurring in initial and medial positions, while /k/ occurred only finally. This can still be seen from the morphophonemic behaviour of the /k/ in word formations, as in the word base /dɔk/ 'say' and the derivations: /dɔhɔ́nɔn/ 'will be said' and /paddɔ́han/ 'what is said'. This theory is also supported by the fact that the Batak script has only one symbol, n^1 , for both /h/ and /k/. It can be safely inferred from this representation of two different sounds, at least so to a modern Batak ear, by a single symbol, that no confusion was caused by it because those two sounds were in complementary distribution. Thus:

/hóda/ = $\pi \times \wedge$ 'horse'

/fɪhur/ = $\pi \pi \pi \backslash$ 'tail'

/ának/ = $\pi \pi \pi \backslash$ 'son'

Lopez (1939) recognises this fact by saying that "h and k are equivalent".

Traces of this once allophonic arrangement of /h/ and /k/ can still be observed from the morphophonemic alternation of words ending in /k/ with a form ending in /h/, as for example:

/ának/ 'son; interest'

/ânahí/ 'that son'

/anáhan/ 'borrow at interest'

¹ See Appendix I, p.

From this we see that the variant with /h/ is taken when followed by a suffix beginning with a vowel. Though we have not yet described what a juncture is, we may make the statement that in a sequence of a word with a final /k/ and one beginning with a vowel, there are two possibilities, depending on the presence or absence of an intervening juncture, represented here by (*), thus, for example, we have:

/élehangimí/ /élek*àngimí/ - 'coax you kid brother'

The variant with /k/ is taken when there is a juncture after it.

There is, however, one condition when /k/ occurs medially; that is when the /k/ is doubled, as in /tókka/ 'taboo', /dókkon/ 'tell', etc. A contrast of /h/ and /kk/ can be found in the pair:

/dohónon/ 'will be said' and

/dókkónon/ 'will be invited or told to come'

In current Batak, maybe due to the influence of other languages through borrowings, the /k/ can now occur in medial position too as shown above, in p.3. When the /k/ occurs in medial position many, especially rustic, people still tend to pronounce it double, e.g.:

(Javanese) /salikúr/ > (Btk.) /salíkkur/ 'a card game; "Black Jack"'

(Indonesian) /toko/ > (Btk.) /tókko/ 'shop'

Even in educated and travelled people this tendency still persists sometimes, so that some people have, for example, /tók/ beside /tókko/.

Where the /k/ occurs in initial position many people from the countryside substitute /h/ for it in familiar words, as in /hácan/ for /kácan/ 'peanut'.

Yet in current Batak /h/ and /k/ are two different phonemes. This can be established by the existence of minimal pairs like:

/hóson/ 'asthma' - /kóson/ 'empty'

/tóho/ 'to one's liking' - /tók/ 'store'

1.3. THE AFFRICATES

/c/ A voiceless alveo-palatal affricate. It has a definitely more front articulation than its English counterpart. It is found in initial and medial positions. In native Batak words it is found only in medial position, as in /páca/ 'platform', /macái/ 'very', etc. In initial position it is found only in loanwords, as in:

/cátu/ 'ration',

/cúka/ 'vinegar', etc.

One peculiarity of the /c/ which may well be mentioned here, is the fact that in medial position it has the effect of a double consonant on the preceding vowel, i.e., shortening it. The doubling of other consonants, as will be shown in p.44, is a significant feature, as it contrasts with the single consonant in the same environment. However,

this does not seem to be the case with /c/, as, in the first place, only one variety of /c/ exists in medial position, and secondly, there is no such difference of length between an initial /c/ and the medial shortening /c/ as there is in the case of the other consonants. Therefore, in this study only one /c/ will be used to symbolise the initial as well as the medial variety of /c/. Many speakers of Batak have /ss/ instead of /c/ in medial position, thus:

/pássa/ for /páca/ 'platform'

/massái/ for /macái/ 'very', etc.

But in the writer's observation this type of pronunciation is rustic and receding. To his ear it sounds loose and affected or feminine. Some people have /ts/, but this is not so frequent.

This /c/ sound is not mentioned in Lopez's study (1939) but then, it is the result of the analysis of the Batak language of the 1900s. In modern Batak /c/ is a full-fledged and more and more extensively used phoneme.

/j/ A voiced alveo-palatal affricate. This is, in contrast with /c/, an original Batak phoneme. It is found in initial and final positions, as in:

/jábu/ 'house'

/jála/ 'and'

/rája/ 'tired'

People who pronounce the /c/ as /ss/ tend to pronounce the /j/ as /z/, especially in medial position, thus:

/zábu/ for /jabu/,

/lóza/ for /loja/.

In Batak the affricates, /c/ and /j/, do not occur in final position, hence the difficulty for a Batak to learn the final /c/ and /j/ in other languages, such as English.

1.4. THE NASALS

/m/ A voiced bilabial nasal. It occurs in all three positions, as in:

/mánan/ 'eat'

/émε/ 'unhusked rice'

/ásom/ 'lime'

/n/ A voiced alveolar nasal. It occurs in all three positions, as in:

/nála/ 'flame'

/gúna/ 'use, n.'

/ípon/ 'tooth'

/ŋ/ A voiced velar nasal. It occurs in all three positions, as in:

/ŋalí/ 'cold'

/dúŋɔ/ 'wake up'

/hépeŋ/ 'money'

The occurrence of /ŋ/ in initial position tends to be a little troublesome for people with a language background in which the /ŋ/ does not occur initially, as English for example. It may therefore be worth noting that there are two words in each of which the /ŋ/ occurs in all three positions, i.e.:

/ŋaŋáŋ/ 'open', and /ŋɔŋɔŋ/ 'idle'.

The Batak alphabet, unlike Latin, has a symbol for this sound, i.e. 'ɔ'.¹

1.5. THE LINGUALS

/l/ A voiced alveolar lateral lingual. It occurs in all three positions, as in:

/láta/ 'seed'

/sáli/ 'borrow'

/síhɔl/ 'longing, n.'

/r/ A voiced alveolar trilled lingual. It occurs in all three positions, as in:

/rája/ 'king'

/mára/ 'danger'

/sár/ 'widespread'

In some areas, notably in the mountain district of Humbang, a velar fricative /R/ occurs. This type of pronunciation has very little prestige and is very rapidly receding.

Another noteworthy characteristic of /r/ is the fact that for most people it changes into /l/ when the following sound is /l/ without an intervening juncture, as for example in:

/pór + lak/ > /póllak/ 'garden'

/mar + lúga/ > /mallúga/ 'row, vb.'

/par + lájja/ > /pallájja/ 'peddler'

The influence of the linguals on each other can be also seen in the common rustic pronunciation of words which contain both /l/ and /r/, in which the /l/ is substituted by /r/, as in:

/saraśar/ for: /saraśal/ 'trousers'

/rafar/ for: /láyar/ 'sail, n.'

In the last example it is to be noted that the more rustic /i/ is

¹See Appendix III.

substituted for the semivowel /y/. Lopez (1939:33) notes that Batak in some words has /r/ for which Primitive Indonesian had /l/ or retroflex /ɭ/, as in:

(Btk.): /rípʊr/ from Pr. In.: /lípʊ/ 'destroy'

(Btk.): /rúar/ from Pr. In.: /lúa/ 'come out'

1.6. THE SEMIVOWELS

The semivowels /w/ and /y/ occur in Batak only in their consonantal function. The semivowel is a rather recent development in the Batak sound-structure. Lopez (1939:17) does not include these two in his list of Toba-Batak sounds. The semivowel came into Batak through loanwords and through the contact of its speakers with other languages. The /y/ is a regular feature in the Mandailing Batak dialect, and the /w/ and /y/ are common phonemes in Modern Indonesian. Brandstetter (1915:17) includes both semivowels in the sound system of Primitive Indonesian. The Toba-Batak dialect seems to have dropped them in the course of its development, and in rustic pronunciation the /w/ still tends to be substituted by /ɔ/ and the /y/ by /i/.

/w/ A voiced bilabial semivowel. Its most common occurrence is in medial position, as in:

/báwaŋ/ 'onion'

/máwas/ 'orangutan'

It also occurs initially in loanwords not quite naturalised yet, as in:

/wáktu/ 'tempo'

/wáras/ 'sane'

or in borrowed personal names, as in:

/waldémar/.

/y/ A voiced alveo-palatal semivowel. It occurs in initial and medial positions, as in:

/yúran/ 'contribution'

/sáyur/ 'vegetable greens'

/láyar/ 'sail'

In the writer's speech it occurs what could be termed obtrusively in the sequence /ái/ plus /ámaŋ/ which become: /âiyamáŋ/ 'an interjection', something like 'By golly'.

The semivowels never occur in final position in Batak, as in the French /fami:y/, nor do they have a vocalic function which would cause their occurrence in final position at the end of complex nuclei.

1.7. THE GLOTTAL STOP

In words like [ŋáʔŋat] '*devour*' there is a stop after the first /a/. This kind of stop is also found in [daʔdának] '*child*' and [láʔlap]. Orthographically the first word is ngatngat, the second word dakdanak and the third word laplap. In the production of this break in the writer's speech, which is by closing the glottis, there is a definite upward movement of the back of the tongue towards the velum without touching it, however, as it would if /k/ is pronounced. For reasons given below it seems to be more justified not to give to it a phonemic status and to call it a glottalised /k/, which would be transcribed [k̚] phonetically, and to represent it by /k/ in phonemic notation. No confusion will arise from transcribing those words /ŋákŋat/ and /dakdának/, as it is normal not to explode the /k/ in such positions, i.e. final in the syllable.

We have seen above that the glottalised /k/ is used for the glottal stop that occurs in ngatngat and laplap. We can infer by means of Batak morphology in which reduplication is a major device, that these two words are reduplications, in which the first parts have [ʔ] as the morphophonemic alternant for /t/ and /p/. In [doʔdók] it is a morphophonemic alternant for /k/. Due to the limitation of its phonotactical possibilities, i.e. it occurs only as the first component of a medial consonant cluster, we have hesitated to give it a phonemic status, and will regard it as an allophone of the consonant which it can substitute in morphophonemic variation. The problem, now, is, to which of the three voiceless stops, /p, t, k/, will it be assigned?

We have proposed above to assign it to /k/, due to its greater phonetic similarity with /k/ than with any other stop, and, secondly, because in the writer's speech there is some activity of the back of the tongue in the general area of the /k/. We will now present another argument for it of a statistical nature.

In reduplications the [ʔ] is found in all instances of words ending in /p/, /t/, or /k/, if the initial sound is a voiced consonant, as, for example:

[gaʔgúk] - '*heaped*'

[gáʔgat] - '*chop*'

[gáʔgap] - '*treason*'

In reduplications with a final /k/ and an initial voiceless consonant we have /k/ consistently in a similar position, and this cannot be said of the /p/ and the /t/, as can be seen from the comparative list below. An asterisk before a form indicates that it is only hypothetical, i.e. that that would be the form, if such an utterance existed (though they may sometimes be found in children's speech).

(1) /k/

/pákpak/ - 'crest'
 /túktuk/ - 'knock'
 /saksák/ - 'snow-white'
 */hákkak/ - --
 /kékkék/ - 'giggle'

(2) /p/

/póppop/ - 'toot of horn'
 /táttap/ - 'wash'
 /sácap/ - 'slice'
 /hókkap/ - 'sacrifice for'
 */kákkap/ - --

(3) /t/

/píppit/ - 'close the eyes'
 /túttut/ - 'sound of horn'
 /sóssot/ - 'insist something on somebody'
 /hóttót/ - 'thick, of broth'
 */kákkat/ - --

The above list shows that the /k/ is found in all five instances of column one: that the /p/ occurs only once in the second column; and that /t/ is found only twice in column three. This means that the formulation of the conditioning of the complementary distribution of [k] and [ʔ] is easier and simpler because of its consistency than it would be for /p/ and /t/, which is sufficient reason to assign [ʔ] to /k/.

The glottal stop, or 'hamzah' in Indonesian, does occur in other Indonesian languages, as for example in Javanese: /tataʔán/ 'coaster'.

Lopez (1939:18) calls this sound phenomenon a laryngeal and assigns to it the symbols: ʔ and ʕ, and says that it is found very frequently in initial, medial and final positions. The example he gives is /ʔásuʕ/. In the writer's speech there seems to be no activity in the glottis before /a/ except the normal narrowing of its opening in preparation of the production of a sound, but after /u/ there is a momentary closure of the glottis. The first is the normal transition from silence to sound called 'vowel onset' and the final laryngeal seems to be only a somewhat abrupt stop as a transition from sound to silence.¹ This glottal activity has no phonemic values, as we can leave it out and drawl out the /u/ and let it die out without causing

¹The absence of a distinct glottal stop can be seen on the appended spectrogram No.3b for the word /so/.

any change in its content. Secondly, in Batak the final /k/ is softened into /h/ if followed by a vowel, thus /lápuk/ 'mold' becomes /lapúhɔn/ 'moldy' because of the following suffix /-ɔn/. If the laryngeal after, for example, /háli/ 'dig' is a genuine glottal stop /ʔ/ and thus a consonant as Lopez considers the laryngeal to be, then, it being homorganic with /h/, it seems natural that it also would be softened into /h/ by suffixing /-ɔn/, thus */halíhɔn/. But it does not work that way, as we have /halíɔn/ 'to be dug' and not */halíhɔn/.

In such situations the glottal activities do exist, but they are at best phonetic phenomena, and are non-phonemic as their presence and absence are not significant. This abrupt stop after final vowels is one of the differences between the Batak and English vowel systems.

From the above descriptions of the consonants a structural arrangement can be set up. In addition to position, the stops are distinguished by voicing. If we include /c/ and /j/ in the stops, we have then a set of four voiceless stops /p, t, c, k/, which is matched by a set of four voiced stops /b, d, j, g/. For the nasals and spirants the distinction of voicing is missing, as all the nasals are voiced and all the spirants (or fricatives) are voiceless. The semivowels also lack this distinction of voicing. The linguals are distinguished by the manner of articulation of the tongue, one is lateral and the other trilled. The resulting arrangement is presented below.

| | | | | | |
|-----------------------|-----------|----|----|---|---|
| | Voiceless | /p | t | c | k |
| STOPS { | Voiced | b | d | j | g |
| NASALS | | m | n | ɲ | |
| SPIRANTS (Fricatives) | | s | h | | |
| SEMIVOWELS (Glides) | | w | y | | |
| | Lateral | | l | | |
| LINGUALS { | Trilled | | r/ | | |

1.8. CONCLUDING REMARKS

In general the articulation of the Batak consonants does not differ too much from that of English. Notable differences can be seen in the pronunciation of /r/ and the voiceless stops. In Batak the /r/ is distinctly trilled, and the voiceless stops are always unaspirated. The distribution differs considerably and in particular the consonant clustering. These differences will be discussed further in another chapter, as a means to discover the causes for learning difficulties. A phonetic classification of the consonants of Batak according to manner and place of articulation is presented on the accompanying chart.

DIAGRAM 1
CONSONANT CHART

| | | Bilabial | Alveodental | Alveolar | Alveopalatal | Velar | Glottal |
|------------|--------------|----------|-------------|----------|--------------|-------|---------|
| STOPS | Voiceless | p | t | | | k | |
| | Voiced | b | | d | | g | |
| FRICATIVES | Voiceless | | | s | | | h |
| | Voiced | | | | | | |
| AFFRICATES | Voiceless | | | | c | | |
| | Voiced | | | | j | | |
| NASALS | Voiced | m | | n | | ŋ | |
| LINGUALS | Lateral, Vd. | | | l | | | |
| | Trill, Vd. | | | r | | | |
| SEMIVOWELS | | w | | | y | | |

2. VOWELS

Languages differ in many respects, amongst others in the number and articulation of the consonants. But the vowels show at least as well marked differences. Even native speakers of the same language show considerable variations in their vowels depending on their geographical and sociological situations. It is the vowels that give learners of the spoken language one of the greatest difficulties. Batak also differs from English more in the vowel sounds than in the consonants.

When we compare the following minimal pairs:

- /míra/ - 'male duck'
- /méra/ - 'purple'
- /mára/ - 'danger'
- /múra/ - 'easiness'
- /móra/ - 'richness'

five different vowels are found contrasting, namely: /i/, /ε/, /a/, /u/, /ɔ/. By comparing:

- /día/ - 'what'
- /déa/ - 'lure, vb.'
- /dúa/ - 'two'
- /dóa/ - 'invoke'

two other vowels are found, namely, /e/, /o/, thus making the number of vowels seven. To establish the phonemic status of these seven each, one should be contrasted with the other six in minimal pairs. This is already accomplished for /i/ and /u/ in the two groups of minimal pairs above. Minimal pairs to contrast /a/ with /e/ and /o/ follow below:

- /a/ with /e/: /híta/ - 'we'
- /híte/ - 'bridge; ladder'
- /a/ with /o/: /tíbba/ - 'scoop'
- /tíbbo/ - 'height'

Many other minimal pairs can be found like the above ones. But for the contrast of /o/ with /ε/, and /ɔ/ with /e/ the writer has after a long search been able to find only one for each pair, namely:

- /o/ /ε/: /dósa/ - 'sin'
- /désa/ - 'rural village'
- /ɔ/ /e/: /alś/ - 'feel a dislike, reluctant'
- /alé/ - 'vocative, you'

However, they are sufficiently dissimilar phonetically to override suspicion, as /o/ is a higher mid back tense vowel, while /ε/ is lower mid front lax, and /ɔ/ is a lower mid back lax vowel, while /e/ is higher mid front tense.

Two additional pairs for the contrast of each vowel - with the above exceptions - with the other six are listed below. The case of /e/ and

/ε/, and that of /o/ and /ɔ/ will be discussed afterwards.

| | | | |
|--------------|--------------------------|----|---------------------------------|
| /a/ and /i/: | /háta/ - 'word' | -- | /híta/ - 'we' |
| | /ála/ - 'cause' | -- | /íla/ - 'shame' |
| /a/ and /u/: | /háta/ - 'word' | -- | /húta/ - 'village' |
| | /ála/ - 'cause' | -- | /úla/ - 'do; work' |
| /a/ and /o/: | /báru/ - 'recent' | -- | /bóru/ - 'daughter' |
| | /dáli/ - 'string beans' | -- | /dóli/ - 'male' |
| /a/ and /ɔ/: | /táru/ - 'carry' | -- | /tóru/ - 'under, n.' |
| | /ra/ - 'probably' | -- | /rɔ/ - 'come' |
| /a/ and /e/: | /séga/ - 'defect' | -- | /sége/ - 'sieve' |
| | /ábat/ - 'hindrance' | -- | /ébat/ - 'visit' |
| /a/ and /ε/: | /máta/ - 'eye' | -- | /máε/ - 'dead' |
| | /pála/ - 'nutmeg' | -- | /páε/ - 'beat or work to death' |
| /u/ and /o/: | /úli/ - 'beauty' | -- | /óli/ - 'marry' |
| | /búru/ - 'hunt' | -- | /bóru/ - 'daughter' |
| /u/ and /ɔ/: | /tuk/ - 'enough' | -- | /tok/ - 'headache' |
| | /sáur/ - 'success' | -- | /sáɔr/ - 'mixed' |
| /u/ and /i/: | /úla/ - 'do' | -- | /íla/ - 'shyness' |
| | /húta/ - 'village' | -- | /híta/ - 'we' |
| /u/ and /e/: | /bégu/ - 'ghost' | -- | /bége/ - 'hear' |
| | /múdar/ - 'blood' | -- | /médar/ - 'lost' |
| /u/ and /ε/: | /túan/ - 'gentleman' | -- | /téan/ - 'inherit' |
| | /sáu/ - 'a fruit' | -- | /sáε/ - 'finished' |
| /i/ and /o/: | /sílu/ - 'hurry' | -- | /sólu/ - 'boat' |
| | /ɲílu/ - 'acidity' | -- | /ɲólu/ - 'life' |
| /i/ and /ɔ/: | /ri/ - 'a kind of grass' | -- | /rɔ/ - 'come' |
| | /tíru/ - 'imitate' | -- | /tóru/ - 'under' |
| /i/ and /e/: | /sfge/ - 'ladder' | -- | /sége/ - 'sieve' |
| | /fda/ - 'see' | -- | /éda/ - 'female cousin' |
| /i/ and /ε/: | /ísi/ - 'fill' | -- | /íε/ - 'who, interr.' |
| | /bihá/ - 'wide open' | -- | /bohá/ - 'how' |

The only minimal pairs that can be found for /o/ and /ε/, and for /ɔ/ and /e/, and the reason why they are to be regarded as different phonemes, have been given above. This leaves us still two suspicious pairs, namely:

/e/ -- /ɛ/, and /o/ -- /ɔ/.

These two pairs consist of members that have phonetic similarity. In articulation, /e/ is different from /ɛ/ only in tenseness and the height of the tongue - and this difference in tongue position can be very small - and in the same way is /o/ different from /ɔ/. Brandstetter (1915:18) notes that in many living Indonesian languages some vowels have two shades ('Nüancen'). This could be the case in older Batak, but in modern Batak we have to do, not with two shades of one vowel, but with two different vowel phonemes. Lopez (1939:17) agrees with Brandstetter's finding when he says that there are five vowels in Toba-Batak.

When we examine the distribution of /e/ and /ɛ/, we see that our suspicions are not entirely without ground. They are not conditioned by a preceding phoneme, vowel or consonant, as we can have almost any consonant or vowel before /ɛ/ and most of them can also precede /e/. But there seems to be a pattern of distribution conditioned by the following consonant,¹ thus:

- (1) before voiced stops (/b, d, g/) and /l/, we have /e/,
- (2) before voiceless stops (/p, t, k/), fricatives, nasals and /r/, we have /ɛ/.

Examples:

/e/: /déba/ - 'part of; partly'
 /médar/ - 'missing; lost'
 /séga/ - 'defect'
 /péle/ - 'worship'
 /ɛ/: /répak/ - 'limping'
 /sétan/ - 'the devil'
 /sékká/ - 'handkerchief'
 /pésan/ - 'destroyed'
 /héhe/ - 'get up'
 /nuáen/ - 'now'
 /éme/ - 'unhusked rice'
 /ménak/ - 'calm'
 /serép/ - 'humble'

However, there are many exceptions to this distribution, as for example:

/édde/ - 'song', in which we would expect /e/
 /héla/ - 'son-in-law', in which we would expect /e/
 /désa/ - 'compass point', in which we would expect /ɛ/
 /béta/ - 'come; let us', -idem-
 /dengán/ - 'good', -idem-

¹See Diagram 2, p.16.

DIAGRAM 2
DISTRIBUTIONAL PATTERN OF /e/ - /ɛ/, AND /o/ - /ɔ/

This tabulation is based on a corpus of 66 e-words (of which 27 have /e/ and 39 have /ɛ/) and 107 o-words (of which 45 have /o/ and 62 have /ɔ/), taken from a story and a sermon in the Batak Church's almanac of 1956, plus a few words that the writer has searched for to illustrate some cases. Only word bases, except one derivational word ending in /-ɔn/, are included in the corpus. The '+' means 'found more than once', and the '-' means 'not found', the '!' means 'found in one word'.

| [CV] Preceding | | CONSONANT | [VC] Following | | [CV] Preceding | | CONSONANT | [VC] Following | |
|-------------------|-----|-----------|-------------------|-----|-------------------|-----|-----------|-------------------|-----|
| /e/ | /ɛ/ | | /e/ | /ɛ/ | /o/ | /ɔ/ | | /o/ | /ɔ/ |
| + | + | p | - | + | ! | + | p | - | + |
| + | + | b | + | - | + | + | b | + | - |
| - | + | t | - | + | - | + | t | + | + |
| + | + | d | + | - | + | + | d | + | - |
| - | + | k | - | + | - | + | k | + | + |
| + | + | g | + | - | + | + | g | + | - |
| + | + | s | - | + | + | + | s | + | + |
| - | + | h | - | + | ! | + | h | - | + |
| - | - | c | - | ! | - | - | c | - | + |
| - | + | j | - | - | + | + | j | + | ! |
| + | + | m | - | + | - | + | m | - | + |
| - | + | n | - | + | - | + | n | - | + |
| - | + | ŋ | - | + | ! | + | ŋ | - | + |
| + | + | l | + | - | + | + | l | + | + |
| - | + | r | - | + | ! | + | r | ! | + |

The semivowels are excluded, due to the rare occurrence with these vowels.

* Found only in weak stressed closed syllables.

Exceptions like these do not show any patterning, hence they cannot be formulated into conditioned exceptions.

The above pattern of distribution, if it were absolute, would be enough to justify concluding the /e/ and /ɛ/ were allophones. Yet in addition to the irregularities already noted, there are also three pairs in which /e/ and /ɛ/ contrast minimally. They are:

- | | | | | |
|--------|-----------------------------|----|--------|---------------------------------|
| /be/ | - 'each' | -- | /be/ | - 'any more' |
| /désa/ | - 'compass point' | -- | /désa/ | - 'rural village' |
| /pe/ | - 'interjection of disgust' | -- | /pe/ | - 'kind of adverb of admission' |

The forms /désa/ and /désa/ can occur in a phrasal minimal pair, thus:

/tu dɛsa día ibàna làɔ/ - 'In what direction did he go?'

/tu dɛsa día ibàna làɔ/ - 'To what village did he go?'

/be/ and /be/ also occur in a non-minimal pair (differentiated by stress), thus:

/nùŋŋa sía bé/ - 'Each has nine already.'

/nùŋŋa sía be/ - 'There are nine already.'

From all this it seems to be best to consider /e/ and /ɛ/ as two different phonemes, because:

- (1) the conditioned distribution as shown above has exceptions, and new words tend to increase their number, and
- (2) minimal pairs exist that show their contrastive distribution.

The case of /o/ and /ɔ/ is very much like that of /e/ and /ɛ/. The distributional pattern, as can be seen on the diagram on p.16, shows partial similarity to that of /e/ and /ɛ/, and like that pattern, the distribution of /o/ and /ɔ/ has exceptions. There are, moreover, two minimal pairs in which the /o/ and /ɔ/ contrast, namely:

/jólo/ - 'front' -- /jólɔ/ - 'auxiliary word used to soften a request, as English 'please''

/do/ - 'first tone of a diatonic scale' -- /dɔ/ - 'an auxiliary word used in affirmative sentences'

Hence, as in the case of /e/ and /ɛ/ and in parallelism to it, it is best to consider /o/ and /ɔ/ as two different vowel phonemes.

Now that the number of vowel phonemes has been decided on, let us compare one of the vowels, /ɔ/, in the following utterances which represent all the possible combinations of /ɔ/ with a following different phoneme, to see how the vowel is influenced by its environment.

- (1) /gɔar/ - 'name'
- (2) /gɔtap/ - 'sever'
- (3) /sɔ/ - 'quiet; motionless'
- (4) /tɔp/ - 'joined'
- (5) /hɔt/ - 'fixed'

- (6) /tɔk/ - 'aching, (of head)'
- (7) /hɔs/ - 'late'
- (8) /sɔr/ - 'almost'
- (9) /tɔl/ - 'spinning top'
- (10) /bɔm/ - 'bomb'
- (11) /bɔn/ - 'on credit'
- (12) /gɔŋ/ - 'inactive; cease action'
- (13) /mɔsɔs/ - 'rubbing'

In the above examples we see that the /ɔ/ is not always the same length in all the utterances. It is longest in /sɔ/, and slightly shorter than that in /gɔar/. There is no perceivable difference in length between the /ɔ/'s of /gɔar/ and /gɔtap/. In /sɔ/ the transition to silence is not gradual but rather abrupt, but there is no definite articulation which might be called a glottal stop. There usually is a closure of the glottis which is a complementary feature of Batak final vowels. In Batak, with final vowels, one does not stop the breath by just ceasing the pressure on the chest cavity, but by obstructing the passage of the breath in the glottis, nor does one let the voice trail off as in English, where an utterance never ends in a monophthong but in a complex nucleus, at least under stress.¹ That this kind of stop is not phonemic has been shown in the section dealing with consonants on p.10. To show impatience we lengthen the /ɔ/ to at least double length and let it trail off into breath without, however, producing /h/, thus only resembling a complex nucleus. This kind of lengthening of the vowel, accompanied by a rise in pitch, applies to all utterances.

The /ɔ/ is shortest in /tɔp/, /hɔt/, and /tɔk/, but any native speaker of Batak can still recognise this /ɔ/ as essentially the same as the /ɔ/ of /gɔar/ and /sɔ/.

In /tɔs/ the /ɔ/ is a little longer than in the last three examples. This degree of length is also found in the /ɔ/ of /bɔm/, /bɔn/, and /gɔŋ/.

In /sɔr/ and /tɔl/ the /ɔ/ is again slightly longer than in /tɔs/. This greater length seems to be the result of the relatively longer time the tongue gives to the articulation of the following lingual.

These differences of length are small and their production automatic for a native speaker. The following five degrees length of /ɔ/ - or any other vowel - are seen in the above examples:²

In open syllables:

1. The /ɔ/, or, in fact, any other vowel, is longest when it occurs in

¹Dr. A.A. Hill in class lecture.

²See also Appendix II, Spectrograms 1-6.

final position.

2. Next in length is the /ɔ/ when followed by another vowel or consonant in open position.

In closed syllables:

3. When the /ɔ/ is followed by a final lingual.

4. When the /ɔ/ is followed by a final /s/ or a nasal.

5. The /ɔ/ is shortest when followed by a final voiceless stop in the syllable.

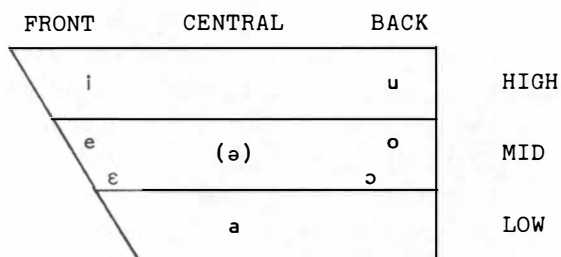
One other feature is observed in the /ɔ/ of /bɔm/, /bɔn/, /gɔŋ/, and /mʰsɔs/, namely, that it is slightly nasalised when preceded or followed by a nasal.

In the above examples the lexical content is not changed by the change in expression, i.e. lengthening or nasalisation. The occurrence of a specific type of /ɔ/ is totally predictable in terms of its environment. Hence we can conclude that length, or quantity, and nasalisation in Batak vowels are non-phonemic, so that the varieties of /ɔ/ above are allophones.

This same kind of comparison can be made for the other vowels with similar results regarding the length of vowels and nasalisation, thus supporting the above finding about their non-phonemic character.

On the chart below, the so-called Vowel Triangle (Gleason 1955:201), the seven vowels are shown in their relative tongue positions.¹

DIAGRAM 3
VOWEL CHART



Thus the vowels can be classified according to the place and manner of articulation as follows:

/i/: high front unrounded, tense

/e/: higher mid front unrounded, tense

/ɛ/: lower mid front unrounded, lax

/a/: low central unrounded, lax

/ɔ/: lower mid back, rounded, lax

¹See also Diagram 6, p.136.

/o/: higher mid back, rounded, tense

/u/: high back rounded, tense

Based on these characteristics of the vowels, we can set up a simpler structural arrangement than the one on the chart above. The distinguishing factors are the front-back and high-low position of the highest point of the tongue in the articulation of the vowels, thus:

| | | |
|---|---|---|
| i | | u |
| e | - | o |
| ɛ | a | ɔ |

We see that this arrangement is a symmetrical one. In weak stressed positions the vowels tend to move towards the centre of symmetry indicated here by the hyphen, the weaker the stress the nearer they come to that position, without, however, ever reaching it.

It may be worth noting here, that Batak is developing a new vowel phoneme through the many loanwords from 'bahasa Indonesia', which the Bataks more and more tend to pronounce in the Indonesian manner. The vowel meant here is /ə/, or pepet /pəpet/. The words with a pepet are there to stay and their number is constantly increasing. Examples of such words are:

- /kàpal tərbán/ - 'flying machine'
- /təgás/ - 'unequivocal'
- /təkát/ - 'slogan'
- /kəmèrdeká + an/ - 'independence'
- /tərán/ - ('evidently', the corresponding native word /tɔrán/ means 'clear, light, evident', while the loanword is used to mean only 'evident')

The pepet is already a phoneme for those who use it, but they are still relatively few. For people who do not use it actively, and they constitute by far the greater majority, it is obviously not a phoneme, though apparently it is at least familiar. For this reason the writer has not included it in the list of vowel phonemes. It may be worth noting here that the familiarity with the 'colorless vowel' as Brandstetter (1915:4) calls it, is an advantage in the study of foreign languages, in particular English, in which its counterpart, the 'schwa', is an important vowel.

To conclude this section on vowel phonemes one more thing needs to be said about the production of Batak vowels, namely, that the Batak vowels are always pronounced in a rather clipped way and that they are entirely monophthongal in contrast with English in which the vowels show signs of the tendency to diphthongisation, which is called drawl in everyday speech.

2.1. THE SUPRASEGMENTAL PHONEMES

In addition to the segmental phonemes, consonants and vowels, there are other features that characterise a language, namely the suprasegmental phonemes. They include the stresses, pitches and junctures, and are called suprasegmental, because they can occur simultaneously with the segmental phonemes and with each other, though strictly speaking the junctures should be excluded from this term. Hence they are usually superimposed over the segmental phonemes in transcriptions.

A foreign speaker is recognised from a native speaker of that language not only by his use of the allophones of the segmental phonemes, but also by his suprasegmental phonemes, especially the pitch contour.

3. STRESS

Brandstetter (1915:86) gives four stress systems ('Betonungssystemen') for the various branches of the Indonesian stock of languages. One of these categories, namely, the one in which the word base is stressed on the penult or ultima according to definite rules, he calls 'Toba Typus'. He refers here to the Toba Batak dialect, the subject of enquiry in this analysis.

When we compare the Batak utterances

/tíbbó/ - 'height' and

/tibbó/ - 'high',

both of which are orthographically alike: timbo, we find that all the segmental phonemes are the same. As these two utterances even in isolation denote different, though related, things, there must be something else which signals this difference. Here it is the difference in position of loudness of the syllables or 'stress', also called 'accent' in everyday speech. The first utterance, which is a noun, has the stronger stress on the first syllable /tib-/ , while in the second one, which could be termed a predicative adjective or what in English would be called a group VII modifier,¹ the stronger stress is on the second syllable /-bo/. From this example it can be seen that in Batak stress is distinctive and hence is a phoneme.

In Batak the stress generally falls on the penult. When an utterance with penultimate stress is expanded by means of affixes, the stress, if necessary, moves to the new penult.

For example:

/dálan/ - 'road'

/mardálan/ - 'walk, vb.'

¹Term used by Dr. A.A. Hill in class lectures.

- /daláni/ - 'walk, vb. transitive'
 /dalánna/ - 'his road, or, way'
 /pardalaníon/ - 'journey'
 /dalannasída/ - 'their road, or, way'
 etc.

Exceptions to this general rule can be found in words derived from monosyllables when the base is the last syllable, and in loanwords of recent borrowing. In these words the stress falls on the ultima, as for example in:

(a) words formed from monosyllables:

- /mardúm/ - 'to chew uncooked rice' from /dum/ - 'munch'
 /pasíp/ - 'quieten' from /sip/ - 'quiet'
 /patúk/ - 'to make to fill the needs' from /tuk/ - 'sufficient'
 Special mention should be made here of the monosyllabic pronouns

and their derivatives, e.g.:

- /an/ - 'that' (i.e. away from the hearer or speaker)
 /i/ - 'that' (i.e. near the hearer)
 /on/ - 'this'
 /tusán/ - 'thereto' (i.e. away from the hearer and speaker)
 /disí/ - 'there' (i.e. near the hearer)
 /disón/ - 'here'
 etc.

(b) loanwords:

- /təkát/ - 'slogan', from Indonesian
 /kapal tərbán/ - 'airplane', from Indonesian
 /ekonomí/ - 'economy', from Dutch
 /parlɛmɛn/ - 'parliament', from English through Dutch.

When otherwise the stress does not fall on the penult, it is to signal a change in content or as in some cases as a result of an ellipsis. One function of the shift of stress is, as above, to contrast a noun with the corresponding 'group VII' modifier. Other examples of this are:

- /gájjan/ - 'length', and
 /gajján/ - 'long'
 /máɔl/ - 'difficulty', and
 /maɔl/ - 'difficult'

Another similar contrast between a verb and its corresponding modifier is also effected by the shift of stress position, e.g.:

- /lápu/ - 'to smear', and
 /lapú/ - 'be smeared'
 /gáttun/ - 'to hang', and
 /gattún/ - 'hanging'

We can find innumerable examples like these. This and the fact that there seem to be no exceptions tend to suggest that this phenomenon follows a rigorous pattern.

Another kind of contrast is seen in the following pair:

/taɲámmu/ and /taɲammú/.

Both mean '*your hand*' in English, but the '*your*' in the first one is informal or familiar '*your*', or the obsolescent '*thy*' (German: *dein*), while the '*your*' of the second is honorific '*your*' (German: *Ihr*). Here again the contrast is effected by the stress position.

The second utterance /taɲammú/ could be considered as the result of an ellipsis of /-na/ from the form /taɲammúna/ '*your hand*'. However, there is a difference between /taɲammú/ and /taɲammúna/. The former is plain respectful speech, and the latter is very formal if used for one person, and both informal and very formal in plural (German *euer* and *Ihr*).

A variation of this last phenomenon is seen in:

/borúm/ -- /borumú/ -- /borumúna/.

All three mean '*your daughter*'. The first '*your*' is familiar, and the last two honorific '*your*', the third being very formal. This variation can be explained morphophonemically by the fact that this kind of '*your* plus modified noun' construction occurs only when the word base, in this instance /boru/ ends in a vowel. This gives us another type of minimal pairs contrasted by the position of the strong stress, as can be seen from the following examples:

/ítom/ - '*black dye*'

/itóm/ - '*your sister, or brother*'

/ótom/ - '*plane*' i.e. carpenter's tool

/otóm/ - '*your folly*'

In this connection it should be mentioned that the personal pronoun /hamú/ '*you*, plural or singular honorific' is stressed on the last syllable and is a variant of the more formal and less frequently used /hamúna/.

When we examine the stresses in a vocative word, we see that they are also not stressed on the penult, but on the last syllable. This furnishes another type of contrast, as e.g.:

/ínaŋ/ - '*mother*'

/inán/ - '*mother!*', as in: Olo, inang. '*yes, mother.*'

/sahála/ - '*a boy's first name*'

/sahalá/ - as in: Ro ho tuson, Sahala '*Come here, Sahala.*'

Other categories of words which have the stress on the last syllable are:

- (1) comparatives ending in /-an/, e.g.:

/ulián/ - 'more beautiful'

/hunián/ - 'more yellow'

This category of words can be paired minimally with words stressed on the penult.

- (2) adjectives with the suffix /-hu/ 'excessively', e.g.

/padokdökkú/ - 'too heavy'

/pahapalhú/ - 'too thick'

These adjectives do not furnish us with minimal pairs.

In all the above words, most of which are disyllables, only one stress, i.e. the strong stress, is indicated, which shows that the other syllable gets a weaker stress or no stress at all. All the examples amount to showing the fundamental fact that stress in Batak is distinctive, hence phonemic.

However, the absence of a stress mark on the other syllables is not meant to imply that they get the same kind of weak stress. Let us, for example, examine the utterance:

/borumuna/ - 'your daughter'.

Here we can hear that /bo-/ is definitely louder than /-ru/ and /-na/, both of which get about the same degree of loudness. This impression of loudness is actually not only produced by the energy of articulation of the syllables concerned, but also by an elongation of the vowel as the syllable nucleus, such that the syllable that is louder also gets the greater elongation. However, as has been said earlier, length is non-phonemic in Batak, and the total predictability of this kind of lengthening also tends to support that statement.

If in the above example we exchange the position of the stresses on the syllables /bo-/ and /ru-/, something, which, though understandable, is not quite Batak will result. Hence it is obvious that also these different degrees of stress are significant features of the language.

Let us now consider this pentasyllabic utterance:

/pardalanion/ - 'journey'.

Here we also encounter three different degrees of stress. The heaviest stress, which we will call primary and represent by /'/, falls on /-ni-/. The next less heavy stress falls on /par-/. This kind of stress we will call tertiary and represent by /`/. The other syllables have about the same degree of weak stress, which we will designate by weak and indicate by /~/ or absence of any stress mark. Thus we can transcribe our last example by:

/pàrdǎlaníǒn/ or /pàrdalaníǒn/.

When we expand one of our examples, thus:

/pàrdalaníǒnnasída/ - 'their journey'

we see that there are two tertiaries, five weaks and one primary there. Ordinarily the stress on /par-/ is slightly stronger than the one on /-ni-/, but they seem to be allophones of the same stress, as, if we give both the same amount of stress or reverse the position of the stresses, it will still sound normal.

Let us now observe the stress phenomena in the word base 'boru' and a few derivatives:

| | |
|-------------------|--|
| /bóru/ | - 'daughter' |
| /borúna/ | - 'his daughter' |
| /parbóru/ | - 'parent of a girl' |
| /bòrunámi/ | - 'our daughter' |
| /pàrborúcn/ | - 'relative on the distaff side' |
| /pàrboruúnna/ | - 'his relative on the distaff side' |
| /pàrborúcnnámi/ | - 'our relative on the distaff side' |
| /pàrborúcnnasída/ | - 'their relative on the distaff side' |

In all these utterances we see that the penult always gets the primary stress, and that the initial syllable, if not penultimate or immediately preceding the penult, gets the tertiary. The other tertiary seems to be placed at such a place that the stronger stresses, namely, the primary and tertiary, are in a rhythmic arrangement with the tertiaries in units of two or three syllables removed from the primary. This rhythmic arrangement, however, is not absolute, for, as we see in the last example, the units can alternate from two or three in one single utterance, if the number of syllables before the penult is not divisible by two or three.

Let us now consider how the stress phenomenon is when we take an utterance that consists of two parts, which are ordinarily called words: ruma 'house' and gorga 'carving'. Read separately, both ruma and gorga have a primary on their first syllables, and a weak on the second. But when we put them together into one unified construction, thus:

/rùma górga/ - 'a house with carved ornamentations'

the strong stress on ruma becomes much weaker than the one on gorga. The stress on the first syllable of gorga remains a primary, but the one on the first syllable of ruma is now tertiary. It is not possible to give those two words the same degree of stress and still mean the same thing. This sequence of two words with a tertiary and a primary is typical of compound nouns consisting of a noun plus a noun modifier. Other examples of this are:

| | |
|--------------|------------------------|
| /hòda pácu/ | - 'race horse' |
| horse racing | |
| /hàta bátak/ | - 'the Batak language' |

/hàlak cína/ - 'a Chinese'

/lòsuŋ áek/ - 'watermill'

The same stress pattern is also found in compound verbs consisting of a verb plus a noun, e.g.:

/màŋaj júhut/ - 'celebrate a wedding'

/màbbuat éme/ - 'to harvest rice'

One compound noun expression to be considered now is:

/àkkat lóbban/ - 'a kind of draughts game'.

This contrasts with:

/àkkat lóbban/ - 'leap across the gorge', i.e. with emphasis on 'gorge'.

Here we have a sequence of a stress that is stronger than a tertiary which we will call **secondary**, and a primary, whereas in the compound noun expression the pattern is tertiary-primary. The normal imperative 'Leap across the gorge' is:

/ákkat lòbban/

with primary on the first syllable of /ákkat/ and tertiary on the first syllable of /lòbban/. The expression with the sequence secondary-primary is used when one wants to emphasise /lóbban/ in meaning that the hearer is not expected to jump over the **shrub**, but across the gorge. This secondary stress is marked by /[^]/. Other such imperatives are:

/búat ín/ - 'take THIS'

/àllaŋ iddáhan/ - 'eat RICE', said for example to a child who eats only his fish.

/àlap ikkáu/ - 'fetch FISH'

This last expression shows minimal, or near minimal contrast with the compound noun, as more often than not the pitch of the last syllable will vary:

/àlap ikkáu/ - 'honorific term for 'tongue'',

which consists of a verb and its object noun.

The occurrence of the secondary can also be seen in phrases like the following:

/gânup lào mäsuk | sài marédde do nasida/

In this sentence a break occurs between /mäsuk/ and /sài/. Within the first half of this sentence there are one primary, one secondary, one tertiary, and three weaks; and in the second half, one primary, two tertiaries and six weaks. In citation form each word would get one primary, but in a context this primary may be reduced to a secondary or tertiary according to the interrelationship of the words, except the function word /do/ which always gets weak stress in context.

In Batak the primary seems to fall on the item of maximum information

as can be seen from the following different renditions of a single sentence:

1. /nùŋŋa láo ibâna/ - 'He has GONE already'
2. /núŋŋa lào ibâna/ - 'He HAS gone ALREADY'
3. /nùŋŋa lào ibána/ - 'HE has gone already'

The normal rendition for a simple statement is the first one.

In ordinary sentences in Batak, contrary to the English pattern which tends to put the primary at the end of each phrase, the primary tends to fall on a word at the beginning of a phrase. This can be seen in the following phrases which happen to be sentences too:

- /na móra do àmanimôpul/ - 'Mopul's father is rich'
 /dak kέα ibàna màrsilecolléon/ - 'He never gives presents'
 /nùŋŋa másak emêna/ - 'His rice is ripe'
 /dàn dipúkka dopè manâbi/ - 'He has not begun to harvest yet'

One other thing needs to be said about the nature of the stress in Batak. It has been stated earlier that stress is the relative loudness that a syllable (or rather, a vowel as its nucleus) gets which contrasts it from other syllables in the utterance. However, other things than mere amplitude are also involved in producing this effect of loudness; one of them, as has also been said earlier, is length. In the section on vowels it has been stated that quantity of a vowel depends on the fact whether it is in an open or closed syllable. Here it should be added that, everything else being equal, a vowel with a stronger stress is also longer than one with a weaker stress; also, that the weaker the stress a vowel gets the more the articulation tends to shift to a central position without, however, ever reaching the central position of a 'schwa'. But, as this kind of lengthening is also totally predictable, it is non-phonemic. The allophonic lengthening and shift of articulation of a vowel¹ can be seen in the following phonetic transcription, in which a dot (·) after the vowel will indicate elongation, and (◀) means fronting:

[ṣ·ṭp̣] - 'foolishness'

[p̣ṭṣ·] - 'foolish'

In the preceding pages some mention has been made about stress pattern; namely, the stress patterns of a compound noun and a unified verb plus object construction, both of which are tertiary-primary. The former will strike an English speaker learning Batak, as it is exactly the reverse of the corresponding compound noun stress pattern in English, which is primary-tertiary for most of the American people.

¹See also Appendix II, Spectrogram 7a and 7b.

Another important stress pattern that may be worth mentioning is that of a 'noun plus modifier' construction (in English the corresponding construction is 'modifier plus noun'), which is secondary-primary, as for example:

/jâbu natíbbó/ - 'a tall house'
 /hâlak nabúrju/ - 'a good person'
 /âñin naúli/ - 'good air'

(This last expression contrasts with:

/âñinnaúli/ - 'name of a place'

in the stress on /âñin/.)

The stress patterns are alike, but the word orders are reversed. The implication of these phenomena for language learning will be discussed in Chapter IV (of the original thesis).

In conclusion of this section a word should be said about the writer's observation of four distinctive stresses and hence his proposition of a four-stress notation. When we observe single and compound utterance units, we can easily distinguish three contrastive stresses: strong, medium and weak. The secondary stress has been introduced above by comparing the stresses in this pair:

/âlap ikkáu/ - 'fetch FISH'
 /âlap ikkáu/ - 'tongue'

The expression for tongue cannot be pronounced otherwise, but the 'fetch fish' expression, though with a different content, can be read with a primary on /âlap/ and tertiary on /ikkàu/, thus:

/âlap ikkàu/.

This is in fact the most frequent pronunciation, as the context is the more common because this is the normal imperative. This is also the case for the other similar examples.

The 'Fetch fish' (with emphasis on 'fish') utterance can also be pronounced, as many people do, with a sequence primary-secondary, thus:

/âlap ikkâu/.

For those people the secondary would be established from the contrast in the second member of the utterances, tertiary in one and secondary in the other.

It should be well to state at this point that the secondary stress has a far smaller functional load than the tertiary, and that it seems to be the least important of the four stresses.

A three stress notation would result from an analysis (Brandstetter 1915:92) along the lines of Van der Tuuk's finding that Batak has no sentence stress ('Satzakzent'). But as can be seen from the above examples, this sentence stress does exist in the writer's speech, by

which the item of maximum information tends to get the primary and by which the syllables getting a primary in words as citation forms often get a reduced strong stress, namely, a secondary or a tertiary. However, it would be easy to have a broad transcription with three different stresses by combining the secondary and tertiary into one 'medium' stress, which would be sufficient for all practical purposes. This would be more easy to do for Batak than for English, because in Batak the secondary stress does not have a great contrastive functional load. But in the writer's opinion a four stress phonemic notation is not an unnecessary complication, on the contrary more clarity is gained by it. Moreover, in Batak stress, accompanied by pitch, is used, though not so extensively as in English, to show prominence in a sentence or to indicate emphasis, as more often word order, but always accompanied by stress, is used to achieve them in Batak. In English only stress is more often used to produce the same effects.

4. JUNCTURE

Most of the examples cited above are single words. Let us now take utterances that consist of more than one word, and written conventionally in two or more separate units.

Adong ma nadjolo sada halak na margoar si Posma

'There was once a man, called Posma'

We can distinctly hear two breaks, or 'junctures', in this sentence. The first and longest break occurs after halak. Its duration is at least twice as long as that of a normal vowel sound, and it is accompanied by a rise in tone (pitch) on the preceding syllable. We will call this, to borrow the English terminology (as in Hill 1958), a 'double bar', and represent it by /||/. The second break occurs after margoar. This break is shorter than the /||/, and there is no appreciable change in the pitch of the preceding syllable perceivable. This juncture will be called, again after the English terminology, a 'single bar', or /|/. The above sentence can also be read with a /|/ after nadjolo. At the end of the sentence there is another break, the longest variety of junctures. It is accompanied by a fall in pitch. This juncture will be called a 'double cross', and represented by /#/. The above sentence can now be transcribed thus:

/adɔŋ ma najôlo sàda hàlak || nà margóar | sipósma#/

The writer has, unfortunately, not been able to measure the exact ratio of the duration of these junctures, but it can be pretty safely assumed that the /|/ is at least as long as the time it takes to pronounce a short vowel, which is about eight centiseconds; the /||/ as long as two normal vowels, i.e. approximately twenty centiseconds; and

the /#/ more than two vowel sounds.¹ These junctures are characterised by pitch; the /|/ by level pitch, the /||/ by rising pitch, and the /#/ by falling pitch. The /#/ by the fall in pitch, gives an impression of finality, and will normally be found at places where in orthography we would expect a period, but this correspondence is far from exact. The other two junctures by the nature of their pitch give an effect of suspense.

The utterances bounded by these junctures are called phrases; and as their occurrence is at the end of phrases, they are called 'terminal junctures.'

A typical place where a /|/ will be found is in normal counting, as for example:

/sáda | dúa | tólu | ópat #/ - *'one, two, three, four.'*

But when we want to count down, as for example, in starting a race, the single bars will be replaced by double bars, thus:

/sáda || dúa || tólu || ópat #/.

The /||/ is normally found at the place where in ordinary spelling we would expect a comma, but here again the correlation is not exact.

Below follow a few sentences to show the occurrence of the terminal junctures:

- (1) /dùk sálpu pítu âri || rò bé ma nàsa pidoŋ | maddapótton làli † #/
'after seven days all the birds came to the chicken eagle'
- (2) /alái | dùŋ adóŋ dùa âri || láo ma babiatì | manôpop bódat #/
'But after two days the tiger went to visit the donkey.'
- (3) /hâlak na pógos || alâi na búrju | jàla na ríngas | dò sipósma #/
'Posma is a poor, but a good and intelligent man.'

Other occurrences of the terminal junctures will be further shown in the section on pitch, as it is pitch sequence together with a terminal juncture that makes intonation.

In the original Batak writings the terminal junctures are not marked. But since the introduction of the Latin alphabet the existence of the terminal junctures has been recognised as evidenced by the use of punctuation marks as the comma, the period, the question mark, etc. However, as mentioned above, many juncture positions have been missed, and the use of those punctuation marks is not consistent phonologically. Let us now consider another type of juncture which is less recognised by Batak orthography.

When we see the following phonemic transcription consisting of segmental phonemes only,

/ríngasnasida/,

¹See also Appendix II, Spectrograms 10 and 11.

we cannot know with certainty what is meant by it, but when we supply the stresses, the utterance takes on meaning. There are two stress patterns possible, namely:

/ringásnasìda #/ 'they are diligent' and

/ringasnasída #/ 'their diligence'

When we listen carefully, these two utterances do not differ in the stresses only, but also in the linear segments. We can hear that the transition from /s/ to /n/ in the first utterance is different from that in the second. In the second example we hear what is termed a smooth transition, whereas in the first one we hear some elongation of the /s/ with a diminution of energy, which a Batak speaker would interpret as some kind of break. This is called open transition (Gleason 1955:42) and is rendered in phonemic notation by /+/. This is called a 'plus juncture'. This juncture, in contrast with the terminal junctures, is called an internal juncture,¹ since it occurs only within a phrase. Its main occurrence is at the boundaries of lexical items or words, although it can and does often occur within words conventionally written as units, and conversely it is often absent between words orthographically written as two units. For example:

hodokna /hodók+na/ 'his sweat'

tungkok i /túkkot í/ 'the stick'

A contrast also occasioned by the different position of the /+/, is seen in the following utterances:

/ásam + ùli + jumôlo #/ 'Estimate the beauty first.'

/ása + múli + jumôlo #/ 'In order to marry first.'

In ordinary speech we can hear an elongation and a laxer closure of the /m/ of /ásam/ in the first sentence compared with the /m/ of /múli/ in the second one. Though in ordinary conversational speed a Batak tends to connect a final consonant with a following initial vowel, the transition between /m/ and /u/ in the two sentences is noticeably different. There is also some difference noticable in the quality of the /a/ preceding /m/. The /a/ in the second utterance is higher and less open. These two utterances are thus contrasted not only by the different stress patterns, but also by the different position of one /+/ in the linear sequence of phonemes.

A contrast can also be seen in the following pair:

/opátsa + tâli #/ 'Bind all four'

/ópat + satâli #/ 'Four for a quarter.'

In the second utterance the /t/ and /s/ are more distinctly pronounced, whereas in the first the transition is so smooth that they resemble the

¹Dr. A.A. Hill in class lectures.

sound of the alveopalatal affricate /c/, but still remain different from it.

The tendency of the /+/ to occur at the traditional word divisions without having an exact correlation, has been mentioned above. The phonological arbitrariness of the Batak orthography can also be seen from the following pair:

di surat /disúrat #/ 'in a letter'

disurat /disúrat #/ 'written'

These two utterances are exact homophones, but in conventional spelling one is separated and other joined. It is, however, true that they have different semantic content.

In /rùma + górga #/ 'carved house'

the /+/ seems to have the effect of increasing the force of the explosion of the following /g/ of /gorga/. This can be compared with the /g/ in, for example,

/agóan #/ 'missing something'.

In reduplications and interative words which are in traditional spelling written as units, we usually have a /+/ separating the two elements, for example:

holangholang /hòlak + kólan #/ 'interstice'

hulinghuling /hùlik + kúlin #/ 'hide, n.'

parikparik /pàrik + párik #/ 'small earthen wall'

siuksiuk /sìuk + síuk #/ 'small ladle'

From these examples one feature of the /+/ can be noted, namely, that assimilation can take place across a /+/, so that the plus juncture is situated in between the assimilated and the assimilating sounds. This can be clearly seen in cases where the final sound of a word is assimilated with the initial sound of a following one, for example:

/sumán/ 'resemble' and /tusí/ 'to it' become:

/sumât + tusí #/ 'like it'

/ópat/ 'four' and /rátus/ 'hundred' become:

/òpak + rátus #/ 'four hundred'

/mífan/ 'dwell' and /sàda + hálak/ 'a man' become:

/mífas + sàda + hálak #/ 'a man dwells'

In the writer's speech, in words with a double consonant as a result of affixation, there are two possible renditions, one with, and the other without a plus juncture in between the two identical consonants. Thus he can have:

/páttu/ or /pát + tu/ 'my foot'

/ròhakkí/ or /ròhak + kí/ 'my mind'

/asómma/ or /asóm + ma/ 'his lime'

The above examples have shown that the nature of the plus juncture

can be seen in elongation of a preceding phoneme, and as a break in between two phonemes. These facts tend to show that the /+/ is a tempo phenomenon. But it is not only that, as it produces also a modification in the articulation of the bordering segmental phonemes, as for the example, the more energetic onset of the following phoneme. The writer has not been able to collect enough evidence of other manifestations of the existence of the plus juncture, but he suspects that there are still other details that have missed his observation.

5. PITCH

Pitch has often been mentioned in the previous sections. Pitch is the musical tone of the voice in speech. Thus only the voiced sounds of an utterance can have pitch, especially the vowels, although some of the consonants can have pitch too. In this analysis the pitches, as the stresses have been, will be attached only to the vowels. In speaking about pitch here, we think of the relative high and low of the tones of a voice rather than the absolute tone-pitch in terms of frequencies.

The normal pitch - normal, taken here in its quantitative meaning - of the voice of a speaker will be called pitch /2/. The lowest pitch the voice has in ordinary conversation will be called pitch /1/, and the pitches higher than /2/ will be called respectively pitch /3/ and pitch /4/. The most frequently occurring pitches are /1/, /2/, and /3/.

In the following utterance we see the occurrence of the three lower pitches:

/dijábu + dò + ibâna #/ 'He is at home.'

The utterance starts with the normal pitch /2/; rises to /3/ for the /-ja-/; falls again to about a little above the level of /2/; keeps gradually falling to about pitch /2/ on /dɔ/, and a little below /2/ on /i-/; rises again sharply to a little above the initial pitch, but not reaching the pitch of /-ja-/; then falls rapidly to /1/ for the last /-na/ and on to silence. When we use the symbol [+] after the pitch number for a higher degree, and a [-] for a lower degree of pitch, and the symbol [↘] for the rapid final fall, we can transcribe the above utterance, phonetically as regards the pitch, as follows:

2 3 2+ 2 2-2+1 ↘

[dijábu + dò + ibâna #]

The transition from one pitch level to another is not abrupt like steps but rather like that found in waves, with differing degrees of steepness, but never perpendicular.

Before proving the validity of the phonemic status of the pitches by their contrastive distribution, a statement will be made here, that there are only three phonemic pitches in the above utterance, based

on the fact that the writer as a native speaker of Batak does not find any significant difference between the pitches of the syllables marked with varieties of /2/, so that now we may transcribe the above example phonemically, thus:

2 3 2 1
/dijábu + dɔ + ibâna #/

or

2 3 2 1
/dijábu + dɔ + ibâna #/

or

2 3 2 2 1
/dijábu + dɔ + ibâna #/

The second notation seems to be preferable, and will be used further in this study, because it is simple and shows better the close relationship between stress and pitch by marking the pitch of the last stronger stressed syllable. In this notation the first pitch point is the first syllable, the second pitch point the syllable with primary, the third comes on the syllable with the last secondary or tertiary, and the fourth on the last syllable; with the understanding that after the first pitch /3/, the pitch always falls back on /2/, except, if the following pitch point has also a /3/. This kind of immediate falling after reaching the peak is also true when the peak is pitch /4/. This number of four pitch points may be reduced to three or two depending on concurrences of two or three of these pitch points, as can be seen in some of the examples in the following pages.

In the above example we have seen that the pitch sequence or pitch pattern is /2321/. A pitch sequence followed by a terminal is called a 'pitch contour' or intonation. This pitch contour /2321#/ is typical of statement sentences, as illustrated in the following examples:

2 3 2 1
/màrmɛámi + dɔ + nasɪda #/ 'They are playing'¹
2 3 2 1
/maddùda + éme + dɔ + inâtta #/ 'Mother is pounding rice'
3 2 1
/dúda + dɔ + hodâna #/ 'He has two horses'
3 1
/ʃɪɔ #/ 'Yes'

Let us take this last example and give it different pitch sequence and see what it does to the content. The example above is a statement

¹See also Appendix II, Spectrogram No.14.

of willingness, or an affirmative answer. When we give a pitch /2/ on the second syllable, usually accompanied by a prolongation of the last phoneme and sustension of the pitch level marked here by the single bar, thus:

3 2
/ʃɪɔ |/,

it will mean something like '*Yes, I know, stop bothering me*', or '*I said yes already, have a little more patience*'. It expresses some, though sort of good-humoured, impatience or irritation. Real impatience or irritation would be expressed thus:

32 32 32
/ʃɪɔ |/, or /ʃ | ɪɔ' #/.

The greater the impatience or irritation, the stronger the stress on the /-ɪɔ/ would be, and the more it would be prolonged.

Another possible pitch contour is:

3 3
/ʃɪɔ ||/.

This means '*Did you say yes?*'. The /3/ on /-ɪ/ is accompanied by a slight rise in pitch, here indicated by the double bar. This is typical of a question to verify an answer by repeating it. This contour can also express a mild surprise at the answer.

A great surprise like '*DID you say yes?*', introduces us to another pitch level, higher than /3/, namely, pitch /4/, which is often accompanied by overloudness, thus:

4 3
/ʃɪɔ ||/.

The different renditions of /ʃɪɔ/ above show that in Batak a difference in pitch produces a difference in content, in other words, that pitch is distinctive or phonemic, and that four pitch levels are adequate to signal these differences. They have furnished minimal contrasts of the four pitches to establish their phonemic status.

Pitch is the most overt feature of the suprasegmental phonemes. Pitch contour habits seem to be the most persistent of the language habits, and pitch contours will pose one of the hardest obstacles in adult language learning. Consequently, a thorough knowledge of the pitch contours is very important, also for pedagogical reasons.

As mentioned before, some pitch contours are typical for some kinds of sentences or phrases. We have seen that /2321#/ is typical of statement sentences. The typical pitch contour of an ordinary question can be seen from the following interrogative sentences:

2 3 2 3
/dijábu + dɔ + ibâna ||/ '*Is he at home?*'¹

¹See also Appendix II, Spectrogram No.15.

The pitch contour is /2323||/. Other illustrative sentences, taken from the previous examples for statements, are:

2 3 2 3
/màrméámi + dɔ + nasída || / 'Are they playing?'
2 3 2 3
/maddùda + éme + dɔ + inâtta || / 'Is mother pounding rice?'
3 2 3
/dúá + dɔ + hodâna || / 'Does he have two horses?'

The answer to the first question can be:

2 3 1
/dijábu #/¹ 'Yes, he is', or literally, 'in the house'

This is a reduced notation of the typical statement contour /2321#/, because the syllable with the primary, the normal pitch point of /3/, is also the last stressed syllable in the utterance. Another reduction of the statement contour has been shown in one of the renditions of /ólɔ/ with a pitch contour /31#/, due to the concurrence of the first three pitch points on /ɔ-/.

The first interrogative sentence can also be read in this way:

2 2 3 3
/dijábu + dɔ + ibâna || /² 'Is HE at home?'

Here the pitch contour is /2233||/, the first five syllables being level on /2/, though the primary still remains on /-ja-/. The /-ba-/ of /ibâna/ has a pitch /3/, and the stress must be a secondary, whereas in the /2323||/ contour it could have a tertiary stress without occasioning any difference in content. As mentioned above, the syllable with a primary stress, where we would normally expect a /3/, gets a pitch /2/. This type of contour puts the /ibâna/ into relief, and thereby signals the emphasis placed on it.

In the ordinary question contour prominence is given to /dijábu/. If we want to put the special emphasis on 'home', the contour becomes /2333||, thus:

2 3 3 3
/dijábu + dɔ + ibâna || /³ 'Is he at HOME?'

Here the pitch remains on /3/ from /-ja-/ to the end. This pitch contour can also express mild surprise.

The answer to this question can have the /231#/ contour or the

¹See also Appendix II, Spectrogram No.18.

²See also Appendix II, Spectrogram No.16.

³See also Appendix II, Spectrogram No.17.

emphatic /241#/ , thus:

2 3 1 2 4 1
/dijábu #/, or /dijábu #/.

When we want to show doubt as to the veracity of the answer we use this contour:

2 4 3 3
/dijábu + dɔ + ibâna || / 'IS he in the house?'

in which the peak is pitch /4/ on /-ja-/ and falls to /3/ on /-bu/, and remains on that level to the end. The answer to this would normally be emphatic, thus:

2 4 1
/dijábu #/¹ 'He IS.'

Further doubt to this answer, or irony, could be expressed with a /242|/ contour, thus:

2 4 2
/dijábu |/² 'Really?'

This type of contour can be often heard also in little girls' speech, when they want to show off something and tease another child with it. If this would occur in a situation in which the child has a new toy in the house, then it would express something like 'Well, I do have it, but it is in our house. YOU don't have one, and YOU may not see it, but I can go in and play with it as much as I like to.' The difference between irony and a little girl's teasing will be mostly signalled by the context and the paralinguistic kinesics.

In a statement sentence that consists of more than one phrase, we have linked contours. For example:

2 3 2 2 2 3 2 1
/dijábu + dɔ + ibâna | alái + dák + karêjo #/
'She is in the house, but she is not working.'

Here the end pitch of the first phrase is the same as the pitch at the beginning of the second phrase, with the single bar in between. The single bar juncture will be often found between two identical pitches.

In utterances, traditionally spelled as single sentences, we can have more than one #, for example:

2 3 2 1 2 3 1
/dijábu + dɔ + inâta # nínna + ibána #/
'Mother is at home, he says.'

¹See also Appendix II, Spectrogram No.19.

²See also Appendix II, Spectrogram No.19.

If we turn this sentence around, we can get the same pitch contours,

2 3 1 2 3 2 1
/nìnnà + ibána # dijábu + dɔ + inàtta #/.

When we give another contour to the first part, thus:

2 3 2 2 3 2 1
/nìnnà + ibána dijábu + dɔ + inàtta #/,

it means '*HE says, that mother is at home*', with the stress on '*he*', with or without the implication that the speaker knows that it is not true, but by which, at any rate, the speaker wants to indicate that he does not want to commit himself.

An example of linked contours with a question in the first part is:

2 3 2 3 2 3 1
/dijábu + dɔ + inàtta nìnnà + ibána #/
'Is mother at home, asked he.'

Linked pitch contours of a question are found, for example:

3 2 3 3
/ʒɔ nímму + dɔ || / 'Did you say yes?'

To express surprise at the 'Yes' answer given the question would have the following contour (see also p.35):

3 3 3 3
/ʒɔ | nímму/ 'DID you say yes?'

When we consider the following possible answer to a question like '*What are you doing there?*',

3 21 1 1
/dán + adɔŋ | ànggiá #/ 'Nothing, brother',

we see that the pitch contour is continued in a monotone beyond the single bar in the vocative material. Here the vocative material has a continuous pitch /1/, with a slight upturn on /-gi-/ and /-a/, followed by a fall so rapid that the voice vanishes and we hear only breath.

Vocative material, monotonous except for the pitch fall of the # juncture, can be seen in this impatient answer:

2 32 2 2
/dân + adɔŋ | rajanami #/ 'Nothing, SIR'

But the vocative material can also have a two-pitch contour. This is the type of pitch contour of a polite answer.

3 22 2 2
/dân + adɔŋ | rajanami #/ 'Nothing, sir.'

Here the pitch gradually falls from /2/ on /ra-/ to /1/ on the last syllable /-mi/.

In ordinary requests the pitch contour ends in /1/, followed by a /#/ , as, for example:

3 21

/rǒ + hamù + tusôn #/ 'Please, come here.'

The difference of this pitch contour from that of a peremptory command is that the fall at the last syllable is more gradual in a request, which tends to cause a prolongation of that syllable. Other means are employed to soften a request, as the use of function words, as /jǒlɔ/, and the toning down of the stresses. A corresponding command to the above request would be:

3 21

/rǒ + hɔ + tusôn #/ 'Come here!'

with a much shorter last syllable.

Above we have seen examples of the occurrences and functions of some categories of sentences. Two more varieties will be discussed below. One is the contour of a question, with an interrogative word helping to signal the question.

2 3 2

/marhúa + hamù + disǐ #/ 'What are you doing there?'

2 3 2

/behá + dɔ + ulàommunaǐ #/ 'How is your work getting on?'

This is the ordinary contour for a civil question, /232#/ or /2322#/. A peremptory question will have a /2321#/ , for example:

2 3 21

/marhúa + hɔ + disòn #/ 'What are you doing here?'

The brusqueness will also depend on overloudness of /-hu/ and the shortness of the syllable /-ɔn/.

Although we can reduce both the first question above and the last one into a one-word question, they will still have the same contour, if the implication is the same, thus:

2 32 2 321

/marhúa #/ and /marhúa #/.

The second question can also be reduced to:

2 32 2 31

/behá #/ 'how' (polite) and /behá #/ 'how' (brusque)

This brings us to an expression of mild surprise commonly employed in conversation, sometimes as a polite interjection,

2 32

/beháǐ |/

which would mean something like, 'Why, is that so?'

We have already seen a few characteristic pitch contours of Batak. Of course not every possibility has been explored above, as there are still other types of sentences, other contexts, and some other shades of the ones mentioned above, which have not been shown. However, the

more important and frequent pitch contours have been included in the discussion.

CHAPTER II

PHONOTACTICS

The term phonotactics is defined as the "area of phonemics which covers the structural characteristics of sequences" (Hill 1958:68). Special attention will be given to sequences of phonemes of the same category, also called clustering, i.e. consonants followed by consonants, vowels followed by vowels, etc. The greater part of this section will concern clusters of segmental phonemes.

In the appropriate sections we have seen that the four pitches cluster with each other and with themselves, and that also the four stresses freely cluster with each other and themselves, except the primary which does not cluster with another primary in the same phrase.

2.1. WORD PATTERNS

The word patterns of Batak will be considered here only in terms of the segmental phonemes. When we use V for vowel and C for consonant, we can summarise the syllable and word patterns of Batak, giving a few examples of each, as follows:

SYLLABLE PATTERNS (The hyphens indicate syllable divisions.)

- (1) V; as in: /há-u/ 'tree'
 /ú-la/ 'do'
 /á-i/ 'for, conjunction'
- (2) CV; as in: (see above).
- (3) VC; as in: /mí-an/ 'dwell'
 /á-in/ 'adopt'
- (4) CVC; as in: /pás-tap/ 'slap'

WORD PATTERNS. Only the patterning of the segmental phonemes of word bases will be given:

- (1) V; as: /i/ 'that, near the hearer'
 /o/ 'Oh!'
 /e/ 'Hey!'

These are the three monophonemic words and minimal utterances in Batak.

- (2) VC; as: /an/ 'that, away from hearer and speaker'
 /ɔn/ 'this'
 /ak/ 'hip'
- (3) CV; as: /tu/ 'to'
 /sɔ/ 'quiet'
 /ra/ 'perhaps'
- (4) CVC; as: /pat/ 'foot'
 /gok/ 'full'
 /sar/ 'widely known'
- (5) VCV; as: /íla/ 'shame'
 /ádu/ 'overtake'
 /úpa/ 'reward'
- (6) CVV; as: /háu/ 'tree'
 /sía/ 'nine'
 /píio/ 'invite'
- (7) CVVC; as: /tíop/ 'hold'
 /táit/ 'pull'
 /súan/ 'plant, vb.'
- (8) CVCV; as: /dúda/ 'pound'
 /dósa/ 'sin'
 /bátu/ 'stone'
- (9) VCCV; as: /úmma/ 'kiss'
 /árta/ 'treasure'
 /óppu/ 'grandparent'
- (10) CVCVC; as: /húdon/ 'pot'
 /pídon/ 'bird'
 /pógos/ 'poverty'
- (11) CVCCV; as: /búrju/ 'goodness'
 /sɔrtá/ 'innocent'
 /páksa/ 'compel'
- (12) CVCCVC; as: /tíktiŋ/ 'banns'
 /pistár/ 'clever'
 /palták/ 'clearly visible'

Trisyllables are not so numerous. The extreme pattern would be CVCCVCCVC, but the writer has not been able to find such a word. The existing words have one or more, up to five, of the consonant positions left vacant. A minimal trisyllable consists of three vowels and one consonant, for example:

- CVVV, as: /baóa/ 'male person'
- VCVV, as: /agía/ 'however, adv.'
- VVCV, as: /aúga/ 'yoke'.

The form VVVC does not seem to exist. A few more examples of trisyllables are the following:

CVCVCV, as: /sahála/ 'might, .'

VCCVCV, as: /appára/ 'brother'

VCVCV, as: /alógo/ 'wind'

CVCVV, as: /satúa/ 'field mouse'

CVVVC, as: /nuáen/ 'now'.

Tetrasyllabic word bases do not seem to be a regular feature of Batak structure. The writer has been able to find only one example, namely,

CVCVVVC, /saraśal/ 'trousers',

which word he often pronounces /saráwal/, thus making a trisyllable of it. He is, unfortunately, not acquainted with its history.

2.2. CONSONANT CLUSTERS

From the examples we have had so far, we can see that any consonant seems to be able to have any vowel after it, although, as we have seen in the section on vowels on p.15 ff., some consonants show a preference but inconsistently, for /e/ or /ɛ/, and for /o/ and /ɔ/. But there are limitations as to the sequence vowel-consonant in the same syllable, because, as mentioned in the section on consonants, there are consonants which never occur finally in a word.

Summing up the phonotactical distribution of the consonants as mentioned in the section on consonants, we have:

- (a) all the consonants can occur initially and medially,
- (b) the consonants that can occur in final position are: /p, t, k, s, (h), m, n, ŋ, r/
- (c) the consonants that do not occur in final position in a word are:
 - (1) the voiced stops: /b, d, g/;
 - (2) the affricates: /c, j/; and
 - (3) the semivowels: /w, y, h/, with the provision that /b, d, j/ can occur finally in a non-final syllable, when the following syllable begins with /b, d, j/ respectively.

The above distributional characteristics of the consonants will be seen in the clusters.

The most frequently occurring syllable in Batak is the CV-type. This fact is reflected, apparently, in the biphonemic character of the Batak alphabet.¹

In Batak the consonants do not cluster in one syllable. There seems

¹See Appendix III.

to be a relationship between this fact and (1) the Batak language habit of pronouncing the vowels rather clearly and distinctly; the centering of unstressed vowels never coming so close to the central neutral vowel [ə], to cause them to lose their identity, which would have meant one step towards the clustering of consonants through a slurring and eventual dropping of a vowel, and, (2) the syllabic character of the Batak alphabet, which tends to help preserve the great syllabicity of the Batak language.

Another way of describing the clustering habit of Batak consonants, is by saying that the consonants do not cluster initially and finally. This fact is also reflected in the Batak language habit in borrowing foreign words which have such clusters, namely, by inserting some vowel in between the component consonants; for example:

Dutch: 'stoker' becomes Batak: /sitókkar/ 'bus conductor'

Dutch: 'klas' becomes Batak: /kalás/ 'class'

or by omitting one of the consonants, as:

Dutch: 'motorfiets' becomes Batak: /mòtɔrpís/ 'motorcycle'

Dutch: 'band' becomes Batak: /ban/ 'tire'.

This means then that in Batak there are only medial, or intervocalic, clusters, for example:

/pástap/ 'slap'.

All the intervocalic consonant clusters consist of only two components and the clustering is always to be regarded as a sequence of the final consonant of one syllable and the initial consonant of the following one.

If any combination of any two consonants were possible, but taking into consideration the limitation of the non-occurrence in final position of seven of the consonants, we would have 170 different clusters. But the number of actual clusters is much smaller, namely 69.

In Batak there are 11 identical consonant clusters, i.e. clusters consisting of two identical consonants. They are phonemic, as can be seen from the following pairs, so that they are not to be regarded merely as an elongation.¹ Those pairs are:

/túlaŋ/ 'uncle' and

/túllaŋ/ 'stab',

/hádaŋ/ 'carry suspended from the shoulder' and

/háddaŋ/ 'fence',

/úma/ 'ricefield' and

/úmma/ 'kiss',

/tópa/ 'forge, (work of a smith)' and

/tóppa/ 'create'.

¹See also Appendix II, Spectrograms No.20 and No.21.

Another evidence that shows that they are clusters and not elongation is trying to segment them by inserting a juncture, a /+ / or /|/. Let us take the third pair, thus:

/úm + a/ or /úmmmm + a/, or

/ú + ma/ or /ú + mmmma/.

We could make the /m/ in both segmentations very long, but they would never be taken as /úmma/ by a native speaker, even though the duration of the single /m/ is longer than the double /m/. Only if the segments are /um + ma/ will they mean 'kiss'.

A list of the intervocalic clusters as found mostly in word bases is given below. Other combinations are possible, at least in careful speech, but always with a distinct /+ / intervening. Such clusters are not included. The ones listed below are what is called genuine clusters, i.e. clusters that can occur without an intervening /+ /. The examples given are all native words with the exception of a few well-established foreign ones if their inclusion is inevitable. The 'foreign', that is, not original but well-established and universally used, clusters will be enclosed between parentheses. If the writer has been able to find only one word where the cluster occurs, it will be indicated by an asterisk placed to the right of the cluster, thus, (/ -nc -/) means 'foreign, but well-established', / -nt -/* means 'one occurrence', and (/ -nt -/)* means 'foreign, one occurrence'.

The clusters are:

| | | |
|-------------|--------------|----------------------|
| / -bb -/ | : /tábba/ | 'increased' |
| / -dd -/ | : /háddaŋ/ | 'fence' |
| / -kb -/ | : /búkbak/ | 'throbbing' |
| / -kd -/ | : /dakdának/ | 'child' |
| / -kg -/ | : /gukgúk/ | 'heaped up' |
| / -kj -/ | : /júkjuk/ | 'touch with a stick' |
| / -kp -/ | : /púkpak/ | 'noisy' |
| / -kt -/ | : /tíktiŋ/ | 'banns' |
| / -kk -/ | : /dákka/ | 'branch' |
| / -km -/ | : /mɔkmók/ | 'fat' |
| / -kn -/ | : /nékneɸ/ | 'pacified' |
| / -kŋ -/ | : /ŋákŋat/ | 'devour' |
| / -kl -/ | : /láklaɸ/ | 'rind' |
| / -kr -/ | : /rékreɸ/ | 'shelf' |
| / -ks -/ | : /páksa/ | 'compel' |
| (/ -pt -/)* | : /sáptu/ | 'Saturday' |
| / -pp -/ | : /óppu/ | 'grandparent' |
| / -ps -/* | : /ɔnópsa/ | 'all six' |
| / -tt -/ | : /páttar/ | 'floor' |

| | | |
|-----------|-------------|-------------------------|
| /-ts-/* | : /ɔpátsa/ | 'all four' |
| /-sp-/ | : /púspas/ | 'noisily busy' |
| /-st-/ | : /pástap/ | 'slap' |
| (/-sk-/*) | : /páska/ | 'Easter' |
| /-sm-/ | : /pósma/ | 'a boy's name' |
| /-sn-/ | : /ásna/ | 'a girl's name' |
| /-sq-/ | : /ɣɔsqós/ | 'worrying' |
| (/-mp-/*) | : /lámpu/ | 'lamp' |
| /-mm-/ | : /úmma/ | 'kiss' |
| (/-nt-/*) | : /tínta/ | 'ink' |
| (/-nc-/*) | : /pánci/ | 'pan' |
| /-nn-/ | : /nínna/ | 'he says' |
| /-nb-/ | : /bunbúɣ/ | 'wide open, vertically' |
| /-nd-/ | : /dándan/ | 'pay heavily' |
| /-ng-/ | : /tángga/ | 'ladder' |
| (/-nk-/*) | : /bánkku/ | 'bench' |
| (/-nt-/*) | : /son tí/ | 'awe-inspiring' |
| /-nm-/* | : /máɣman/ | 'vow' |
| /-nn-/ | : /nónnon/ | 'immerse' |
| /-ng-/* | : /núɣga/ | 'already' |
| /-nl-/* | : /láɣlaɣ/ | 'uninhabited place' |
| /-nr-/ | : /ríɣriɣ/ | 'joint' |
| (/-ns-/*) | : /báɣsi/ | 'adz' |
| /-nj-/ | : /jɔɣjɔɣ/ | 'stand' |
| /-jj-/ | : /éjjak/ | 'mount something' |
| /-lb-/ | : /túlbas/ | 'replenish' |
| /-ld-/ | : /dáldal/ | 'bee' |
| /-lg-/ | : /balgá/ | 'big' |
| /-lj-/ | : /jáɣlaɣ/ | 'chop' |
| /-lp-/ | : /talpé/ | 'on the brink' |
| /-lt-/ | : /palták/ | 'clearly visible' |
| /-lm-/ | : /jɔɣlma/ | 'man, generic' |
| /-lq-/ | : /talɣá/ | 'uneven' |
| /-ls-/ | : /sóɣsɔɣ/ | 'repentance' |
| /-lh-/ | : /púlha/ | 'break dam to drain' |
| /-ll-/ | : /túɣlaɣ/ | 'stab' |
| /-rb-/ | : /súrbu/ | 'burn' |
| /-rd-/ | : /sóɣdak/ | 'lock' |
| /-rg-/ | : /gɔɣrga/ | 'carve' |
| /-rp-/ | : /gáɣpu/ | 'fork' |
| /-rt-/ | : /árta/ | 'treasure' |
| /-rm-/ | : /sóɣrmin/ | 'reflect' |

| | | |
|------------|------------|---------------------|
| /-rn-/ | : /burnán/ | 'puffed up' |
| /-r -/ | : /sírnom/ | 'lost, unfortunate' |
| /-rs-/ | : /ársak/ | 'sorrow' |
| /-rh-/ | : /sórha/ | 'spinning wheel' |
| /-rj-/ | : /burjú/ | 'good, kind' |
| /-rr-/* | : /pírrik/ | 'whistle, n.' |
| (/-rk-/) : | /parkákas/ | 'tool'. |

One cluster has been excluded above, because it needs some explanation, namely:

| | | |
|---------|-------------|---------|
| /-ss-/* | : /massádi/ | 'cease' |
|---------|-------------|---------|

This is the only occurrence of it in the writer's idiolect, but in the speech of many people it will be found in words where the speaker also uses /-c-/ in medial position.

A tabulation of these clusters is presented in p.49. It gives us an overall view of the intervocalic clustering patterns of Batak.

Treating all the clusters, native or foreign, frequent or rare, on the same level, we can observe several things from the table. The patterning is not so regular as to be easily reduced to neat formulas, but the following structural statements can be made.

In terms of the selection of the second component, the patternings seem to be as follows:

- (1) Voiced stops can have only an identical voiced stop in position 2.
- (2) Voiceless stops can take only an identical voiceless stop, /s/, or a /t/, as a second component, with the exception of /k/, which also takes other consonants (see below).
- (3) The fricative /s/ takes any voiceless stop, the fricative /s/, and all the nasals.
- (4) The nasals take only a homorganic voiceless stop, with the exception of (/nc-/) and the single occurrence of /-ŋt-/*.
- (5) The nasals take only an identical nasal as a second component, with the exception of /ŋ/, which can take the other nasals also.
- (6) Clusters with a lingual in position 1 constitute the biggest class. They seem to be more easily stated in terms of the phonemes they do not take in position 2:

(i) /l/ takes all consonants, except /k/, /c/, /n/, and /r/. The non-occurrence of /k/ here is phonotactically evident, as it does not occur natively in initial position. The non-occurrence of /c/ can be understood because of the peculiarity of its distribution, i.e. in medial position it acts as a double consonant, and in initial position only in loanwords. The non-occurrence of /n/ and /r/ after /l/ can perhaps be explained by their phonetic similarity, which seems to account for the

ease of their assimilation. This is in fact what happens when in a sequence of /n/ and /l/ an intervening /+/ is dropped, so that /-l+n-/ becomes /-ll-/, and /-n+l-/ becomes /-ll-/, for example:

/bál+na/ > /bálla/ 'his ball'
 /làon+láon/ > /làolláon/ 'eventually'

This is also true in the case of /r/ and /l/, so that /-r+l-/ becomes /-ll-/, for example:

/pór+lak/ > /póllak/ 'garden'
 /mar+lájja/ > /mallájja/ 'carry with a pole'

But in the case of /r/ and /l/, the intervening /+/ is more frequently preserved.

(11) /r/ takes any consonant except /l/ (see above), and /c/, though in very recent borrowings we can see words like

(/árca/) 'museum'.

But due to the highly specialised character of such words and hence their rare occurrence in everyday speech, this cluster is not included in the list. Moreover, in popular words, /-rs-/ is substituted for it, as in:

elegant: /kárcis/ > ordinary: /kársis/ 'ticket'

(7) Another big class is made up of the velars /k/ and /ŋ/ in position 1. They take all consonants as a second component except /c/ and /h/. It should be mentioned here that in clustering before voiced consonants - i.e. voiced stops, nasals and linguals - only the glottalised allophone of /k/, i.e. [k̚], occurs.¹

Another thing about the /-kX-/ clusters (here /X/ stands for any possible consonant), is the fact that most of the examples are reduplications, due to the fact that the glottalised /k/ is a regular morpho-phonemic alternant for the voiceless stops /p/ and /t/ before voiced consonants, and for the velar nasal /ŋ/ before voiceless consonants.

In terms of the phonemes which can occupy position 2, we can make these statements:

- (1) The most frequently occurring consonants in position 2 are the voiceless stops, the fricative /s/, and the nasals.
- (2) The voiced stops occur only in identical clusters, and after /ŋ/ and the linguals.
- (3) /c/ occurs only after /n/, i.e.: (/nc-/).
- (4) The glottal fricative /h/ occurs only after a lingual.
- (5) The linguals occupy position 2 only in identical clusters.

¹See page 9.

DIAGRAM 4
 DIAGRAMMATIC TABLE OF THE CONSONANT CLUSTERS

Underlining, e.g. nc, means 'foreign, but established'.

An asterisk, e.g. ts*, means 'single occurrence'.

Horizontal index = second component.

Vertical index = first component.

| | b | d | g | j | p | t | k | c | m | n | ŋ | s | h | l | r |
|---|----|----|----|----|------------|------------|------------|-----------|----|----|----|-----------|----|----|-----|
| b | bb | | | | | | | | | | | | | | |
| d | | dd | | | | | | | | | | | | | |
| g | | | — | | | | | | | | | | | | |
| j | | | | jj | | | | | | | | | | | |
| p | | | | | pp | <u>pt*</u> | | | | | | ps* | | | |
| t | | | | | | tt | | | | | | ts* | | | |
| k | kb | kd | kg | kj | kp | kt | kk | -- | km | kn | kŋ | ks | -- | kl | kr |
| c | | | | | | | | — | | | | | | | |
| m | | | | | <u>mp*</u> | | | | mm | | | | | | |
| n | | | | | | <u>nt*</u> | | <u>nc</u> | | nn | | | | | |
| ŋ | ŋb | ŋd | ŋg | ŋj | -- | <u>ŋt*</u> | <u>ŋk</u> | -- | ŋm | ŋn | ŋŋ | <u>ŋs</u> | | ŋl | ŋr |
| s | | | | | sp | st | <u>sk*</u> | -- | sm | sn | sŋ | ss* | | | |
| h | | | | | | | | | | | | | — | | |
| l | lb | ld | lg | lj | lp | lt | -- | -- | lm | -- | lŋ | ls | lh | ll | -- |
| r | rb | rd | rg | rj | rp | rt | <u>rk</u> | -- | rm | rn | rŋ | rs | rh | -- | rr* |

The structural statements above contain some exceptions and show some patterning imbalances. However, this is not very uncommon in language analyses, as there seem to be few, if any, one hundred per cent rigorously symmetrical structural patterns in language. But on the whole the structural patterning of the Batak consonant clusters can be said to be rather regular.

2.3. VOWEL CLUSTERS

As we have seen above the Batak consonant clusters are rather limited, not so much in variety, but in their actual occurrence in word bases and especially in type, as there is only one type, namely the two-consonant intervocalic cluster. Consonant clustering is a much more frequent, more complex and more important feature in English. But on the other hand vowels do not cluster in English, whereas in Batak vowel clustering is a frequent and important linguistic feature.

In the segmental patterning of words at the beginning of this chapter we have seen that in Batak there are sequences of two or three vowels in one word. In this section the two-vowel clusters will be discussed first.

As with consonant clusters above, we will concern ourselves only with genuine vowel clusters, i.e. sequences of vowels without an intervening /+/. By their very nature as syllable nuclei the vowels will be in different syllables.

A list of the two vowel clusters is presented below. Only original clusters, i.e. clusters found in word bases, are listed; the clusters resulting from affixation will be discussed later. The examples are chosen in such a way that they will show the different occurrences of the clusters with regard to their position in the word, initial, medial, or final. One example will be given for each type of vowel cluster according to position. If no example is found for some position, it will be indicated in this way:

- (i) * before a cluster means 'non-occurrence in initial position',
- (ii) * after the cluster means 'non-occurrence in final position',
- (iii) parentheses (...) means 'non-occurrence in medial position'.

LIST OF TWO-VOWEL CLUSTERS:

/a/ in position 1:

| | | | |
|---------|---|--------|-----------------------|
| /au/ | : | /áup/ | 'wash away' |
| | | /ráut/ | 'knife' |
| | | /báu/ | 'odour' |
| *(/ao/) | : | /dao/ | 'far' |
| /aɔ/ | : | /áɔr/ | 'spread' |
| | | /máɔs/ | 'repeatedly' |
| | | /táɔ/ | 'lake' |
| /ai/ | : | /áit/ | 'pull toward oneself' |
| | | /ráit/ | 'structure' |
| | | /dáí/ | 'taste' |
| *(/ae/) | : | /daé/ | 'improper' |
| /æ/ | : | /aék/ | 'water' |
| | | /dáek/ | 'gather' |
| | | /sáε/ | 'finished' |

/i/ in position 1:

| | | |
|--------|----------|--------------|
| */iu/ | : /sʲuk/ | 'ladle, vb.' |
| | /pʲu/ | 'twist' |
| */io/ | : /tʲop/ | 'hold' |
| | /pʲo/ | 'invite' |
| */io/* | : /dʲor/ | 'search' |
| /ia/ | : /iás/ | 'clean' |
| | /bʲaŋ/ | 'dog' |
| | /pʲa/ | 'kidney' |

/e/ in position 1:

| | | |
|--------|----------|-------------|
| */eu/ | : /péut/ | 'fall' |
| | /mahéu/ | 'exhausted' |
| (/ei/) | : /éi/ | 'Hey!' |
| /ea/ | : /deák/ | 'a lot' |
| | /déa/ | 'lure' |

/ɛ/ in position 1:

| | | |
|-------|----------|-----------------------|
| */cɔ/ | : /péɔp/ | 'keep' |
| | /léɔ/ | 'kind of water plant' |
| /ɛa/ | : /éa/ | 'gait' |
| | /mɛat/ | 'perch' |
| | /gɛa/ | 'worm' |

/u/ in position 1:

| | | |
|---------|----------|----------------|
| *(/uo/) | : /túo/ | 'toss up coin' |
| */ui/* | : /túit/ | 'vanity' |
| (/ue/) | : /ué/ | 'Alack!' |
| | /bué/ | 'abundant' |
| /ua/ | : /úas/ | 'thirst' |
| | /túak/ | 'palm wine' |
| | /dúa/ | 'two' |

/o/ in position 1:

| | | |
|--------|----------|------------------------|
| /ou/ | : /óu/ | 'Yes, coming.' |
| | /kóum/ | 'kinsman' |
| | /jóu/ | 'call' |
| /oi/ | : /ói/ | 'Hey! or Ah!' |
| | /dóit/ | 'sting' |
| | /bói/ | 'may, or can' |
| */oa/ | : /bóan/ | 'take, bring' |
| | /dóa/ | 'invocation' |
| (/oe/) | : /oé/ | 'Ah, me!, in despair.' |

/ɔ/ in position 1:

| | | |
|---------|----------|-------------|
| *(/ɔi/* | : /sʲit/ | 'hip' |
| */ɔɛ/* | : /kɔɛŋ/ | 'very ugly' |
| */ɔa/ | : /gɔar/ | 'name' |
| | /rɔá/ | 'ugly, bad' |

To sum up the vowel inventory, we have: 3 front vowels /i, e, ε/; one low central vowel /a/; and 3 back vowels /u, o, ɔ/. Two are high, namely /i, u/; two are higher mid /e, o/; two are lower mid /ε, ɔ/; and one is low /a/.

Several things can be observed about the combinatorial structure of the vowels. In terms of the phonetic qualities of the vowels and the selection of the second components, we can see the following clustering patterns (see Diagram 5 below):

(1) The high front vowel /i/ takes only the central and back vowels in position 2, thus: /ia, iu, io, iɔ/.

(2) The higher mid front vowel /e/ takes only /a/ and /u/ in position 2, with a single occurrence of /i/ in the interjection /ei/, thus: /ea, eu, ei/.

(3) The lower mid front vowel /ε/ takes only /a/ and /ɔ/ in position 2, thus: /εa, εɔ/.

(4) The central vowel /a/ is the only vowel that can take any other vowel in position 2, thus we have /ai, ae, aε, aɔ, ao, au/.

(5) The lower mid back vowel /ɔ/ only takes /a/ and the front vowels /i, ε/, thus: /ɔa, ɔi, ɔε/.

DIAGRAM 5

DIAGRAM OF THE TWO-VOWEL CLUSTERING PATTERNS

Horizontal Index = Second Component.

Vertical Index = First Component.

Bold Line from top left corner to bottom right corner is the Line of Symmetry.

| | i | e | ε | a | ɔ | o | u |
|---|------|------|----|----|----|----|----|
| i | | ? | | ia | iɔ | io | iu |
| e | (ei) | | | ea | | ? | eu |
| ε | | | | εa | εɔ | | |
| a | ai | ae | aε | | aɔ | ao | au |
| ɔ | ɔi | | ɔε | ɔa | | | |
| o | oi | (oe) | | oa | | | ou |
| u | ui | ue | | ua | | uo | |

Note: The clusters in parentheses are found only in one interjection.

The question mark (?) indicates a missing counter-cluster.

(6) The higher mid back vowel /o/ takes /i, e, a, u/ in position 2, thus: /oi, oe, oa, ou/.

(7) The high back vowel /u/ takes /i, e, a, o/ in position 2, thus: /ui, ue, ua, uo/.

(8) Vowels on the same structural level can cluster, namely /iu, eo, but not /eo/.

(9) The only structurally vertical combinations that occur are /ei, ou/.

(10) The clustering habits of the vowels in the selection of the second component seem to be reciprocal, i.e. if X can take Y, then Y can also take X; with the exception of the clustering of the interjections /ei/ and /oe/. This fact can be seen in the symmetrical character of Diagram 5, p.52.

DERIVED TWO VOWEL CLUSTERS. The term derived is used here for the non-original clusters because they are 'derived' from affixation. In Batak as a member of the Indonesian branch of the polysynthetic Malayo-Polynesian family of languages, affixation is the major morphological device. The most used are the prefix and the suffix.

The addition of prefixes and suffixes to a word results in the formation of vowel clusters. Thus, a prefix ending in a vowel (e.g. /di-/) added to a word beginning with a vowel (e.g. /ádu/) produces a vowel cluster (i.e.: /ia/ in /diádu/); and a suffix beginning with a vowel (e.g. /-on/) added to a word ending in a vowel (e.g. /ádu/) produces a vowel cluster (i.e. /uo/ : /adúon/ 'to be overtaken'). The vowels that occur at the end of a prefix are /i, u, a/, as in /di-/, /hu-/, /ta-/, etc. There is no need to dwell on the prefix ending in /a/, as /a/ already clusters originally with all other vowels, and because we will confine ourselves only to the discussion of new cluster types formed, i.e. by a /Ci-/ and /Cu-/ prefix (C stands here for 'consonant'). It should be added here, that in the case when the end vowel of a prefix is identical to the initial vowel of the following word, no cluster results, as in such a sequence there is always a /+/ intervening.

The new cluster types derived from affixation, with a few examples, are listed below:

| | | |
|--------|-------------|---------------|
| /ie/ : | /diebáti/ | 'visited' |
| | /diebákkon/ | 'shown off' |
| /ie/ : | /diéton/ | 'calculated' |
| | /diéla/ | 'persuade' |
| /ue/ : | /huéton/ | 'I calculate' |
| | /huélek/ | 'I coax'. |

The new cluster types that are formed by the addition of the suffixes /-an/, /-ɔn/, and /-i/ are as follows:

| | | |
|------|---------------|--------------------|
| /eɔ/ | : /leléɔn/ | 'to be chased' |
| | /peléɔn/ | 'to be worshipped' |
| /oɔ/ | : /pàddobóɔn/ | 'fraud' |
| | /patibbóɔn/ | 'raise' |
| /uɔ/ | : /adúɔn/ | 'to be overtaken' |
| | /burúɔn/ | 'to be hunted'. |

THREE-VOWEL CLUSTERS. There are two kinds of three-vowel clusters, namely the clusters found in word bases and those resulting from affixation. As with the two vowel clusters we will call the former 'original clusters' and the latter 'derived clusters'.

The original three-vowel clusters are very few. We can distinguish two structural types, namely:

(1) The cluster that consists of a sequence of a back vowel, a central vowel, and a front vowel. The only example the writer has been able to find is:

/nuáɛɛ/ 'now'

(2) The cluster that consists of a sequence of a central vowel, a back vowel, and again the central vowel. Three examples of this have been found:

/baóa/ 'male person'

/saóan/ 'vial'

/saraóal/ 'trousers'

There are for some speakers more examples of this type, namely for those speakers who have /ɔ/ for the commoner /w/ in such words as:

/baóan/ for /báwan/ 'onion'

/maóas/ for /máwas/ 'ape'

The writer uses only the variant with /w/ for these words. But as mentioned earlier, for /saraóal/ he often has /saráwal/. This kind of free variation never occurs in his speech for the other two examples.

The subject of derived three-vowel clusters will be touched only lightly, as they arise under the same conditions as the derived two-vowel clusters, and on the consideration of space, because there are fifty types of those clusters.

There are here also two classes:

(1) The clusters formed by the addition of a prefix ending in a vowel before a word with an initial two-vowel cluster, for example:

/huaúhɔn/ 'I accept'

/diéáhi/ 'go after'

/dipaías/ 'cleaned'

(2) The clusters formed by the addition of a suffix beginning with a vowel to a word ending in a two-vowel cluster, for example:

/joúɔn/ 'to be called'

/duái/ 'to double'

/buéan/ 'in abundance'

The suffixes that can be used here are /-an/, /-ɔn/ and /-i/. In conclusion it should be mentioned that the derived classes constitute the majority of the vowel clusters.

NOTES

1. For ease of typing during the preparation of the dissertation manuscript, the vowel symbols /ɛ/ and /ɔ/ were replaced by /E/ and /O/ respectively. The latter symbols are retained in this edition. Therefore, in Chapters III, IV, and V and in Appendix I, only /E/ and /O/ will be used.

2. The dissertation (Chapters III, IV, and V) uses a simplified 3 stress notation, the primary stress symbol /'/ standing for both primary stress and secondary stress symbolised /' and ^/ in Chapters I and II. The conflated primary stress symbol /'/ is to be read primary ['] when it is the first /'/ in the macrosegment, otherwise it is to be read [^].

CHAPTER III

MORPHOPHONEMICS

3.1. WORDS

Beginning at this point we write a (non-phonemic) space between successive words.

The beginnings and ends of words show alternations of phonemic shape depending on environment. The basic shape of the end of a word, from which all alternations can be predicted, appears when the word is followed by nothing (in the same intonation segment); the basic shape of the beginning of the word appears when the word is initial.

In terms of their basic shapes, words consist of one or more syllables: the word may begin with a vowel, or with any consonant (except that /y/ and /w/ are rare, found only in loanwords), or with the cluster /dd/; it may end with a vowel or with /p t k s m n ŋ r l/. Word-medial consonantisms include (underlining indicates the involvement of a loanword):

TABLE 1

| | | | | | | | | | | | | | | | |
|----|----|----|----|-----------|----|-----------|-----------|-----------|----|----|----|----|----|----|-------|
| bb | | | | | | | | | | | | | | | |
| dd | | | | dj | | | | | | | | | | | |
| | | | | pp | | <u>pt</u> | | | | | | | | | |
| | | | | | | tt | | tc | | | | | | | |
| kb | kd | kg | kj | kp | kt | kk | <u>kc</u> | km | kn | kŋ | ks | | | kl | kr |
| md | | | | <u>mp</u> | | | | mm | | | | | | | |
| | | | | | | <u>nt</u> | | <u>nc</u> | | nn | | | | | |
| ŋb | ŋd | ŋg | ŋj | | | <u>ŋt</u> | ŋk | ŋm | ŋn | ŋŋ | ŋs | | | ŋl | ŋr |
| sb | sd | sg | | | sp | st | <u>sk</u> | sm | sn | sŋ | ss | | | sl | sr |
| lb | ld | lg | lj | lp | lt | | | lm | ln | lŋ | ls | lh | ll | | |
| rb | rd | rg | rj | rp | rt | rk | <u>rc</u> | rm | rn | rŋ | rs | rh | | | rr ry |

3.2. EXTERNAL SANDHI

The alternations of shape where words come together are subsumed by the following ten rules, which must be applied in the order given:

- ES1. h- → p- after -p, -m
 t- after -t
 k- after -k, -n, -ŋ
 s- after -s
- ES2. s- → c- after -t
- ES3. -n → -s before s-
 -r before r-
 -l before l-
- ES4. -m → -p }
 -n → -t } before p-, t-, c-, k-, s-
 -ŋ → -k }
- ES5. -t → -p before p-
 -k before k-
- ES6. -n → -m before b-, m-
 -ŋ before g-, ŋ-
- ES7. -m → -ŋ before g-
- ES8. -m → -b before b-
 -n → -d before d-, j-
- ES9. -k → -h before V- (any vowel)
- ES10. -p }
 -t } -k([q]) before b-, d-, j-, g-, m-, n-, ŋ-, r-, l-
 -k([k]) }

Because of the decision to treat syllable-final [q] as an alternant of the /k/ phoneme, one part of ES10. is a sub-phonemic alternation.

Table 2 summarises the results of these alternation rules. The row headings show the base form of the end of the first word, the column headings show the base form of the beginning of the second word, and the entries show what is actually pronounced (the word space is omitted for clarity). Cases actually affected by the rules (including the sub-phonemic effect of part of ES10.) are underlined.

TABLE 2

| | p | t | c | k | s | h | b | d | j | g | m | n | ŋ | r | l | v |
|---|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----|
| p | pp | pt | pc | pk | ps | pp | <u>kb</u> | <u>kd</u> | <u>kj</u> | <u>kg</u> | <u>km</u> | <u>kn</u> | <u>kŋ</u> | <u>kr</u> | <u>kl</u> | pV |
| t | <u>pp</u> | tt | tc | <u>kk</u> | <u>tc</u> | <u>tt</u> | <u>kb</u> | <u>kd</u> | <u>kj</u> | <u>kg</u> | <u>km</u> | <u>kn</u> | <u>kŋ</u> | <u>kr</u> | <u>kl</u> | tV |
| k | kp | kt | kc | kk | ks | <u>kk</u> | <u>kb</u> | <u>kd</u> | <u>kj</u> | <u>kg</u> | <u>km</u> | <u>kn</u> | <u>kŋ</u> | <u>ke</u> | <u>kl</u> | hV |
| s | sp | st | sc | sk | ss | <u>ss</u> | sb | sd | sj | sg | sm | sn | sŋ | sr | sl | sV |
| m | <u>pp</u> | <u>pt</u> | <u>pc</u> | <u>pk</u> | <u>ps</u> | <u>pp</u> | <u>bb</u> | md | mj | <u>ŋg</u> | mm | mn | mŋ | mr | ml | mV |
| n | <u>pp</u> | <u>tt</u> | <u>tc</u> | <u>kk</u> | <u>ss</u> | <u>kk</u> | <u>bb</u> | <u>dd</u> | <u>dj</u> | <u>ŋg</u> | <u>mm</u> | nn | <u>ŋŋ</u> | <u>rr</u> | <u>ll</u> | nV |
| ŋ | <u>kp</u> | <u>kt</u> | <u>kc</u> | <u>kk</u> | <u>ks</u> | <u>kk</u> | ŋb | ŋd | ŋj | ŋg | ŋm | ŋn | ŋŋ | ŋr | nI | nV |

(no alternations with other word-final phonemes)

3.2.1. Examples of External Sandhi

The list comprises only those that manifest alternations.

- p h- : /álap pàlah í/ 'fetch them', from /álap/ and /hàlak í/;
 m- : /ràk módom/ 'sleep together';
 n- : /álak na (di hùta í)/ 'get the one (in the village)';
 ŋ- : /rak ŋót/ 'get up together';
 r- : /rak ró/ 'come together';
 l- : /rak láo/ 'go together';
 b- : /álak bàtu í/ 'fetch the stone';
 d- : /álak dakdànah (i)/ 'fetch the children';
 j- : /manàruk jábu/ 'to roof a house';
 g- : /maŋàlak jádŋ/ 'to get yam'.
 -t p- : /ràup parnábuŋ/ 'a very sharp knife', from /ráut/ and /parnábuŋ/;
 k- : /huhúk kótór/ 'dirty at the same time';
 h- : /huhút tútur/ 'shake at the same time';
 s- : /huhút cÉat/ 'cut at the same time';
 m- : /huhúk mársak/ 'sad at the same time';
 n- : /huhúk nólnóI/ 'stare (at it) at the same time';
 ŋ- : /huhúk ŋót/ 'rise at the same time';
 r- : /júk rŋhána/ 'he is exasperated';
 l- : /huhúk láo/ 'go at the same time';
 b- : /huhúk bóla/ 'split at the same time';
 d- : /huhúk dók/ 'say at the same time';
 j- : /huhúk jóu/ 'call at the same time';
 g- : /huhúk gótap/ 'break at the same time'.
 -m p- : /ónOp pE/ 'just six';

- m t- : /h0hóp tabúat/ 'let's take (it) stealthily';
- c- : /h0hóp cÉt/ 'paint (it) silently';
- k- : /làp k0tÓr/ '(become) more dirty';
- s- : /holóp saÓtik/ 'somewhat dark';
- b- : /làb balgá/ 'become bigger';
- g- : /làg gogó/ 'become stronger'.
- n p- : /bùlap purnáma/ 'full moon';
- t- : /sìat tÓru/ 'from below';
- c- : /sàbut cúci/ 'laundry soap';
- k- : /sòηOk kúli/ 'like a cooly';
- h- : /sìak kúta/ 'from the village';
- s- : /sìas súηE/ 'from the river';
- m- : /sòηOm mánuk/ 'like a chicken';
- η- : /hòlaη ηalúkna/ 'only the misery';
- r- : /sìar rúra/ 'from the valley';
- l- : /sòηOl lóbuu/ 'like a cow';
- b- : /sòηOb bátu/ 'like a stone';
- d- : /sòηOd dólok/ 'like a mountain';
- g- : /sòηOη gÉa/ 'like a worm'.
- η p- : /dak pÓla/ 'not necessary';
- t- : /dak tabó/ 'not delicious';
- c- : /dak cúkkup/ 'not enough';
- k- : /dak k0tÓr/ 'not dirty';
- h- : /dak kEbbáη/ 'not spread';
- s- : /dak sáηa/ 'not enough time';
- k h- : /hàlak kúta/ 'villager';
- v- : /hàlah Ón/ 'this person';
- s h- : /là0s sudÓk (ma)/ 'and I said'.

3.3. INTERNAL MORPHOPHONEMICS

We provide for the (basic) shapes of words by setting up basic shapes for their constituent morphemes and specifying the alternations of shape that these morphemes undergo in the various environments, within the word, in which they can occur. Insofar as this does not involve either the mention of specific morphemes or the introduction of special symbols (that is, symbols not used in the phonemic transcription), the alternations are regular; other alternations are irregular.

3.3.1. Regular Internal Alternations

By and large the internal alternations follow the same rules as external sandhi. For internal combinations we will write a hyphen at

morpheme boundary, and a dash for uncompleted forms. The rules given below apply specifically to internal combinations not covered by the irregular alternations, and are to be applied in the order given; the output is then run through the ES rules, with the additional provision that at the end of the whole process the hyphens are deleted.

IS1. -ŋ → -n before consonants, except ŋ-, h-

IS2. -n → -t before s-, h-

IS3. -r → -l before l-

Examples:

/paŋ-dangur-h0n/ → /pan-dangur-h0n/ (IS1)

/pad-dangur-h0n/ (ES8)

/paddangúr-h0n/ 'throw (them)'

/di-dalan-h0n/ → /didalatt0n/ (IS2, ES1) '(he) carries out'

/mar-laŋE/ → /malláŋE/ (IS3) 'to swim'

3.3.2. Irregular Internal Alternations

The following rules (IIS) apply only to particular morphemes, which are here labelled by means of subscripts or capital letters. IIS rules are applied before IS and ES rules: the output is still to be run through the latter two sets of rules. /maŋ-/ is an active prefix in transitive verbs; /maŋ₁-/ is a derivational prefix in intransitive verbs; /UM/ is, amongst others, the comparative affix in adjectives; /NI/ is a completive-participial affix in transitive verbs; /-hu/, /-mu/, /-na/, /-ta/ are possessive pronominal suffixes, respectively, 'my', 'they', 'his, her, its', 'our (incl.)'; /-sa₁/ is an inclusive suffix after numerals ('all ...'); /-sa₁/ is an agentive or objective suffix.

IIS1. pV--
maŋ- { mV-- } → mamV--

 tV--
maŋ- { sV-- } → manV--

maŋ- { rV-- } → maŋa- { rV--
 lV-- }
 nV-- }

Exceptions to this rule are found with some stems beginning with /b/ yielding /mamV--/; namely, /búnu/ 'kill', /búri/ 'wash', /bÉrEŋ/ 'see', /bódil/ 'shoot', /búro/ 'scare away birds from a rice field', and a few other derived transitive verbs.

IIS2. $\text{ma}\eta_1\text{-hV--} \rightarrow \text{manV--}$
 $\text{-bV--} \rightarrow \text{mamV--}$
 $\text{-sV--} \rightarrow \text{man-sV--}$

otherwise, IIS1 applies.

Exceptions are:

$\text{/ma}\eta_1\text{-huli}\eta\text{/} \rightarrow \text{/makkuli}\eta\text{/}$ 'to speak'
 $\text{/ma}\eta_1\text{-siamun}/ \rightarrow \text{/ma}\eta\text{i}\text{amun}/$ 'to turn right'
 $\text{/ma}\eta_1\text{-singalak}/ \rightarrow \text{/maningalak}/$ 'to lie on one's back'

IIS3. UM- is infixed /<um>/ after the first consonant in morphemes beginning with /t c k s d j g r l h/.

UM- \rightarrow un- before n--, η --
 umm- before V--

There is variation between the two rules with /nEá η / 'light (not heavy)', yielding both /numEá η / and /unnEá η / 'lighter'.

IIS4. NI- is infixed /<in>/ after the first consonant in morphemes beginning with /p t k c b j g s h/

IIS5. --V \rightarrow --Vn before $^m\text{-ta/}$, $^m\text{-sa}_1\text{/}$

IIS6. -- η \rightarrow --k before $^m\text{-ta/}$, $^m\text{-sa}_2\text{/}$

IIS7. --V
 --Vn } \rightarrow --V η before $^m\text{-hu/}$

There are three morphemes ending in /n/ that show variation between regular (IS2) and irregular alternation (IIS7); namely, /sóban/ 'fire-wood', /úban/ 'grey hair', and /napúran/ 'betel'. One morpheme, /ípOn/ 'tooth', always follows the regular alternation, yielding /ipÓttu/ 'my tooth'.

IIS8. --na \rightarrow -ra after --r
 -lá after --l
 -ma after --m
 - η a after -- η

IIS9. $\text{-hu} + \begin{cases} i \rightarrow \text{-hi} \\ 0n \rightarrow \text{-hOn} \\ an \rightarrow \text{-han} \end{cases}$
 $\text{-mu} + \begin{cases} i \rightarrow \text{-mi} \\ 0n \rightarrow \text{-mOn} \\ an \rightarrow \text{-man} \end{cases}$

IIS10. -mu \rightarrow -m after --V

3.3.3. Stress Alternation

There are five special rules pertaining to stress that are to be applied after all the other rules.

SR1. In a word containing suffixes, except for those cases covered by SR3 and SR4 below, the heaviest stress falls on the penultimate; in IIS10 the deleted vowel is counted.

SR2. The heaviest stress in an uninflected adjective, which otherwise falls on the final syllable, shifts to the penultimate

- (a) after /na/ 'attributive particle', and /na sái/ 'so ...';
- (b) when affixed with /UM/ 'comparative affix';
- (c) before /hían/ 'extremely'.

SR3. The heaviest stress in the word (adjective) moves to the suffixes

- (a) /-hu₁/ (in combination with /pa-/ forms the excessive degree inflection, i.e. 'too ...');
- (b) /-an/ (another comparative morpheme).

SR4. In a vocative form, the heaviest stress is placed on the final syllable.

SR5. In an attributive phrase or a compound, the primary stress falls on the last determiner, including the enclitic demonstratives /i/, /0n/, /an/; the stressed syllable of the first word carries a secondary stress.

3.3.4. Examples

In the first few examples the application of the rules are shown step by step; the others give only the input and output with a specification of the rules applied; the slanting lines are omitted.

UM-naíí → un-ηaíí (IIS3)
 un-ηaíí (ES6)
 unηaíí (Hyphen Deletion)
 unηáli (SR2) 'colder'

pa-luón-hu₁ → pa-luón-tu (IS2)
 pa-luón-tu (ES1)
 paluón-ttu (H.D.)
 palu uttú (SR3) 'too lonely'

paη-súan-h0n → pan-súan-h0n (IS1)
 pat-súat-h0n (IS2, twice)
 pat-súat-t0n (ES1)
 pat-cúat-t0n (ES2)
 patcúatt0n (H.D.)
 patcuátt0n (SR1) 'plant (them)'

CHAPTER IV

MORPHOLOGY

4.0. INTRODUCTION

Henceforth, in citing words or morphemes we will give them in their base forms, (§3.1.), in bold type; when useful for clarity, the actual phonemic shape of the word or phrase will be added between slant lines.

4.1. PARTS OF SPEECH

The parts of speech are:

CONTENTIVES:

Nouns: *jábu* 'house'; *sukkutcúkkun* 'question'.

Verbs: *módom* 'sleep'; *pásak* 'beat'.

Adjectives: *deḡán* 'good'; *daó* 'far'.

Adjuncts: *annón* 'presently'; *t0kt0ḡ* 'always'.

Interjections: *dagó* 'oops'; *na* 'here you are'.

FUNCTORS:

Substitutes:

Pronoun: *áu* 'I'; *On* 'this'.

Proverb: *marhúa* 'do what'.

Pro-adjective: *bEhá* 'of what sort'.

Pro-adjunct: *disón* 'here'; *addígan* 'when'.

Pro-clause: *dá0ḡ* 'no!'; *ípE* 'therefore'.

Directive Particles: *di* 'at'; *mó10* 'if'.

Modal Particles: *d0*; *ma*.

Connectives: *alái* 'but'; *mánaḡ* 'or'.

Auxiliaries: *s0r* 'almost'; *hólan* 'only'.

Some of these fall into sub-classes in terms of differences of inflection or syntax; see below.

Nouns, verbs, and adjectives are distinguished by different types of inflection and syntax, the others by syntax.

4.2. INFLECTIONAL CATEGORIES

Inflection is made for six types of categories:

Substantive Class: person, number;

Substantive Reference: allocation, actor, object;

Voice and Resolution: active, passive, reciprocal; intransitive, transitive;

Mode and Aspect: imperative, simple, completive (-participial), distributive, promissory, potential, simple, iterative, explicit plural, intensive-instrumental;

Degree (in adjectives): comparative, excessive;

Ordinal (in numerals): simple, ranking.

One and the same inflectional affix may reflect more than one category.

4.2.1. Substantive Class

There are two categories of substantive class:

Person: first (speaker), second (hearer), third, and indefinite;

Number: singular and plural.

The categories of substantive class appear in personal pronouns, and through pronominal affixes in the inflection of nouns (for allocation) and of verbs (for actor and object): *pat 'leg' : páttu 'my leg';* *buát0n 'will be taken' : buát0nnámi 'will be taken by us'.*

Some transitive verbs are inflected with *paŋ-* for explicit plural object: *daŋgúrh0n 'throw' : dipaddaŋgúrh0n ibána (ma bàtu í) '(the stones) are thrown by him';* *léan 'give' : palléan òn (tu nasída) 'give these (to them)'.*

Some intransitive verbs are inflected with *mar-...-an* for explicit plural subject: *húddul 'sit; sit down' : marhuddúlan (ma nasída) '(they) are sitting down';* *pÉák 'lie' : marpÉáhan (d0 búkku í) '(the books) are lying about'.*

4.2.2. Substantive Reference

The categories of substantive reference are: possessor of a noun (allocation), actor of a passive transitive verb, and object of an active transitive verb. The reference is personal-anaphoric; in the case of allocation or of the actor of a promissory passive verb it specifies the person and number of the possessor or actor; in the other cases of actor reference and in the case of object reference, the specification is defective as to person and number.

The actor reference in simple passive verbs indicates:

first person exclusive, first person inclusive, and

non-first-person actor: *buat 'take' : hubuat 'taken by me, or*

by persons including me, excluding you', *tabúat* 'taken by us (inclusive)', *dibúat* 'taken by persons excluding me'.

The actor reference in potential passive verbs (and in a small set of verbs not inflected for mode) distinguishes only between indefinite and third person singular, -sa: *tarhiddat* 'can be lifted' : *tarhiddátca* 'can be lifted by him (her, it)'; *dáp0t* '(be) found' : *dápótca* 'found by him'.

The object reference in simple and completive active verbs also makes only this two-way distinction: *makkíddat* 'to lift' : *makkiddátca* 'lift him (her, it)'; *uppóppar* 'have as descendant (said of males)' : *uppóppársa* 'have him as descendant'.

4.2.2.1. Allocation

All nouns occur in unpossessed form, and most nouns occur in possessed forms. The possessed forms show the person and number of the possessor by means of pronominal suffixes: *jábu* 'house' : *jabúna* 'his (her) house', *jabumúna* 'your (pl.) house', *jabúтта* 'our (incl.) house'.

With some nouns of relationship, a prefix inflection, *da-*, is used for first person singular possessor: *ínaṇ* 'mother' : *daínaṇ* 'my mother'.

4.2.2.2. Actor

Passive verbs contain inflectional reference to an actor.

Simple passive verbs have prefixes referring to first person exclusive actor, *hu-*; first person inclusive, *ta-*; non-first person, *di-*. The actor, if expressed by a separate word, comes immediately after the verb: *búat* 'take' : *dibúat* 'taken (by non-first person)', *dibùat* *hàlak í* 'taken by the man'. When the actor is first person singular, the actor morpheme is zero: *hubúat* 'taken by me'; when the actor is first person inclusive, the actor morpheme is zero: *tabúat* 'taken by us (incl.)'; when the actor is third person singular, the actor morpheme can be represented by zero: *dibúat* or *dibùat* *ibána* 'taken by him (her)'.

Promissory passive verbs are inflected for pronoun actor: *buát0n* 'will be taken; to be taken' : *buatókku* 'will be taken by me', *buat0nnámi* 'will be taken by us (excl.)', *buatótta* 'will be taken by us (incl.)'.

Completive passive verbs, in nominal clauses, are inflected for pronoun actor: *binúat* 'has been taken' : *binuaknámi* 'has been taken by us (excl.)' in *na binuaknámi* 'what has been taken by us (excl.)'.

4.2.2.3. Object

In non-initial position in the sentence, active transitive verbs are inflected, with *-sa*, for third person singular object. This inflection is optional in the simple mode, obligatory in the completive mode:
mabbége 'hear' : *mabbegésa* 'hear it (him, her)'; *ubbége* 'have heard' :
ubbegésa 'have heard it (him, her)'; *tádda* 'know (= recognise)' :
manaddása or *manàdda* *ibána* 'know him (her)', *tumaddása* 'have known
 (= know) him (her, it)'.

4.2.3. Voice and Resolution

Verbs are intransitive or transitive: *mardálan* (d0 *ibána*) 'he walks'; *manùh0r búkku* (d0 *ibána*) '(he) is buying a book'; *dibùat ibána* (d0 *búkku í*) '(the book) is taken by him'.

4.2.3.1. Voice

Transitive verbs are inflected for active (*maŋ-*, etc.), passive (*di-*, etc.), and reciprocal (*masi-...-an*):

jou 'call' : *madjóu* 'to call (somebody)';
dijóu 'be called (by somebody)';
masijoúan 'call each other'.

When the stem ends in the suffix *-i*, the suffix is dropped for the reciprocal inflection; e.g. *pudí* 'turn one's back to' (from *púdi* 'back part' + *-i*) : *masipudían* d0 *nasída* 'they are standing with their backs to each other'; *r0ái* 'defame' (from *r0á* 'ugly' + *-i*) : *masir0áan* 'defame each other'.

4.2.4. Mode and Aspect

Transitive verbs are inflected for mode and aspect.

4.2.4.1. Mode: Active Transitive

Active transitive verbs appear in four modes: imperative (with no inflectional affixes), simple (*maŋ-*), completive (*-participial*) (*UM-*), and distributive (*masi-*). The third is rarer than the others.

jóu 'call' : *jóu* (*ibána*) 'call (him)!';
madjóu (d0 *ŋan*) 'call (a friend)';
jumòu (d0 *ŋan*) 'have called (a friend); calling
 (a friend)';
masijòu (d0 *ŋánna*) 'each calls (his own friend)'.

4.2.4.2. Passive Transitive

Passive transitive verbs are inflected for a different set of modes: simple, completive (-participial), promissory and potential.

The simple mode shows inflection for person of the actor: hu- 'first person exclusive'; ta- 'first person inclusive'; di- 'non-first person'. The actor is placed immediately after the verb; when the actor is first person singular or first person inclusive, it is represented by zero; when the actor is third person singular, it is optionally represented by zero:

búat 'take' : hubúat 'taken by me';
 hubúat hámi 'taken by us (excl.)';
 tabúat 'taken by us (incl.)';
 dibúat hÓ 'taken by thee';
 dibúat hamú 'taken by you';
 dibúat or dibúat ibána 'taken by him (her)';
 dibúat nasída 'taken by them';
 dibúat hàlak í 'taken by the man'.

The completive mode shows a different inflection for person of the actor: hu- 'first person singular' (cf. simple mode above); ta- 'first person inclusive'; NI- 'other persons (including indefinite person)'. The verb, further, has an inflectional ending for pronoun actor; in the case of hu- 'first person singular' and ta- 'first person inclusive', the actor is represented by zero; noun actors are marked by ni:

búat 'take' : (na) binúat '(what) has been taken';
 (na) hubúat '(what) has been taken by me';
 (na) binuaknámi '(what) has been taken by us (excl.)';
 (na) tabúat '(what) has been taken by us (incl.)';
 (na) binuákmu '(what) has been taken by thee';
 (na) binuákna '(what) has been taken by him (her, it)';
 (na) binuaknasída '(what) has been taken by them';
 (na) binuakníba '(what) has been taken by one';
 (na) binúat ni hàlak í '(what) has been taken by the man'.

When the stem ends in the suffix -i, this is replaced by -an for the NI- inflection of the completive-participial mode; e.g. ulósi 'cover (with a blanket)' (from úlós 'blanket') + -i : niulósan 'have been covered'.

The most frequent uses of the completive are: when the actor is indefinite; in absolute passive forms (i.e. when no actor is expressed); and as the verbal element in a nominal clause (introduced by na as above).

The promissory has the mode marker -On (-an, when the underlying forms ends in the suffix -i), and takes the inflectional ending to show

the actor, a noun actor is marked by ni:

búat 'take' : buát0n 'will be taken';
 buatÓkku 'will be taken by me';
 buat0nnámi 'will be taken by us (excl.)';
 buatÓtta 'will be taken by us (incl.)';
 buat0mmu 'will be taken by thee';
 buat0mmúna 'will be taken by you';
 buatÓnna 'will be taken by him (her, it)';
 buat0nnasída 'will be taken by them';
 buat0nníba 'will be taken by one';
 buàt0n ni hàlak í 'will be taken by the man'.

ulÓsi 'cover with a blanket' '' ulÓsan 'will be covered with a blanket'; ul0sákku 'will be covered with a blanket by me'; etc.

The promissory expresses a promise, a duty, or future execution of something: buatÓkku d0 í 'it will be taken by me' = 'I will take it'; (ìkk0n) olóan d0 nat0rasníba 'one's parents must be obeyed' = 'one should obey one's parents'.

The promissory is also used absolutely: jóú0n d0 ibána 'he will be called'. When the underlying form ends in -h0n, the inflectional ending in absolute use has the form -0tt0n (instead of -0n); guruh0n 'learn' : guruh0nÓtt0n (d0 nàsa na turÉ) '(all what is proper) is to be learned'.

A potential passive verb is inflected for mode with tar- (ha-...-an, when the underlying form ends in the suffix -i); it further, optionally, has an inflectional ending for third person singular actor, -sa; the actor is placed immediately after the verb:

búat 'take' : tarbúat (áu) 'can be taken (by me)';
 tarbúat hó 'can be taken by thee';
 tarbuátca 'can be taken by him (her, it)';
 tarbúat ibána 'can be taken by him (her)';
 tarbúat hàlak í 'can be taken by the man'.

dakdátí 'continue' : hadakdàtan (áu) 'can be continued (by me)',
 hadakdàtan hó 'can be continued by thee', etc.

4.2.4.3. Aspect

Transitive verbs are inflected for three aspects: iterative (-i), explicit plural object (paŋ-), and intensive-instrumental (-h0n); a transitive verb without any of these inflections has simple aspect:

pásak 'beat' : pasáhi 'beat repeatedly';
 búat 'take' : pabbúat 'take (plural object)';
 pásak 'beat' : pásákk0n 'beat with (something); DO beat (something)'.

The explicit plural object aspect can occur with the intensive-instrumental:

pasak 'beat' : pappasákk0n (nasída) 'DO beat (them)';
 pappasákk0n (balòbas í) 'beat with (the rulers)'
 or 'beat with the ruler many times'.

From the last example we see that it is sometimes not easy to tell whether we have to do with something done many times or with many things used individually.

Modal inflection is superimposed on aspectual inflection. The modal inflections of the iterative aspect are given below in the second column, the first column is the simple aspect:

| | | | |
|----------|--------------|-------------|--|
| Active: | imperative | jóu 'call!' | : joúi 'call repeatedly!'; |
| | simple | madjóu | : madjoúi '(to) call repeatedly'; |
| | completive | jumóu | : jumoúi 'have called repeatedly'; |
| | distributive | masijóu | : masijoúi 'each calls repeatedly'; |
| Passive: | simple | e.g. hujóu | : hujoúi 'called repeatedly by me'; |
| | completive | e.g. jinóu | : jinoúan 'have been called repeatedly'; |
| | promissory | e.g. jóú0n | : joúan 'will be called repeatedly'; |
| | potential | tarjóu | : hajoúan 'can be called repeatedly'. |

The modal inflections are different in the last three modes (see §4.2.4.2. above).

An example of the modal inflection of the intensive-instrumental aspect: jóu 'call' : jinóu (d0 ibána) '(he) has been called', jinoúh0n (d0 ibána) '(he) HAS been called'. The inflection for the promissory mode in absolute use (-0n) appears in a different form (-0tt0n) with the intensive-instrumental aspect: jóú0n d0 ibána 'he will be called' : jouh0n0tt0n d0 ibána 'he WILL be called' (see also §4.2.4.2. above).

The explicit plural object aspect is inflected for mode in a regular fashion; e.g.

| | | |
|--------------------|-------------|--------------------------------------|
| imperative active | lÉan 'give' | : pallÉan 'give (many things)'; |
| simple passive | hulÉan | : hupallÉan '(pl.) are given by me'; |
| potential passive | tallÉan | : tarpallÉan '(pl.) can be given'; |
| completive passive | nilÉan | : pinallÉan 'have been given'. |

4.2.5. Degree

The category of degree appears in adjectives; it falls into two classes: comparative (UM- or -an), and

excessive (pa-...-hu):

balgá 'big' : ubbálga or balgaán 'bigger', pabalgahú 'too big';
 r0á 'ugly' : rum0a or r0aán 'uglier', par0ahú 'too ugly';

daó 'far' : dumáo or daoán 'farther', padaohú 'too far';
 padditá 'preacherlike' : uppaddíta or padditaán 'more preacherlike',
 papadditahú 'too preacherlike'.

With ordinary adjectives like the first three examples the UM- comparative is more common; with forms which are only occasionally used as adjectives like the last example, the -an comparative is preferred. In colloquial speech, we sometimes find a combination of the two comparative inflections; uppadditaán 'more preacherlike'.

4.2.6. Ordinal

This category appears in numerals (a sub-class of pronouns). There are two ordinal categories: simple (pa-...-h0n) and ranking (pai-).

Simple ordinal indicates a place in a series; ranking ordinal is used when there is some judgment of excellence or prestige involved in the set (e.g. in a race).

dúa 'two' : paduáh0n 'second', paidúa 'the second; number two';
 tólu 'three' : patolúh0n 'third', paitólu 'the third; number three'.

Irregularly, we have the simple ordinal parjólo 'first' (from par- + jólo 'front').

The ordinal affixes are I.C. partners to complex numerals; the immediate constituents of paòpat pùlu sadáh0n 'forty-first' are pa-...-h0n and òpat pùlu sáda 'forty one'; paisappùlu pítu 'the seventeenth' consists of pai- and sappùlu pítu 'seventeen'.

Simple ordinal inflection is also made with expressions of measure: patòlu makkúkk0n 'the third cupful', paòpat halíh0n 'the fourth time'.

For birth rank among siblings, it is more common to use special forms (nouns): siaháan 'eldest' (si- + háha 'older sibling same sex' + -an), siappúdan 'youngest' (presumably, si- + UM- + púdi 'back' + -an), and, if the number is odd, silitóŋa 'the middle' (si- + di- + tóŋa 'middle').

4.3. CONTENTIVES

4.3.1. Nouns

Nouns fall into several classes. Nouns which can take the plural auxiliary ákka will be called common nouns; others are personal proper nouns, non-personal proper nouns, measure nouns, and mass nouns, which besides not taking the plural auxiliary have the following characteristics.

Personal proper nouns are marked by a preposed particle si or are formed with the (teknonymous) formatives náí 'mother of', ámani 'father

of', óppuni 'grandparent of'; e.g. si sáhat; nài badía 'Mother of Badia'.

In direct address, the particle si is not used; thus sahát 'Sahat!'. si is not used with family names (a sub-class of personal proper nouns) beginning with si-; e.g. sihóbbi. Since the si- of family names is not an auxiliary, it is retained in the vocative.

Non-personal proper nouns are not inflected for allocation; e.g. bútar, balíge 'place names'.

Measure nouns take a preceding attributive numeral. The numeral sáda 'one' appears bound to a free measure noun in the form sa-: dúa híraŋ 'two basketfuls', sahíraŋ 'one basketfull'. Irregularly, we have sadári 'one day' (ári 'day'). Bound measure nouns are linked by an accretive /ŋ/ to a numeral ending in a vowel: sa- → saŋ-, dúa 'two' → duaŋ-; e.g. -harippan 'piece (of warty root, like ginger)' : sakkaríppan 'one piece', dùakkaríppan 'two pieces'.

A mass noun does not take a bare numeral; it is quantified only by the combination numeral + measure noun: samákkuk kópi 'a cup of coffee', dùa tákkar dahánOn 'two cans (measure) of rice'.

Some nouns of relationship have different forms, obligatory, or optional, in the vocative, in the following list, the second is the vocative:

| | |
|---------------------------------|--------------------------------------|
| ína 'mother' | : inán 'mother!' (obligatory); |
| áma 'father' | : amán 'father!' (obligatory); |
| óppu 'grandparent' | : oppún 'grandparent!' (obligatory); |
| ángi 'younger sibling same age' | : angiá (optional); |
| háha 'older sibling same sex' | : hahán or akkán (obligatory); |
| ibóto 'sibling other sex' | : itó (obligatory). |

In polite speech, the vocative forms (except for angiá) are used for non-vocative functions, presumably because of the unfavourable connotation of the common forms through their use in cursing: e.g. inám 'they mother!' (a very rude and insulting exclamation). With a noun from the set above, the possessive prefix da- 'my' is attached only to the vocative form.

4.3.2. Verbs

To summarize the discussions on the categories of verbs we present a morpheme order table of the inflection of transitive verbs on the following page. In the table, positions -2 and +2 are mutually exclusive or both are occupied by the parts of a discontinuous morpheme; paŋ- (-1) is mutually exclusive with -i (+1); aspect suffixes bear subscript numerals: the mode inflection with a corresponding subscript numeral (if any) is to be chosen in the case of an aspectual underlying form; discontinuous morphemes are indicated by parentheses and dots.

TABLE 3

| | -2 Mode/ Actor ref. | -1 Aspect | stem | +1 Aspect | +2 Mode | +3 Actor/Object |
|----------------------------|---------------------------|--------------|------|---------------------|------------|---|
| <hr/> | | | | | | |
| ASPECTS: | | | | | | |
| Iterative | | | | -i ₁ | | |
| Intensive- instrumental | | | | -h0n ₂ | | |
| Explicit plural | | | paŋ- | | | |
| ACTIVE MODES: | | | | | | |
| Simple | maŋ- | | | | | -sa (Obj.) |
| Completive- participial | UM- | | | | | -sa (Obj.) |
| Distributive | masi- | | | | | |
| PASSIVE MODES: | | | | | | |
| Simple | hu- | | | | | |
| | ta- | | | | | |
| | di- | | | | | |
| Completive- participial | hu- | | | | | -nami, etc. (actor, except 1st sg. & 1st incl.) |
| | ta- | | | | | |
| | NI- | | | | | |
| | (NI | | | -an) ₁ | | |
| Promissory | | | | -0n | | -hu, -nami, etc. (Act.) |
| | | | | -an ₁ | | |
| | | | | -0tt0n ₂ | | (absolute use) |
| Potential | tar- | | | | | |
| | (ha | | | -an) ₁ | | |
| RECIPROCAL VOICE: | (masi- | | | -an) | | |

Examples: pásak 'beat'; búat 'take'

| | A S P E C T S | | | |
|----------------------------|---------------|------------|------------------|-----------------|
| | Simple | Iterative | Intensive-instr. | Explicit plural |
| <hr/> | | | | |
| ACTIVE: | | | | |
| Imperative | pásak | pasáhi | pasákk0n | pabbúat |
| Simple | mamásak | mamasáhi | mamasákk0n | --- |
| Completive- participial | uppásak | uppasáhi | uppasákk0n | --- |
| Distributive | masipásak | masipasáhi | masipasákk0n | masipabbúat |
| PASSIVE: | | | | |
| Simple | dipásak | dipasáhi | dipasákk0n | dipabbúat |
| Completive- participial | pinásak | pinasáhan | pinasákk0n | pinabbúat |
| Promissory | pasáh0n | pasáhan | pasakk0n0n | pabbuát0n |
| Potential | tarpásak | hapasáhan | tarpasákk0n | tarpabbúat |
| RECIPROCAL: | masipasáhan | --- | --- | --- |

A small set of verbs are inflected for actor in the same way as the potential passive; these verbs have a passive denotation:

súda 'finished; used up' : sudása 'finished by him';
 júppaŋ 'found (after a search)' : juppáksa 'found by him';
 dáp0t 'found; got' : dapótca 'found by him; obtained by him';
 tólap 'capable of doing' : t0lápsa 'can be done by him'.

One verb forms the promissory passive in an irregular way: d0k 'say; tell' : d0h0n0n 'will be said'.

One transitive verb has a defective inflection as to mode and is peculiar in other ways. The stem is niŋ- 'say (quotative)'. The occurring forms are: níkku 'I say (said)', nítta 'we (incl.) said'; nínna or nínna ibána 'he said', nínna hámi 'we (excl.) said', nímmu or nínna hó 'thou said', nínna hamú 'you said', nínna nasída 'they said', nínna hálak í 'the person said'. The phrase nínna hamú is used as a token of high formality: nùŋga ró ibána 'he has come' (an ordinary statement), nùŋga ró ibána nínna hamú 'he has come' (said to a person with whom one is on a very formal relationship).

Intransitive and transitive verbs do not only differ in syntax and inflection, but also in derivational patterns (see §4.6.4. below).

4.3.3. Adjectives

Adjectives are simple stems. Syntactically they behave like intransitive verbs. They differ in that they are inflected for degree (see §4.2.5. above).

Predicative adjectives not inflected for degree are stressed on the last syllable; e.g. tibbó 'high' in tibbo (d0 jabu 0n) '(this house is) tall'; compare with tumíbb0 d0 jàbu 0n sian án 'this house is taller than that', and jàbu na tíbb0 d0 0n 'this is a tall house'. An attributive adjective comes after the noun head and is marked by the preposition na (see also the previous example): jàbu na bùruk na tibbo í 'the tall old house'. In hálak na hatóp mardálan 'a person who walks fast', hatóp 'quick; fast' is predicative in the attributive clause (note the stress pattern).

4.3.4. Adjuncts

An adjunct is a stem which is not a noun, a verb, or an adjective, and which occurs in the frames lá0 d0 ibána ---, --- d0 lá0 ibána, or --- lá0 d0 ibána 'he goes ...'

There are two types of adjuncts: verb adjuncts, which are attributive to verbs, and clause adjuncts, which are attributive to clauses (examples below). Verb adjuncts are different from predicator

auxiliaries (§4.4.5.) in that the latter are always next to the verb head, whereas the former usually occupy a different position. Verb adjuncts fall into two classes, which we will call *time adjuncts* and *adverbs*. Time adjuncts have a freer distribution than adverbs: of the three frames above, the last cannot be occupied by an adverb. Time adjuncts are morphologically complex and are morphophonemically irregular; e.g. *nakkiníŋ* 'just now' (from *na + i + i*), *bodári* '(in the) evening' (from *bót + di + ári*), *annón* 'presently' (from *an + 0n*). Adverbs are simple stems. They share some syntactical privileges of occurrence with adjectives, but are not inflected for degree.

Examples:

Time Adjuncts: *annón* 'presently' in *annón d0 ibána múlak* 'he will come back presently'; *namasahái* 'the other day' in *pajúppaŋ d0 hámi namasahái* 'we met the other day'; *nasogót* 'this (past) morning' in *nasogót d0 ibána lá0* 'he went this morning'.

Adverbs: *t0kt0ŋ* 'always' in *t0kt0ŋ d0 ibána marpógos* 'he is always (= remains) poor'; *sáut* 'go through (of a plan)' in *sáut d0 ibána ró* 'he came as planned'; *dóh0t* 'do something together with another' in *dóh0t d0 ibána marÉddE* 'he is singing too (along with others)'.

Clause Adjuncts: *ra* 'probably' in *rá ró dó ibána* 'he will probably come'; *ulániŋ* 'I wonder' in *ró d0 ibána ulániŋ* 'I wonder whether he will come'; *t0ttú* 'certainly' in *t0ttú ró d0 ibána* 'it is certain that he will come'.

4.3.5. Interjections

Interjections are exclamatory words (words used in cursing, animal-call words, words for the attraction of attention, and the like) and other words which regularly occur as one-word clauses (minor clause-types).

Examples:

Exclamatory words: *buriapús* or *apús* 'I'll be damned'; *bayá* or *puán* 'buddy'; *húrree* 'chicken call'; *húdjee* 'pig call'; *seh* 'to drive away dogs'; *peh* 'exclamation of disgust'. Many of these words (e.g. the last four examples) manifest deviant phonological behaviour.

Others: *béta* 'come (i.e. let's go)'; *b0háí* 'indeed?'; *t0lE* 'come on!'; *hóras* 'greeting, usable at any time', *síu* 'please' (used at the end of a sentence, only to children).

4.4. FUNCTORS

4.4.1. Substitutes

The substitutes fall into several classes according to their domain-tie: pronoun, proverb, pro-adjective, pro-adjunct and pro-clause.

4.4.1.1. *Pronouns*

These fall into several types: personal, demonstrative, interrogative, inclusive, partitive, and indefinite pronouns, and numerals.

The personal pronouns are:

| | Singular | | Plural | |
|---------------------|-----------|------------|-----------|------------|
| | Free form | Bound form | Free form | Bound form |
| First person: excl. | áu | hu-, -hu | hámi | -nami |
| incl. | | | híta | ta-, -ta |
| Second person | h0 | -mu | hamú | -muna |
| Third person | ibána | -na | nasída | -nasida |
| Indefinite person | íba | -niba | | |

The bound forms are used in inflectional reference to possessor, actor, and object (cf. §4.2.1.2.).

The plural of the second person (hamú) and of the third person (nasída) are also used to refer to one person to show respect to the person referred to; the ambiguity is usually resolved by the context. This is a matter of stylistics; likewise, the indefinite person (íba) is used in some situations to refer to the speaker; e.g. when a child makes an objection to his parents: *nùŋŋa lója íba 'one is already tired'*, that is *'I am already tired (so don't make me do any more work)'*. With nouns of relation (and others like jábu 'home; house'), the first person inclusive suffix (-ta) refers to the speaker or the hearer depending on the context.

The demonstrative pronouns fall into two sets.

The first set consists of 0n 'this', i 'that', an 'that (yonder)', adu or adui 'that (over there)'; 0n is closer to speaker than to hearer, i is closer to hearer than to speaker, an is far from both, adú(i) is very far from both. Examples: *mút0r d0 0n 'this is a car'*; *ummárga d0 í sian án 'that is more expensive than that (yonder)'*. These demonstratives are also used attributively: *jàbu 0n 'this house'*, *dòlok án 'that mountain (yonder)'*, *dòlok adú 'that mountain (over there)'*.

i is also used without any definite demonstrative force (much like 'it' and the definite article in English): *nùŋŋa ditùh0r ibána bùkku í 'he already bought the book'*; *dàŋ hubót0 ì 'I don't know it'*.

0n, i, an are enclitically linked to a noun inflected for first and second person singular possessor: *jabúkku 'my house' + i → jàbukkí 'my house there'*; *d0ŋámmu 'thy friend' + 0n → d0ŋammón 'thy friend here'* (cf. §3.3.2.).

The second set of demonstratives consists of ddión 'this one', ddi 'that one', ddián 'that one yonder', ddadú or ddadúi 'that one over

there'. These are more emphatic than the first set. Morphologically, they seem to have developed from *na + di + On (i, an, adú, adúi) '*that which is here (there, etc.)*'.

The interrogative pronouns are ísE '*who*', áha '*what*', díá '*which*'; sadíá '*how much (price)*': ísE manùhOr ón '*who bought this?*'; ísE gOármu '*who (= what) is your name?*'; áha dO ón '*what is this?*'; áha ì '*what is it?*'; díá duméngan '*which is better?*'; jàbu díá na tumíbbo '*which house is taller (or tallest)?*'; sadíá òn '*how much (price) is this*'.

The inclusive pronoun is sudé or sasudé '*all*': sudé dO ró '*everybody came*'; túhOr ma sudé '*buy all*'; sudé jólma ìkkOn mátE dO '*all men must die*'.

The partitive pronoun is déba '*some*': déba dO ró '*some came*'; túhOr ma déba '*buy some (part) of it*'.

The indefinite pronouns are saOtik or Otik '*a little; some; a few*'; (si) anu '*so-and-so*' (aha is also used for this purpose); aha '*such-and-such a thing*'; godán '*many (people or things)*'. Examples: túhOr ma saótik '*buy a few (a little)*'; léan òn tu si ánu '*give this to so-and-so*'; túhOr ma àha í '*buy the what-do-you-call-it*'; túhOr ma godán '*buy many (much)*'. The indefinite personal pronoun íba '*one*' is mentioned above: ìkkOn garárOn dO utaŋŋíba '*one's debts must be paid*', that is '*one must pay one's debts*'.

Numerals. There are two classes of numerals, cardinal and ordinal. The simple cardinal numerals are: sáda '*one*', dúá '*two*', tólu '*three*', ópat '*four*', líma '*five*', ónOm '*six*', pítu '*seven*', uálu '*eight*', síá '*nine*', púlu '*ten*', rátus '*hundred*', ríbu '*thousand*', lóksa '*ten thousand*' (this is going out of use), júta '*million*'. Higher numbers are formed by the pattern number of unit + unit + number of next lower unit + (next lower) unit + ... + digit (if any); when the number of unit is one, the shortened form sa- of sada (saŋ- before púlu '*ten*') is prefixed to the unit; otherwise, the full forms are preposed to the unit; thus, sappúlu '*ten*', dùa púlu '*twenty*', sappúlu pítu '*seventeen*', dùa púlu pítu ríbu tòlu ràtus uálu púlu '*twenty seven thousand three hundred and eighty*'.

The lower cardinal numerals beginning with dúá '*two*' (rarely higher than eight) take an inclusive suffix -sa, in which case the numerals ending in a vowel have the accretive ending /ŋ/ (see also §3.3.2.). Thus, duátca '*both*', Opátca '*all four*', Onópsa '*all six*'. With higher numbers, inclusiveness is expressed by sudé '*all*' + na '*preposition*' + number + demonstrative pronoun: sudé na òpat púlu í '*all forty*', sudé na sappúlu ònOm ón '*all these sixteen (i.e. all sixteen here)*'. Perhaps, we also have the inclusive suffix -sa in sasadása '*(all) by himself*'; the morphemic structure seems to be sa- '*one*' (bound form) + sáda '*one*'

+ -sa (inclusive suffix).

A sub-class of the cardinal numerals is the interrogative píga 'how many': píga bukkúna 'how many are his books (how many books does he have?)'; pìga háli 'how many times'. It is regularly inflected for ordinal: papigáhOn 'the howmanieth (= what number)', paipíga 'what number?'.
 For ordinals, see §3.5.

4.4.1.2. Proverbs

The proverbs are formed from the root √-hua: active intransitive marhúa 'do(ing) what', stative intransitive mahúa 'what happened to it', and transitive hahúa 'do something to (something)'. Examples: marhúa ibána disí 'what is he doing there?'; dah marhúa híta tusí 'we are not doing anything there (i.e. we have no business going there)'; mahua ibana 'what happened to him?'; dàṇ mahúa ibána 'nothing happened to him'; ísE makkahúa hó 'who has done what to you?' (active voice, simple mode); in the following examples of the passive, the translations are by necessity in the active: dīhahúa ibána hó 'what did he do to you?'; dàṇ hahuaónna hámi 'he won't do anything to us (excl.)'; dàṇ hinahúa ibána 'nothing has been done to him'; dàṇ tarhahuása ì 'he cannot do anything to (or, about) it'. The above examples show that the transitive stem hahúa is regularly inflected for all the modes in the passive and for the simple mode in the active. The reciprocal masihahuáan 'do what to each other' does occur in the interrogative, but it is rare; so is the distributive active masihahúa 'each does what'; the completive active *humahua does not seem to occur.

4.4.1.3. Pro-adjective

The pro-adjective is bEhá 'of what sort': bEhá dO pakkilalaámmu 'of what sort is your feeling (i.e. how do you feel?)'; jàbu na bEhá 'what kind of house?'; bEhá ulína 'of what sort is its beauty (i.e. how beautiful is it?)'

4.4.1.4. Pro-adjuncts

Interrogative pro-adjuncts are addígan 'when (future)' and naddígan 'when (in the past)': addígan hò láO 'when are you going?'; naddígan hò láO 'when did you go?'

Pro-adjuncts of location are formed from the propositions di 'at', tu 'to' and the demonstrative pronouns On 'this', i 'that', an 'that yonder', adú(i) 'that over there' with a linking /s/: disón 'here' (di + /s/ + On); tusón '(to) here'; disí 'there'; tusán '(to) there

*yonder*¹; etc.

4.4.1.5. Pro-clauses

The pro-clauses are ó10 'yes', dá0ŋ 'no' (note the difference from the negative particle daŋ 'not; no'), b0ása 'why' (some say bEása), ása 'so', jádi 'so; therefore', ípE 'so now' or ÓppE (from 1 + pE and On + pE), attÓŋ 'so now': ása | lá0 ma híta 'so let's go'; jádi | lá0 ma ibána 'therefore he went (away)'; ípE | bórhát ma hó 'so now, go thou'; attÓn | marÉddE ma híta 'so now, let's sing'.

4.4.2. Directive Particles

The directive particles fall into two classes: **prepositions** and **subordinating conjunctions**. The axis of a preposition is a word (mostly nominals), and that of a subordinating conjunction a clause.

4.4.2.1. Prepositions

The prepositions fall into several groups according to their common meaning or function: locative, instrumental, attributive, possessive-agentive, objective, comparative, and resultative.

4.4.2.1.1. Locative Prepositions

di 'at': di jábu 'in the house; at home'; di bútar 'at Butar (place name)'; di ibána 'with him (at his person)'.

tu 'to': tu Ónan 'to the market'; tu ibána 'to him'; tu gídjaŋ 'to the top; upward'; tu áha 'what for; to what'.

sían 'from': sían jábu 'from the house'; sían tóru 'from the bottom; from below'; paùli títtin sían más 'to make a ring out of gold'.

Some prepositional phrases function as a whole as prepositions of location. These phrases are made up of the preposition di or tu followed by a noun; e.g. di átas 'on (top)', tu bagásan 'into': di átas méja d0 bùkku í 'the book is on the table'; di bagásan lamári d0 bùkku í 'the book is in the cupboard'. Some phrasal prepositions are formed by di, tu, sían + a noun + ni 'of' (possessive preposition); e.g. di tòru ni méja 'at the bottom of (= under) the table'; di làbbuŋ ni jábu í 'at the side of (beside) the house'; the structure is 'locative preposition + possessed noun + possessive preposition (marker) + possessor noun'. When the possessor is a personal pronoun, it appears in its bound form suffixed to the possessed noun: di labbúkku 'beside me' (làbbuŋ + -hu); di pudím 'behind you' (púdi 'back' + -mu); di jolótta 'in front of us' (jolo 'front' + -ta).

Some locative prepositions are derived from verbs:

humaliŋ (UM- + √haliŋ, root of haliáŋi 'surround') 'around' adón d0 háu humaliŋ jábu í 'there are trees around the house'.

maddapótt0n 'to' (maŋ- + dapótt0n 'go towards') is used in letter addresses also; ró ma ibána maddapótt0n áu 'he came towards me'.

maddók 'to' (maŋ- + d0k 'say') is used only with some form of d0k 'say' preceding in the sentence: didók ma madd0k áu '(he) said to me...'

4.4.2.1.2. Instrumental Prepositions

dóh0t 'by means of; with': dilítciŋ ibána d0 áu d0h0t balóbas 'I was beaten by him with a ruler'.

marhite 'by means of; through': marhite hap0rsEá0n d0 jólma bóí malúa 'through faith, man can be saved': hutóŋ0s pE í marhite p0s 'I will send it by mail'. In the last example, sían can be substituted for marhite.

4.4.2.1.3. Attributive Preposition

There is one attributive preposition, na; it links a noun head with an attributive adjective, verb, numeral, or a phrase: háu na tíbbo 'tall tree'; hálak na mòdom í 'the sleeping man'; dakdànak na dùa í 'the two children'; hálak na manùh0r bùkku í 'the man who bought the book'; jábu na di hùta í 'the house in the village'.

4.4.2.1.4. Possessive-Agentive Preposition

This preposition, ni, marks possession when it is preceded by a noun: tàrup ni jábu 'roof of a house'; jábu ni ísE 'whose house'; jábu ni dóŋan 'a friend's house'; in these phrases the possessor follows the preposition.

When ni is preceded by the completive-participial or the promissory mode of a passive verb, it marks the noun following it as actor of the verb: tuhòr0n ni si ónd0 d0 í 'it will be bought by Ondo (boy's name)'; ón d0 tinùh0r ni d0ŋàtta í 'this is (what is) bought by our friend'.

4.4.2.1.5. Objective Preposition

This preposition, di, marks the goal of a transitive verb in constructions in which the object does not immediately follow the verb: manúkkun ma áu di hó 'I am asking you (now)'; na madjukkàti d0 hó di áu 'are you teasing me?' (in this style, more emphasis is given to 'teasing' than in the ordinary na madjukkàti áu d0 hó). Apparently, di performs the same function (goal marker) with special expressions like l0mó róha 'like' in l0mó d0 r0hána di hó 'he likes you', in which

lómú is an adjective (it can be inflected for degree), and rOhána 'his mind'.

sáma marks reciprocity; it seems to belong in this class. It is found, in most cases, between two identical pronouns, and is often preceded by a reciprocal verb: masisukkúnan ma hámi sama hámi 'we asked each other'; marbáda d0 nasída sama nasída 'they are quarreling amongst themselves'; híta sama híta ma í 'let it be between us'.

4.4.2.1.6. Comparative Prepositions

Of the three prepositions in this class, the last two are interchangeable, sían being the more common.

sÓŋ0n 'like': sÒŋ0n surusurúan d0 ulína 'her beauty is like an angel('s)', that is, 'she is as beautiful as an angel'.

sian 'than': humát0p d0 Ón sian mótor 'this is faster than a motor-car'.

ása 'than': duméngan d0 Ón asa án 'this is better than that yonder'.

4.4.2.1.7. Resultative Prepositions

There are two prepositions in this group: gábe 'for; as', báEn 'for; as'. The latter is also a verb, meaning 'make; put'; when it occurs in the clause, the preposition gabe is used: báEn ma ibána gábe d0ŋámmu 'make him (for) your friend'; túh0r dÈkkÈ í bàEn paŋanótta 'buy the fish for our food'; ibána d0 ditòdo nasída gábe kapàla káppuŋ 'he was chosen by them for village chief'.

4.4.2.1.8. Some Phrases Function as Prepositions

ála ni 'because of' (ála 'reason; cause') is a phrase that can function as a preposition: àla ni áha lá0 ibána 'because of what (= why) did he go?'; hupasídiŋ d0 ibána àla ni jukkákna 'he is avoided by me because of his naughtiness'.

taríŋ0t tu 'concerning' (tar- + íŋ0t 'remember'): hupaboa d0 tu ibána taríŋ0t tu pardalanakkí 'I told (to) him about my travel'.

4.4.2.2. Subordinating Conjunctions

In the transcriptions below and in the following sections, we will use additional symbols: . represents a falling terminal contour, , a sustained T.C., ? a rising T.C.

mó10 'if; when' (this has a more general function than the other words for 'if' below): (lá0 d0 áu,) mò10 ró ibána '(I will go,) when he comes'.

áng0 or ía or iáng0 'if (conditional); as for'; there is usually a

negative in one of the clauses: láo d0 áu, àng0 s0 ró ibána 'I will go, if he does not come'; àng0 sa0nnári, dàñ oloákku bE hú 'as for now, I won't obey thee any more'.

áut or autcúra 'if' (contrary to fact): láo d0 áu, àut adón hEpÉkku 'I would go, if I had money (literally, 'if my money existed')'. The same effect is produced by means of the verb auxiliary sugári (contrary to fact), or the combination àut sugári: láo d0 áu, sugàri adón hEpÉkku 'I would go, had I money'.

nañ 'although': nàñ ró ibána, láo d0 áu 'although he comes (may come), I will go'; láo d0 áu, nàñ s0 ró ibána 'I will go, although he does not come'.

agía 'even if' (stronger than nañ); it is mostly used with a negative: láo d0 áu, agía s0 sáE hEpÉkku 'I will go, even if I have not enough money'.

disí 'as soon as': hujóu pE hó, disí sáhat ibána 'I will call you, as soon as he arrives'.

tíkki or tágan 'while' (tíkki as a noun means 'time'): bórhat ma hámi, tíkki makkatái nasída di jábu 'we left, while they were talking in the house'. In a special style, this word can occur before an independent clause, in which case the other clause is dependent: tíkki makkatái d0 nasída di jábu, bórhat hámi 'it is while they were talking in the house, that we left'.

dun 'after': bórhat ma hámi, dùñ ró ibána 'we left, after he came'.

paíttE or paíma 'until' (both occur as verbs, meaning 'wait'): únañ marísap hò, paíma sáhat híta tu jábu 'don't smoke until we arrive at home'.

marádu 'so ... that': lálap d0 ibána marmEámi, marádu s0 díñ0t múlak tu jábu 'he is so absorbed in playing, that he forgets to go back home'.

ubbáEn (or dibáEn) usually followed by na introduces a resultative clause; the preceding clause is also usually introduced by na: na hòna údan d0 ibána, ubbàEn na marsáhit '(it was because) he was caught in the rain, that he is sick'. We also find both clauses introduced by ubbáEn na: ubbàEn na hòna údan d0 ibána, ubbàEn na marsáhit (same meaning).

ala 'because': mársak d0 ibána, àla s0 adón hEpÉñña 'he is sad, because he has no money'.

ái 'for; because': dàñ tallÈan áu í tu hó, ài s0 bukkúkku í 'I can't give it to thee, for it is not mine'; únañ p0la mabfar hò, ài na jínak d0 bían í 'thou needn't be afraid, because the dog is a tame one'.

apála 'not ..., but ...' (signals contrast): apála ibána manùh0r bùkku í, áu d0 'it is not he (who) bought the book, but I (did)'.

sótun 'lest': únaṇ segái bükku í, sòtun dimurúhi hó 'don't ruin the book, lest you be rebuked (by him)'.

ása 'so that': tibú ma hó ró, ása rap lá0 híta 'come early, so that we go together'.

mánaṇ 'whether': dibòt0 hó d0 mánaṇ naun ró ibána? 'do you know whether he has come?'

sána 'before (it is too late and ...)': burí ma pìngan í, sàṇa ró inátta 'wash the dishes before (it is too late and) mother comes'. This implication of threat is also seen in: sàṇa dibàEn hó í sahàli nári 'if you do it once again (you will suffer the consequences)!'; sàṇa s0 tibú h0 ró 'if you don't come early!'

4.4.3. Modal Particles

The modal particles fall into two classes: **predicative particles** and **imperative particles**. Predicative particles are found in non-imperative clauses. Each occurs with a certain predication mode. Imperative particles are found in imperative clauses; they indicate various shades of the imperative. If no translation is possible, we will indicate the mode with which the particles are associated.

4.4.3.1. Predicative Particles

The first three come between the predicator and the subject, the next two occur at the beginning of the clause.

d0 'affirmative': módom d0 ibána 'he is sleeping'; huída d0 ibána 'he is seen by me'; hó d0 mabbàEn ón? 'did YOU do this?'; ringás d0 ibána 'he is diligent'; gúru d0 ibána 'he is a teacher'.

ma 'narrative': marsiájar ma ibána '(and so) he studied'; ditùh0r ibána ma bükku í '(and then) he bought the book'.

pE 'concessive': hugárar pE bükku í '(I promise) I will pay for the book'; mardàlan pát pE áu '(it's all right) I will walk'; ón pE hulÉan 'I will just give this (i.e., there is no need to find another)'.

The first example can be expressed by the promissory mode of the verb: gararókku d0 bükku í 'I will pay for the book'; the use of pE is felt to express a stronger commitment.

daṇ 'not; no' (dáda and dátun are more emphatic negative particles; cf. s0 below): dàṇ huída ibána 'he is not seen by me (= I don't see him)'; dàṇ gúru ibána 'he is not a teacher'.

núnṇa 'already' (explicit completive): núnṇa lá0 ibána 'he has already gone'; núnṇa hujóu ibána 'he has already been called by me!'

The above predicative particles are mutually exclusive. A concessive negative (corresponding to pE) is dàṇ póla: dàṇ póla ró hó 'you need

not come'; dàṅ p0la hubúat í 'it won't be taken by me (= I won't take it)'. The negative corresponding to núṅṅa 'completive' is dǎṅ bE 'no longer': dàṅ bE jukkát ibána 'he is no longer naughty'.

Except in special styles, these particles occur only in independent clauses. In subordinate clauses or phrases, the negative particle is s0: lá0 d0 áu, m0l0 s0 ró ibána 'I will go, when he does not come'; the completive is duṅ (náuṅ in nominal clauses): lá0 d0 áu, m0l0 duṅ ró ibána 'I will go, when he has come'; bÉrEṅ j0l0 màṅṅa nauṅ ró ibána 'go see whether he has come'.

s0 is also used in independent clauses as an emphatic negative; cf. dàṅ ad0ṅ hEpÉkku 'I have no money' and s0 ad0ṅ hEpÉkku 'I HAVE no money'.

dOpÉ 'still' (progressive): makkatái dOpE nasída 'they are still talking'. This belongs marginally to this class, because it is mutually exclusive with the non-negative particles above; however, it can occur with the negative daṅ: daṅ máṅṅa dOpE nasída 'they have not yet eaten', and it is found in subordinate clauses: lá0 ma híta, m0l0 makkatái dOpE nasída 'let's go, if they are still talking'.

4.4.3.2. Imperative Particles

An imperative clause occurs without a particle: ró h0 'come (thou)!'; gárar utaṅmí 'pay your debt!'. Other shades of the imperative contain a particle. The first two come after the verb (j0l0 can come after the subject h0 'thou'), and the last three occur initially in the clause.

j0l0 'mild imperative': ró j0l0 hamú tusón 'come (you) here, please'; búat j0l0 bùkku í 'take the book, please'; ró j0l0 hó or ró h0 j0l0 'please, come'.

ma 'polite imperative; request': húddul ma hamú 'please, have a seat'; túh0r ma bùkku í '(will you) buy the book, please'.

nàng0 'suggestive': nàng0 mardàlan pát h0 'why don't you walk (on foot)', nàng0 marsógot h0 ró 'why don't you come tomorrow!'

unaṅ 'don't' (negative imperative): unaṅ ró h0 tusón 'don't come here'; unaṅ túh0r bùkku í 'don't buy the book'. Milder shades of prohibitions are effected by a combination of unaṅ and ma or j0l0: unaṅ ma ró h0 tusón or unaṅ j0l0 ró h0 tusón 'please, don't come here'. Like daṅ 'not' above, unaṅ combines with p0la and bE: unaṅ p0la ró h0 tusón 'you don't have to come here!'; unaṅ bE ró h0 tusón 'don't come here any more!'

s0tuṅ 'don't' (emphatic negative imperative): s0tuṅ ró h0 tusón 'don't come here by any means!'. Unlike unaṅ above, it appears with an inflected transitive verb with the subject expressed: s0tuṅ ditùh0r hó bùkku í 'don't buy the book by any means'.

4.4.4. Connectives

Connectives are particles that function as markers of elements in co-ordinate construction. The first three mark clauses only.

dùŋi 'and; then': másuk ma ibána, dùŋi didók ma, 'he came in, and (he) said ...'; másuk ma ibána tu jábu, dùŋi harúar ma ibána musÉ 'he went into the house, and he came out again'.

alái 'but': hujóu d0 ibána, alai dan ól0 ró 'I called him, but (he) did not want to come'.

hapÉ 'but' (stronger than alái, and connotes surprise or irritation): dan hÉá huurúpi ibána, hapÉ ól0 d0 ibána manurúpi áu 'I had never helped him, but he was willing to help me'.

jála 'and': burjú jala ringás d0 ibána 'he is good and diligent'; dibúkka ma pittu í, jála harúar ma ibána 'he opened the door, and he came out'.

huhut 'and at the same time': húddul ma ibána, huhút mÉkkEí 'he sat down, and (at the same time he) smiled'; lója huhut malÉ d0 dihilàla ibána 'he feels tired and hungry'.

mànaŋ 'or': Ón d0 tuh0rómmu? mànaŋ án. 'are you going to buy this or that?'; gararómmu d0 utaŋmí? mànaŋ hualúh0n ma hó. 'are you going to pay your debt, or shall I sue you?'; angím mànaŋ akkáŋmu d0 Ón. 'is this (= he) your younger (brother) or your older brother?'

dóh0t 'and' (does not mark clauses): lá0 d0 áu dóh0t ibána tu Ónan 'he and I went to the market'; dibòt0 hó d0 manúrat dòh0t madjáha? 'do you know (how to) write and read?'; búat ma na di gídjan dòh0t na di tóru ni mèja í 'please, take what is on and under the table'.

4.4.5. Auxiliaries

Auxiliaries are attributive particles. They fall into several classes. The first consists of particles that modify any form-class which can fill the position before the predicative particle d0 in the clause. In the direct word order, this position is filled by the predicator, which, in most cases, is a verb. We will call them **predicator auxiliaries**. The other classes consist of particles modifying one form-class. We will group them according to the head.

4.4.5.1. Predicator Auxiliaries

They are placed immediately before or after the head.

pítt0r 'immediately': pítt0r bórhát ma ibána 'he left immediately'; pítt0r kapála d0 ibána 'he immediately (became) a chief'.

íkk0n 'must': íkk0n lá0 d0 hó tu húta 'you must go to the village'; íkk0n dókt0r d0 hó 'you must (become; be) a doctor'; íkk0n matcài m0rá

d0 ibána 'he must (be) very rich'.

sòr 'almost': sòr lá0 ma ibána 'he almost left'; sòr gúru ma ibána 'he almost (became) a teacher'; sòr sappúlu ma dilÉan ibána 'he almost gave ten'.

gábe 'become; result in': gábe mékke1 d0 ibána 'he winds up laughing'; gábe gúru d0 ibána 'he became a teacher (in the end)'; nùŋŋa gábe burjú ibána 'he has become nice'. This particle also occurs as a preposition (see §4.4.2.1.7. above).

sái 'always': sái mardàlan pát d0 ibána tu sikkóla 'he always walks (on foot) to school'; sái jukkát d0pE dakdànak í 'the child is still always (= still remains) naughty'.

náEŋ 'want to (become)': náEŋ lá0 d0 ibána 'he wants to go'; náEŋ túkkaŋ d0 ibána 'he wants (to become) a carpenter'.

be 'distributive': là0 bé d0 nasída 'each of them went away; they went each his own way'; gùru bé d0 nasída 'they are teachers, each of them'; m0rà bé d0 nasída 'they are rich, each of them'. It also occurs with the distributive mode of a transitive verb after the object: masigàrar utàŋŋa bé ma híta 'let each of us pay his own debt'. When there is a numeral in the clause, be usually comes after the numeral: tòlu bé d0 nasída hulÉan or hulÉan d0 nasída tòlu bé 'I gave them three each'.

musÉ 'again': là0 musÉ ma ibána '(so) he went again'; dijòu musÉ ma d0ŋàŋna í 'he called his friend again'. This auxiliary can also occupy another position: lá0 ma ibána musÉ or lá0 ma musÉ ibána (same meaning as above with a slight difference in emphasis).

hián 'used to (be); previously': mòdom hián d0 ibána 'he was asleep previously'; gùru hián d0 ibána 'he used to be a teacher'; m0rà hián d0 ibána 'he used to be rich'.

tar 'pretend to' (this is used with a verb head only; see below for a homophonous form with an adjective): tar táŋis d0 ibána 'he pretended to cry'; tàr dilíftciŋ ma angína í '(so) he pretended to beat his younger brother'.

4.4.5.2. Noun Auxiliaries

They all come before the noun head.

ákka 'plural': ákka jábu 'houses'; lÉan ma ón tu ákka d0ŋammí 'give this to your friends'. It also pluralises nominal phrases lÉan0n d0 ón tu ákka na pógos 'this will be given to the poor'. The plural auxiliary is omitted, when there are other clues concerning the number in the context: godán d0 jábu disón 'there are many houses here'.

nása 'all': nása jólma 'all men (persons)'; diplo ibána d0 nása d0ŋàŋna 'all his friends are invited by him'.

hòlan 'only': hòlan róti d0 dipànan nasída 'only bread is eaten by them (i.e., they eat only bread)'; when this auxiliary enters in construction with a plural-marked noun, it precedes ákka: hòlan ákka d0ḡánna d0 dipìò ibána 'only his friends are invited by him'.

gánup 'each (every)': marsikkóla d0 ibána gánup ári 'he goes to school every day'; godán d0 dilÈan ibána tu gánup hálak 'he gave much to each person'.

apála 'real' is used only with nouns of relationship to indicate primary relationship: apála angína d0 í 'that is his real younger brother (or, her real younger sister)', that is, 'not just a person called by that term'.

si is used before personal proper nouns, but not in direct address: si tio 'Tio (girl's name, meaning 'clear'))'. It is also not used with teknonymous proper nouns: nattfo 'Mother of Tio'.

4.4.5.3. Adjective Auxiliaries

All these auxiliaries indicate degree.

matcái 'very': matcài tìbbó 'very high'; matcài dengán 'very good'.

saháli 'very' (from sa- + háli 'time'): tìbbò saháli 'very high'.

situtú 'very' (from si- + tutú 'true'): tìbbò situtú 'very (truly) high'.

tun 'extremely': tun ulí d0 ibána 'she is extremely beautiful'.

lòbi 'more': lòbi tìbbó 'higher'. This is used for stylistic purposes (emphasis) instead of the inflected comparative: tumíbbò 'higher'. We also find the combination lòbi tumíbbò 'still higher; much higher'.

lam 'more and more': lám m0rá d0 ibána 'he is (becoming) richer and richer'; lám lelén, lám pogós 'the longer, the poorer'.

apála 'really; truly'; more commonly found with attributive adjectives: jàbu apála na tìbbò 'a really tall house'.

hían 'extremely': ùli hían d0 ibána 'she is extremely beautiful'; this expresses a higher degree than tun above.

tar 'somewhat': tàr m0rá d0 ibána 'he is somewhat rich'.

A very high degree of some quality is also expressed by preposed *bólat ni* (literally, 'the limit of'): *bòlat ni ùli ibána 'she is extremely beautiful'*; and also by *na sái* (sa- + i 'as that'): *na sài gógo ibána 'he is so strong'*; in which *ùli* and *gógo* are nouns.

4.4.5.4. Adjunct Auxiliary

marÉak 'toward; nearly': marÈak b0tári d0 ibána múlak 'he came back toward the evening'; núḡḡa marÈak manógot 'it is (already) nearly morning'.

4.4.5.5. Substitute Auxiliary

Of the noun auxiliaries above, hòlan 'only', gánup 'each', and apála 'real' are also used with some pronouns: hòlan ibána 'only he'; gánup híta 'each of us'; apála ibána 'he himself (not somebody else)'; hòlan áha 'only what'; hòlan tólu 'three only'.

4.4.5.5.1. Numeral Auxiliaries

nári (or náí): búat j0l0 tòlu nári 'take three more, please'; adón d0 lĩma náí 'there are five more'.

náng0 'at least': túh0r ma náng0 ópat 'please, buy at least four'; dòn adón hÉpÉkku náng0 sarupía 'I don't even have one rupiah (money)'; (cf. náng0 as an imperative particle, §4.4.3.2.).

hirahíra 'approximately': hirahíra sarátus 'about one hundred'; this is also used with pro-adjuncts: hirahíra tusón d0 madábu '(it) fell approximately here'.

Two auxiliaries are used with interrogative substitutes (pronoun, proverb). They come before the head and make it emphatic:

agía: bóí d0 masuk agía ísE 'anybody may enter'; dòn adón agía áha (or agiáha); 'there is nothing whatsoever'; this is more emphatic than mánaŋ below;

mánaŋ: dòn adón huída mánaŋ áha 'I don't see anything whatever'; búat ma mánaŋ díá 'please take whichever (you like)'; dán dibót0 manan marhúa 'he doesn't know to do whatever (= he doesn't know how to do anything whatever)'.

4.4.5.6. Two auxiliaries function as attributes to negative modal particles: póla and bE (see examples above §4.1.).

4.4.5.7. Some particles, or combinations of them, seem to modify the clause rather than just part of it. They express some emotion or attitude of the speaker, or a shift in predication mode.

tun (... d0) 'amazement': tún díál0 d0 nat0ràsna í 'he DID (have the heart to) oppose his parent!'.

úa (...d0) 'impatience': úa búat d0 búkku í 'DO take the book!'; úa só d0 hó 'DO be still!'.

tahÉ 'denial': bukkúna d0 í tahÉ 'it IS his book (although he said it was not)'.

nián 'regret; emphasis': láo d0 ibána nián 'he actually did go (but to no avail)'; láo ma hó nián 'DO go now'.

(sótun ...) má0n 'doubt': s0tun múruk ibána má0n 'he MIGHT get angry'.

ásal na 'derogatory; discrepancy between name and reality': àsal na paragáma d0 ibána 'he is just by name a religious man!'.
 póla ma 'mild disapproval': póla ma dimurùhi hó anakmí 'thou needn't have scolded thy son!'.
 (ma ...) attón 'satisfaction': hutídjú ma ibána attón '(and so) I punched him (which serves him right)'.

4.5. STRUCTURE OF WORDS

Words consist of a single morpheme, an underlying form (stem) plus inflectional affixes, an underlying form (base) plus derivational affixes, or two forms in composition. Inflection is discussed above (§4.2.), and derivation below (§4.6.).

4.5.1. Composition

There are four types of composition: the first involves a demonstrative pronoun, the second involves a numeral, the third type is found in teknonymous personal proper nouns, the fourth involves two nouns in attributive or co-ordinate construction.

(1) The case of enclitic demonstratives is mentioned in §3.3.2., IIS9: jabukkí '(that) my house', from jabúkku 'my house' plus i 'that; the'.

In the formation of pro-adjuncts of location, the prepositions di 'at' and tu 'to' are linked by /s/ to a following demonstrative (On 'this', i 'that', an 'that yonder', adú(i) 'that over there'): e.g. disón 'here'; (see §4.4.1.4. above).

The demonstrative i 'that' occurs linked to a preceding disi 'there' by /n/: (na) disiní 'that (what is) there'. This word without the connecting phonemes is di + i + i. The demonstrative On 'this' is connected with a preceding i 'that' by /n/: na hùd0k í '(that) what I said': na hudòk inón '(that) what I just said'; jabukkí '(that) my house': jabukkinón 'my house just mentioned'. It may, historically be this type of combination that we have in some adjuncts of time (see also §4.3.4.); e.g. nahinán 'formerly', from na + i + an.

(2) The linking of numerals to free and bound nouns of measure is mentioned above (§4.3.): sakílo 'one kilogram'; tòlukkabbóna 'three pieces (used for plants)'.

(3) Teknonymous personal proper nouns are formed by preposing náí 'mother of', ámani 'father of', óppuni 'grandparent of' to a name (see §4.3.1.). When the name begins with /s t d j/, the formatives are proclitically linked to the name in the shapes nan-, aman-, and oppun- respectively; thus: naddaúlat 'Mother of Daulat', amatcáhat 'Father of

of which *huddúlan* is a secondary and *b0rtían* is a primary derivative.

The morphophonemic symbols used in the following examples are explained in §4.6.3. below:

Prefix + root:

- ha- + √subba : *hasúbba* 'indigo plant' (noun);
- an- + √dalú : *addálu* 'pounder (made of wood)' (noun);
- maŋ- + √appir : *maŋáppir* 'feel cramp' (intransitive verb);

Partial reduplication + root:

- CV- + √tariŋ : *tatáriŋ* 'hearth'
- + √sagun : *saságun* 'snack made of rice flour, coconut, and sugar' (noun);

Initial + medial + root:

- sa- + Ltn- + √pEpE : *saruppÉpE* 'a creeper plant' (noun);
- si- + maŋ- + √hudap : *simakkúdap* 'mouth' (noun);
- UM- + aL- + √gutcaŋ : *gumalútcaŋ* 'be in violent wavy motion' (intr. verb);
- ma- + ht- + √l0l0ŋ : *mahil0l0ŋ* 'desert husband' (intrans. verb);

Root + suffix:

- √tibal + -an : *tibálan* 'sheaf (of rice)' (noun);
- √bataŋ + -i : *batáni* '(rice field) dike' (noun);
- √beuk + -an : *beúhan* 'rot (said of eggs)' (intransitive verb);

Prefix + root + suffix:

- par- + √badda + -an : *parbaddáan* 'graveyard' (noun).

4.6.2.1. Root Compounds

There are two types of root compounds: doubling (full reduplication) and a compound of two different roots.

Doubling of roots:

- R + √lappu : *lappuláppu* 'butterfly' (noun);
- √uŋut : *uŋutúŋut* 'complaint' (noun);
- √hEa : *hEahEá* 'pant' (intransitive verb);
- √jalak : *jalakjalák* 'wander about searching for something' (intransitive verb).

Compound of two different roots:

- √atta + √jau : *attajáu* 'tree bearing pearlike fruit' (noun).

Root Compounding and Affixation:

- R + √uliŋ + -an : *hulikkulíŋan* 'riddle' (noun);
- + √uti + -an : *utiutían* 'machinations; tricks' (noun).

4.6.3. Noun Stems

Nouns are simple stems, primary derivatives, or secondary derivatives. Derived nouns fall into classes depending on the affix; when useful, we sub-classify on the basis of meaning.

The morphophonemic symbols found in this and the following sections (and not in §3) are:

L : /r/ before a root containing /r/;
/l/ otherwise.

† : /u/ before a root with /i e É a/ (non-back vowels) in first syllable;
/i/ otherwise.

() : a parenthesised symbol may be represented by zero.

4.6.3.1. Simple Noun Stems

Examples:

jábu 'house'; dóŋan 'friend'; táŋan 'hand';
bfaŋ 'dog'; údan 'rain'; pat 'foot; leg'.

4.6.3.2. Primary Derived Noun Stems

Three or four examples are given for each type or class; if there is less than three, it means that the writer has not been able to think of (or find) other examples.

(1) Root Doubling:

(animals): b0r0ŋb0r0ŋ 'wasp', lappuláppu 'butterfly', uttuŋúttuŋ 'a non-stringing kind of beetle'.

(other): habahába 'thunderstorm', uŋutúŋut 'complaint', 0saŋ0saŋ 'chin', sErasÉra 'budding jack fruit'.

Doubling + -an: hulikkulíŋan 'riddle', utiutian 'trick, machinations', uníuníŋan '(musical) instrument'.

(2) Reduplication + Root: tatáriŋ 'hearth', saságun 'delicacy made of rice flour, coconut, and sugar'.

(3) Root Compounds: attíŋáŋ0 'winged insect found in rice fields', addalíman 'a plant', jŋŋulmÉ0ŋ 'beetle often found in dung heaps'.

(4a) si- + Root:

(animals): sigák 'raven', sipÉt 'woodpecker', sij0bbiŋ 'spider', sipásiŋ 'a water insect'.

(plants): sihúpi, siála, sipÓrh0t (means also 'hair bun' from the ancient custom of putting the fragrant leaves of the plant in women's hair buns).

(family names): sih0bbiŋ, siráit, sitórus.

(place names): sipírok, sipúltak, silídduŋ, sigót0m.

(other): sirábun 'ash', silóŋiŋ 'complete silence'.

(4b) si- + complex root:

sitapáni 'drainage cut in a rice field dike', (from √-tapaŋ + -i);
simaló10ŋ 'eye' (from -ma- + √-1010ŋ), sipariama 'place name',
simakkúdap 'mouth' (from -maŋ- + √-hudap).

(5) (s)+(n)- + Root: sibbóra 'lead', lbbúlu 'body hair',
sulábbak 'animal like a porcupine', subbíá 'arrow'.

(6) pa- + Root: pahála 'good fortune', parápat 'place name';

(6a) pa-(h)+- + Root: pauséaŋ 'dowry', pahiló10ŋ 'deserter of
husband'; the corresponding verbs begin with ma-(h)+-.

(7) par- + Root: parbúe 'fruit'.

(7a) par- + Root + Suffix: parbaddáan 'graveyard', paruhá0n 'family
relationship' (from par- + √hua + -0n, with metathesis).

(8) paŋ- + Root: pakkúliŋ 'sound', paŋáppir 'cramp', patcádi
'way of ceasing'; these nouns have a corresponding verb with the prefix
maŋ-.

(9) ha- + Root: hasúbba 'indigo', halífo 'plant bearing pearlike
fruit', hasí0r 'plant yielding spice', haúma 'rice field', halílu
'shadow'.

(9a) ha- + Root + -an: hamisáran and haluáan (symbols in Batak
script), haluá0n 'freedom, salvation' (the corresponding verb is malúa
'be free').

(10a) <aL> + Root (aL is infixed after the first consonant of the
following syllable): balátuk 'ladder', baréaŋ 'sulphur', alógo 'wind',
saléddan 'scarf', haríhir 'nose ring (of animals)'.

(10b) <aL-> + (h)l- + Root: halisúksuŋ 'whirlwind', hariára 'a
banyanlike tree'.

(10c) <aL-> + (h)i + Root + -an: halibutóŋan 'rainbow'.

(11) lan- + Root: lakkítaŋ 'snail, shell', labbíak 'belly',
lattéuŋ 'a thorny plant bearing round eggplant-like fruit'.

(12) an- + Root:

(birds): appórik 'sparrow', abbaróba 'bird with yellow breast',
addúhur 'pigeon-like bird'.

(other): abbálaŋ 'sling (as weapon)', addálu 'pestle (for pounding
rice)', addóra 'chest (body part)'.

(13) Root + -an: tibálan 'sheaf (of rice)'; lapátan 'meaning';
sibáran 'fate'; dalíhan 'tripod potstand (in cooking)'; b0rtían 'womb'.

(14) Root + -i: batáhi 'staff (of a shepherd)'; batáni 'dike (in
a rice field)'; these should, perhaps, be called secondary derivatives:
bátak is an obsolescent verb, meaning 'riding on horseback'; bátaŋ
occurs as a noun, meaning 'coffin; trunk (of a tree)'.

4.6.3.3. Secondary Derived Noun Stems

In the following examples: N = noun, V1 = intransitive verb; Vt = transitive verb, A = adjective.

(1a) Doubling of noun; the meaning is 'something resembling the base' or a diminutive: *hóda* 'horse', *hodahóda* 'something resembling (or used as) a horse (i.e. a saw horse)'; *dólok* 'mountain', *dolokdólok* 'hill'; *lánit* 'sky', *laniklánit* 'ceiling'.

(1b) Doubling of transitive verb stem; the meaning is 'something with which the action expressed by the verb is carried out (instrument)': *gugagúga* 'stirring ladle' (*gúga* 'stir'); *litciŋlítciŋ* 'whip; whipping rod' (*lítciŋ* 'to whip'); *sukcutcúkkun* 'question' (*súkkun* 'ask'); *togutógu* 'rope by which one leads an animal; leash' (*tógu* 'pull').

(2a) *si-* + N: *sitúak* 'honey; sweet essence' (*túak* 'palm wine'); *siláŋge* 'place name' (*láŋge* 'a plant of same family as ginger').

(2b) *si-* + A: *silÉpE* '(water-buffalo) with downturned horns' (*lEpÉ* 'drooping'); *sihúnik* '(yam) with yellow meat' (*huník* 'yellow'); *sibádar* 'an albino' (*badár* 'creamy (white)').

(2c) *si-* + Phrase: *sipelebégu* 'heathen' (*péle* 'to worship', *bégu* 'spirit; ghost'); *sipOŋgOlúlu* 'headhunter' (*pOŋgOl* 'cut', *ulu* 'head'); *sibOttarmáta* 'a white person (European)' (*bOttár* 'white', *máta* 'eye').

(2d) *si-* + Promissory mode of Vt: *sijoúOn* '(person) to be called' (*joúOn* 'will be called'); *situhóron* '(things) to be bought' (*tuhóron* 'will be bought'); *sipaŋánOn* 'food' (*paŋánOn* 'will be eaten').

(3a) *par-* + N; it has the meaning 'actor' or 'owner' (cf. V1 beginning in *mar-*): *pargúru* 'pupil' (*gúru* 'teacher', *margúru* 'to study (V1)'); *pardálan* 'walker' (*dálan* 'road; path', *mardálan* 'to walk'); *parhaúma* 'owner of rice field' (*haúma* 'rice field'); *pardÉkke* 'fisherman' (*dÉkke* 'fish').

(3b) *par-* + V1: *parmínun* 'a habitual drinker' (*mínun* 'to drink'); *parmábuk* 'a habitual drunkard' (*mábuk* 'be drunk'); *parhúddul* 'a person who likes to sit idle' (*húddul* 'sit').

(3c) *par-* + Vt: *parÓppa* 'sling for (or person) carrying a child on the back' (*Óppa* 'carry a child on back'); *partídju* 'pugilist' (*tídju* 'punch'); *parísap* 'smoker' (*ísap* 'to smoke (something)').

(3d) *par-* + Numeral (cf. homophonous Vt); these serve as denominators of fractions: *parÓpat* 'fourth' (*Ópat* 'four'; *tòlu parÓpat* 'three-fourths').

(4a) *par-* + N + *-an*; it makes nouns of place: *pargadÓŋan* 'yam field'; *parjabúan* 'site of house'; *pardEkkÉan* 'fishing place'.

(4b) *par-* + N + *-On*; it makes abstract nouns: *parŋolúOn* 'way of life'; *pardOŋánOn* 'friendship'; *pardEkkÉOn* '(way of) fishing'.

(4c) par- + Vt + -an; denotes place or act; these nouns have corresponding intransitive verbs of the form mar- + Vt: paridían 'bathing place' (from dídi 'to bathe (a person)', which probably developed from *diidi with stem idi); paralóan 'fight, strife' (from álo 'oppose'); pardegéan '(rice) threshing place' (from dége 'tread').

(5a) ha- + N + -On; denotes quality: hagurúOn 'teachership'; harajáOn 'kingship; kingdom'; hadatúOn 'magicianship'.

(5b) ha- + A + -On: haulfOn 'beauty'; harOáOn 'ugliness'; haOtóOn 'folly, foolishness'.

(5c) ha- + V1 + -an: hasuddútan 'west (place of sunset)'; habitcáran 'east (place of sunrise)'; haróan 'feast (place or event to which many people come)'; halaóan 'expenditure (of money)'.

(6) ha + R + V1: habObórhath 'departure' (bórhath 'depart'); hatutúbu 'birth' (túbu 'be born'); hasasáhat 'arrival' (sahath 'arrive').

(7) NI- + Vt: sináOk 'small snack made of roasted rice'; pináhan '(household) animal' (páhan 'feed (an animal)'); sinódduk 'husband' (sódduk 'keep in the house'); iddáhan 'cooked rice' has an irregular form (regular would be *nidahan) from dáhan 'to cook'.

(8) NI- + ha- + A: denotes degree of some quality: hinaúli '(degree of) beauty'; hinadáó '(extent of) distance'; hinamálo '(degree of) cleverness'.

(9a) paŋ- + N: denotes actor; most of these nouns have a corresponding V1 verb of the form maŋ- + N: panórtOr 'dancer', panúrat 'writer', panúhOr 'buyer', paŋaléle 'chaser'.

(9b) paŋ- + Vt + -an: paŋOmóan 'profit' (from ómo 'gain'); panurátan 'place for writing'; paŋalOppáan 'place for cooking'; paddegéan 'foot print'.

(10a) N + -an: pagáran 'outlying settlement; parish' (from págar 'fence'); paráŋan 'follower; subordinate' (from póraŋ 'war' with a vowel change); batáŋan 'main support' (from bátaŋ 'tree trunk').

(10b) V + -an: pOppáran 'descendant' (from póppar 'have as descendant'); tudósan 'parable' similarity' (from tudós 'be similar to'); huddúlan 'seat' (from húddul 'sit'); paŋáŋan 'dish; plate' (páŋan 'eat'); tanóman 'grave' (tánOm 'bury'); gObbáran 'picture' (góbbar 'draw (a picture)').

(11) Vt + -On: gOtílon 'harvest' (from gótil 'to harvest'); sabúrOn 'sowing season'; uláOn 'work; feast' (from úla 'to work'); halÉOn 'famine' (presumably from háli 'dig'; so, a time when digging yam is necessary, because there is not enough rice stock).

When the verb ends in the suffix -i, the nominalising suffix has the shape -an: hapósan 'a trustworthy person' (hapósi 'to trust', pOs 'trust'); habiáran 'fear; something to be feared' (habiári 'to fear',

bíar 'fear'); inánan 'place; dwelling' (ináni 'dwell in; inhabit from /inan).

(12) Miscellaneous: si- + ha- + buní 'hidden' + -an: sihabunfan 'secret place'; par- + si- + tíjur 'spittle' + -an: parsitijúran 'spittoon'; par- + ta- + buní 'hidden' + -an: partabunfan 'hiding place'. The last two examples without the suffix are normal noun forms corresponding to verb forms in mar-; in these cases, marsitíjur 'to spit', and martabúni 'to hide (oneself)'; paṇaláho 'conduct; behaviour' (presumably from láho, an older form of láo 'go'); parúlan 'behaviour, act' (from par- + úla 'do' + -an with a coalescence of vowel).

4.6.4. Verb Stems

The two classes of verbs, intransitive and transitive, also differ in derivation.

4.6.4.1. Intransitive Verb Stems

Intransitive verbs are simple stems, primary derivatives, or secondary derivatives. They fall into two classes. The first class consists of those derived verbs beginning in /m/, which correspond to nouns beginning in /p/ denoting actor, action, or instrument; e.g. mardálan 'to walk' : pardálan 'walker'. We will call these verbs active. Other intransitive verbs are then stative. This classification intersects the division into primary and secondary derivatives.

4.6.4.1.1. Simple Intransitive Verbs

Examples: láo 'go'; ṇot 'wake up'; hábaṇ 'fly'; hḡddul 'sit'.

4.6.4.1.2. Primary Derived Intransitive Verbs

Active Intransitive Verbs:

(1a) maṇ- + root: makkúliṇ 'speak; make a sound' (from √huliṇ, also found in hulikkulíṇan 'riddle'); matcádi 'cease' (from √sadi, also in pasádi 'stop' Vt); maṇítcu 'fake; act misleadingly' (from √itcu, presumably the same morpheme as in itcuátca, a challenging exclamation in a riddle game meaning something like 'guess what').

(1b) maṇ- + tṇ- + root: maṇutcálṇaṇ 'sprawl' (from √-salṇaṇ); maṇitcúrut 'withdraw' (from √-surut, also found in the stative verb sumúrut 'retreat'); maṇiddórap 'suffer deeply' (from √-d0rap).

(2) ma- + (h)t- + root: mahuEppE (from √-EppE), mahil0l0ṇ (from √-l0l0ṇ), both mean 'desert husband'; mahusáEm 'sacrifice in order to avert misfortune' (from √-saEm); maitúrun 'leave parents' home and

stay with future in-laws (said of a girl, who wants to force her parents to agree with her choice of future husband)', from $\sqrt{-}$ turun 'go down'.

(3) $\text{mar-} + \text{LV}^1\text{-} + \text{CV}^1\text{C}$ (from onomatopoetic roots of the form CVC; the meaning is 'make such and such a sound'; L is a morphophoneme as in §4.6.3.; v^1 indicates that the vowel is the same as in the root): mallitík; mallasák; mardapár; mardubúr.

Stative Intransitive Verbs:

(4) $\text{m(a)-} + \text{root}$: mabálu 'be a widow or widower' (from $\sqrt{-}$ balu); míttop 'extinguished, go out' (from $\sqrt{-}$ ittop with a loss of vowel in the prefix; the same root is found in the transitive verb ittópi 'extinguish'); múllop 'emerge' (from $\sqrt{-}$ ullop, also found in the transitive verb paúllop 'make to emerge'); míjur 'descend' (from $\sqrt{-}$ ijur, also found in paíjur 'to lower'). The following verbs seem to belong here: másuk 'enter', mátE 'die', másak 'be done (of something cooked)', perhaps with a loss of vowel or syllable.

(5) $\text{UM-} + \text{root}$: humórdit 'make a slight movement' (from $\sqrt{-}$ hórdit); humítir 'tremble'; humísik 'make a rustling sound'; there seems to be a relationship between this formative UM- and the completive modal affix UM- in transitive verbs.

(6) $\text{UM-} + \text{ha-} + \text{root}$: humatíltal 'prance about' (from root $\sqrt{-}$ tiltal; humatíon 'run to and fro excitedly'; humasÉOk 'run about frantically'.

(7) Reduplication + root + -On: hEhEámOn 'yawn' (from root $\sqrt{-}$ hEam); sOsObbÓpOn 'sob'.

(8) $\text{ha-} + \text{root}$: hasÉa 'be useful' (from $\sqrt{-}$ sEa, presumably the same morpheme found in parhasÉan 'to consider valuable; save from being thrown away'); harúar 'emerge; come forth; be over (said of school or church service)', from rúar, whose status as a root has become doubtful since rúar has become to be used for dúru 'outside part' presumably through the influence of Indonesian lúar 'outside part'.

4.6.4.1.3. Secondary Intransitive Verbs

Active Intransitive Verbs:

(1a) $\text{maŋ-} + \text{N}$: maŋabbíraŋ 'turn left' (from habbíraŋ 'left'); maŋÓrtOr 'to dance' (from tÓrtOr 'dance (N)'); maŋiámun 'turn right' (from siámun 'right (hand)').

(1b) $\text{maŋ-} + \text{A}$: maŋíbul 'become hard; become as of one piece' (from hibúl 'whole, of one piece'); maŋÉbbaŋ 'to spread; develop' (from hÉbbáŋ 'spread out'); mamÓttar 'become white; pale' (from bÓttár 'white').

(1c) maŋ- + reduplication + Vt: matcisúlu or matcusúlu 'basak (before a fire)' (from súlu 'light (Vt)').

(1d) maŋ- + tñ- + Vt: manut cáddE 'lean (against something)', from sáddE 'support (Vt)'; (cf. §4.6.4.1.2., (1a)).

(2) ma- + ni- + A: manisíó 'seek cover from rain', from sió 'sheltered from rain'; manihórtan 'sit with stretched legs', from hórtan 'stretched; elongated'.

(3a) mar- + N: this is the most productive derivation of intransitive verbs; the meanings are: (i) make use of what is expressed by the noun base, (ii) own or have what is expressed by the noun base.

(i) mardálan 'walk, go on foot', from dálan 'road, path'; marbáju 'be dressed', from báju 'clothing'; marsikkóla 'go to school', from sikkóla 'school'; martahúlu 'wear a hat', from tahúlu 'hat'; marsáhit 'be sick', from sáhit 'sickness'; marháta 'confer; hold talks', from háta 'word'.

(ii) marhÉpEŋ 'possess money', from hÉpEŋ 'money'; martáŋan 'have hands', from táŋan 'hand'; martúa 'be fortunate', from túa 'good fortune'; marjÉa 'be unfortunate', from jÉa 'misfortune'.

(3b) mar- + Vt: these have corresponding inflected transitive verbs, e.g. with maŋ-: marísap 'smoke (i.e. have habit of smoking)', from ísap 'smoke (something, e.g. a cigarette)', martídju 'fight; be in a fight', from tídjú 'punch', mallójoŋ 'run', from lójoŋ 'run after (somebody)'.

(3c) mar- + A: marhúnik 'grow (or, be) yellow', from huník 'yellow'; marpógos 'be poor', from pogós 'poor'; marúnE 'come to terms; be in agreement', from unÉ 'convenient; appropriate'.

(4) mar- + medial + base: martabúni 'hide; be in hiding', from buní 'hidden (A)'; marsiájar 'study', from ájar 'teach (Vt)'; marsiádu 'run in a race; be in a competition', from ádu 'chase; overtake'; marsitíjúr 'spit', from tíjúr 'saliva'; martalúkkup 'sleep with one's head covered', from lukkúp 'completely covered'; marbinótó 'know; be cognizant', from bótó 'know (something)'; marbinége 'have sense of hearing; hear', from bége 'hear (something); listen to (something)'; marhujíŋjaŋ 'prance; jump up and down', from jíŋjaŋ (or, more commonly, hajíŋjaŋ) 'prancing movements'; marhurája 'entreat', from rája 'king'; marhuási 'be forgiving; relent', from ási 'pity; grace'; marnalója 'tire', from lója 'be tired'; marnasáE 'stop', from sáE 'finished; over' (the last two examples are only found in the negative: e.g. dàŋ marnalója íbána marmÉam 'he does not get tired of playing'). In these examples the segments consisting of the medial and the base follow the pattern of nouns in this derivation, but, except for binótó 'knowledge' (usually in a derogatory sense), they do not occur as nouns in

that form. Verbs like *margalása* 'writhe' (mar- + <al> + √gasa) are not included here, because it is covered by (3a) above, *galása* 'writhing movement' formed in the pattern of §4.6.3.2., (10).

(5a) ma- + si- + N (the meaning is 'try to obtain what is expressed by the noun base'): *masihóda* 'search for, or try to buy, a horse', from *hóda* 'horse'; *masihÉpEn* 'try to earn, or borrow, money', from *hÉpEn* 'money'; *masibóru* 'be in search for a wife', from *bóru*, short for *tungànE bóru* 'wife; female spouse'.

(5b) ma- + tu- + N: *matubába* 'have shortage of food; suffer famine', from *bába* 'mouth'; *matumóná* 'eat first fruit of crop', from *bóna* 'beginning' with a change of consonant.

(6) mar- + si- + A + -i; used with a plural subject, the meaning is 'be unequal in what is expressed by the adjective base': *marsitibbói* 'be of unequal height', from *tibbó* 'high'; *marsigadjáni* 'be of unequal length', from *gadján* 'long'; *marsipogósi* 'be of unequal poverty', from *pogós* 'poor'. The corresponding noun has the form par- + si- + A + -i; it means 'the difference in the quality expressed by the adjective'.

Stative Intransitive Verbs:

(7a) m(a)- + N: *mabúgan* 'suffer wounds; be wounded', from *búgan* 'wound'; *matóras* 'ripe; mature', from *tóras* 'ripeness; maturity'; *masíh0l* 'yearn; be homesick', from *síh0l* 'homesickness; longing'; *mársak* 'be sad' from *ársak* 'sadness'; *mÉkkeI* 'to laugh', from *ÉkkeI* 'a laugh'.

(7b) ma- + Vt: *madábu* 'fall', from *dábu* 'drop; let fall'; *mabá0r* 'float downstream; be washed away', from *bá0r* 'wash away'; *maríbak* 'be torn', from *ríbak* 'tear (something)'.

(8a) pa- + V; used with a plural subject: *pasítik* 'be at odds', from *sítik* 'be in an irritable mood'; *pagátti* 'be exchanged by mistake', from *gátti* 'change'; *pajúppan* 'meet with (somebody)', from *júppan* '(be) found, recovered'.

(8b) pa- + tu- + V; (expresses a high degree of something undesirable): *paturápar* 'suffer extreme famine; starve', from *rápar* 'go without food; suffer hunger'; *patunó0n* 'stand completely idle', from *00n* 'do nothing'; *patusá0k* 'be in utter confusion', from *sá0k* 'fry without fat; turn over and over as in such frying'.

(9) UM- + base: *humútur* 'be shaking', from *hútur* 'shake (something)'; *gumórsin* '(has) become brown', from *g0rsin* 'brown'; *humúrtik* 'make a slight movement', from *húrtik* 'move slightly; disturb, touch'; *humaláput* 'be in a hurry', from *haláput* 'haste (N)'.

(10) N + -0n (rarely, -an); the meaning is 'suffer from what is expressed by the noun base': *batúh0n* 'have a cold', from *bátuk* 'a cold'; *ruán0n* 'have holes', from *rúan* 'hole; cavity'; *nalían* 'feel cold', from

ḡáli 'coldness'; halimatá0n 'have eye disease', from halímáta 'kind of eye disease characterised by redness'.

(11) ha- + N + -an; it expresses 'being overtaken by something undesirable, expressed by the noun base': haudánan 'be caught in rain', from údan 'rain'; hab0rḡínan 'be overtaken by night', from b0rḡin 'night'; hasiánan 'overtaken by daylight; get up late'; hapanakk0án 'suffer burglary or theft', from panákk0 'thief'; habubúhan, the only one of this kind (ha- + reduplication + base + -an), 'be no longer a virgin', seems to belong in this group, from búha 'open (Vt)' with a coalescence of vowel.

(12a) tar- + Vt (or, N); it means 'accidentally suffer from what is expressed by the base': tardége 'be trod upon accidentally', from dége 'tread on (something)'; tart0bb0m 'be rammed; collide against something', from t0bb0m 'ram; strike forcefully with body'; tarári 'be victim of a bad day', from ári 'day'; tarp0s0 'be born too young; be in a state in which there are no elder people (said of a family group)', from p0s0 'youngness' (or p0s0 'young'); tarsúga 'be pricked by a thorn', from súga 'thorn'.

(12b) tar- + su- + A: tarsuláddit 'slip; skid', from laddít 'slippery'.

(12c) In tarsíḡ0t 'remember; come to mind', from íḡ0t 'remember (something)', there is an intrusive /s/, if it belongs to (12a) above, or there is a loss of vowel /u/, if it belongs to (12b). The form taríḡ0t 'can be remembered' is the potential mode of the transitive verb íḡ0t 'remember'; it also occurs in the phrasal preposition taríḡ0t tu 'concerning'.

(13) There are two intransitive verbs beginning with /m/ matched by forms beginning in /p/: máḡan 'eat', páḡan 'eat (something)'; módom 'sleep', papódom 'put to sleep', podómi 'sleep upon (something)'.

(14) Doubling: pEakpEák 'lie idle', from pEák 'lie'; huddulhúddul 'sit idle', from húddul 'sit'; modommódom 'lie (without really going to sleep); lie down for some rest', from módom 'sleep'; taḡistanḡis 'cry piteously', from táḡis 'cry'; mEkkElÉkkEl 'laugh for a long time', from ÉkkEl 'laugh (N)' (cf. mÉkkEl 'laugh (V1)'); mittopíttop 'flicker', from míttop 'go out' (with root √ittop); joujóú 'make a sound like one continuously calling', from jóú 'call'.

4.6.4.2. Transitive Verbs

Transitive verbs are simple stems, primary derivatives, or secondary derivatives.

4.6.4.2.1. Simple Transitive Verbs

pásak 'beat'; sÉat 'cut'; d0k 'say'; búat 'take'; jóu 'call'; álap 'fetch'.

4.6.4.2.2. Primary Transitive Verbs

There are few primary derivatives among the transitive verbs.

(1) par- + root: parmEam 'play with (something)', from √-mEam, which is also found in mEammEam 'toy' and marmEam 'play (Vi)'; parsitta 'wish for; desire', from √-sitta, also found in sittasítta 'wish (N)'.

(2) pa- + root: pabóa 'tell', from √-boa, also found in boabóa 'proclamation; announcement'; pasídiñ 'avoid', from √-sidiñ, also found in marsidiksídiñ 'keep oneself away from any contact'; paúll0p 'make to emerge', from √-ullop, also found in múll0p 'emerge'.

(3) root + -h0n: tarúh0n 'take (something somewhere)', from √taru also found in patáru 'take (i.e. accompany a person)'; luáh0n 'carry away', also found in mañalúa 'elope'; agóh0n 'lose (something)', from √ag0, also found in mág0 'be lost (Vi)'.

(4) root + -i: ittópi 'extinguish', from √ittop, also found in míttop 'go out; extinguished'; podómi 'sleep on (something)', from √podom, also found in papódom 'put to sleep'; dakdátí 'continue (doing)', from √dakdat-.

(5) pañ- + root+ -i (cf. intransitive mañ- + root): pakkulíñi 'to address (a person)', from √huliñ; pañihúti 'to follow (a person)', from √ihut (also found in ihútt0n 'follow').

4.6.4.2.3. Secondary Transitive Verbs

(1a) par- + N; the meaning is 'make use of, or do something in connection with, what is expressed by the noun base': pardóñan 'befriend; accompany', from dóñan 'friend'; parábit 'wear; dress oneself with', from ábit 'clothing'; pargúru 'ask for (or, follow) the advice of', from gúru 'teacher'.

(1b) par- + N + -i; the I.C. structure is (par- + N) + -i, (cf. (9) below): parr0háí 'make wise' give wisdom to', from parr0há 'wise man', róha 'mind'; parbadiái 'make holy', from parbadía 'holy being', badía 'holiness'; parabíti 'dress (somebody)', from parábit 'one who has clothes', ábit 'clothing (N)'.

(1c) par- + N + -h0n; the constituent structure seems to be (par- + N) + -h0n, (cf. (10) below): parr0háh0n 'pay attention to', róha 'mind'; pardebatáh0n 'worship (or, consider) as a god', debatá 'god'; parbadáh0n 'quarrel about', from báda 'quarrel'.

(1d) par- + ha- + A: parhatútu 'confess', from tutú 'true'; parhamá0l 'cherish; value', ma0l 'dear; difficult'; parhaséan 'find a use for; not throw away', sEán '(it is a) pity'.

(1e) par- + numeral; the meaning is 'divide into such and such a number of parts': pardúa 'divide in two', dúa 'two'; partólu 'divide into three parts', tólu 'three'.

(2) par- + si- + Vt: parsiajári 'study; learn', from ajári 'teach'; parsigáttuñ 'suspend oneself from', from gáttuñ 'hang'.

(3a) pa- + A, Vi: the meaning is 'make or cause what is expressed by the base: pagádjan 'lengthen; make longer', gadján 'long'; pabálga 'make bigger', balgá 'big'; patíbbo 'raise', tibbó 'high'; papÉak 'lay (down)', pEák 'lie'; pahúddul 'set (down); make to sit', húddul 'sit'; pajÓñj0ñ 'make to stand (up)', jÓñj0ñ 'stand'; painum 'give to drink', mínum 'drink (Vi)', ínum 'drink (Vt)'. The simple mode in the active voice is not formed by prefixing mañ-, but by suffixing -h0n; thus, paj0ñj0kk0n jábu d0 ibána 'he is building (erecting) a house'.

(3b) pa- + N: pabára 'put in stable', from bára 'stable'; pasáruñ 'sheathe (a knife)' from sáruñ 'sheath'; padúru 'excommunicate', dúru 'the outside'.

(4) pa- + N + -hon: patukkákk0n 'have (something) fixed by a carpenter or mechanic', from túkkañ 'carpenter; mechanic'; padatúh0n 'have (something) cured by a magician', from dátu 'sorcerer; medicine man'.

(5) pa- + Vt-doubling; the meaning is 'do repeatedly what is expressed by the base in a playful or irresponsible manner': pad0kdók 'chatter about (something)', d0k 'say'; pain0tíñ0t 'keep remembering', íñ0t 'remember (something)'; pahiddattíddat 'lift up playfully and repeatedly', híddat 'lift'.

(6a) N + -i: t0rt0ri 'dance before (or, around) something', t0rt0r 'dance'; ul0si 'cover with a blanket'. úl0s 'blanket'; pudí 'stand with (or, turn) one's back to', púdi 'back (part)'; ruáñi 'make a hole in', rúañ 'hole'.

(6b) Vi + -i: huddúli 'sit on (something)', húddul 'sit'; ñóti 'wake up (a person)', ñ0t 'be awake'; rapári 'make hungry; starve', rápar 'go without food'.

(7a) N + -h0n; the meaning is 'do what is appropriate with (or, use) the object expressed by the noun base': abítt0n 'wrap around (like a dress)', ábit 'clothing'; t0rt0rh0n 'use, or give, in dancing (e.g. a ceremonious blanket)', t0rt0r 'dance'; gurúh0n 'learn; study', gúru 'teacher'.

(7b) Vi + -h0n: pEákk0n 'put down', from pEák 'lie'; gattúkk0n 'hang up (something)', from gattuñ 'suspended; be hung'.

(8a) ha- + N + -i: hasihóli 'long for', síh0l 'longing'; habiári 'fear; feel fear for', bíar 'fear'; hapósi 'trust', pòs (ni róha) 'trust (of heart)'.

(8b) ha- + V1 + -i: hap0rsEái 'believe in (something)', from p0rsÉa 'believe; have faith (V1)'; hasiáti 'be sympathetic towards (a person)', from síat 'fit (into something)'.

(9a) ha- + V1 + -h0n: hamatÉh0n 'die of; keep doing something till one dies of it', mátE 'die'; haróh0n 'come for', r0 'come'; haj0ñjók0n 'bear witness of', jók0ñ 'stand'.

(9b) ha- + N + -h0n: haluñútt0n 'yearn for', lúñun 'longing; solitude'; haburjúh0n 'persevere in', búrju 'seriousness; goodness'; hadengátt0n 'accept (as good)', déngan 'goodness'.

4.6.5. Stems of Other Word Classes

There is no extensive derivation in the other word classes. The few derived stems found have been mentioned passim (§4.3.2.-4.5.).

4.7. ROOTS

In §4.6. above, we make mention of roots: √tariñ in tatáriñ 'tripod potstand'; √badda in parbaddáan 'graveyard'; √gak in sigák 'raven'. By and large, roots consist of two syllables, as do the majority of simple stems.

4.7.1. Pairs of words like hábañ 'to fly' : háb0ñ 'wing'; mós0k 'burnt (i.e. in cooking)' : másak 'done (in cooking)', which show a similarity in form and meaning, raise the question whether one may not be derived from the other. There are other such pairs, but too few for any conclusions: sákk0p 'lid of earthen pot' : lákk0p 'a cover'; dúpañ 'Y-shaped branch' : síppañ 'intersection'; h0pít 'narrow' : hápit 'squeeze (Vt)'; tádda 'know (recognise)' : taddí 'different; stand out'; múlak 'return (V1)' : túlak 'reject'; b0rát 'heavy' : s0rat 'suffer under a heavy load'; mátE 'die' : páte 'unredeemable (of sale)'; álañ 'not quite reach the end' : hálañ 'bar (Vt)' : h0láñ 'distant'. Descriptively, we consider such words as not related.

4.7.2. Many of the disyllabic simple stems and roots seem clearly to be the result of reduplication: dokdók 'heavy'; táttap '(√tap- √tap) 'wash'; sátcap '(√sap- √sap) 'slice clean'; láklak 'bark; skin (of fruit)'; dándañ 'pay a heavy sum'; √s0s0 in s0s0i 'urge'; √dakdak '(√dat- √dat) as in dakdáti 'continue'; √bakkak as in mabákbak 'stream down (e.g. tears)'. Some of these manifest peculiar alternation

patterns, which are not covered by §3. In our discussions on derivation, we treat these reduplication forms as simple stems or roots without mentioning their inner structure.

Other types of such reduplication are: (1) reduplication of onset and peak; e.g. *tátap* 'peer (at something distant)'; *tútun* 'burn (Vt)'; *√1010ŋ* as in *mahil010ŋ* 'desert husband'; *súsun* 'arrange in a row or pile'; and (2) with a different vowel; e.g. *píp0t* 'guilt; sin'; *dódak* 'inner skin of rice, shed through pounding'; *sósa* 'erase'; *túltal* 'quarrel'; *súksaŋ* (*√suŋ- √saŋ*) 'curly feathers (of a chicken)'; *súrsar* 'in disarray; crisscross'; *gúga* 'stir'; *dúda* 'pound (rice)'.

CHAPTER V

SYNTAX

5.0. INTRODUCTION

In sections §5.1.1.-2., construction markers, pure and impure, will be placed between brackets for clarity. Intonation will not be transcribed; when intonation is relevant, the phrase with the intonation is given between slant lines, but we retain (non-phonemic) spaces as word boundaries. When useful, the immediate constituent cut will be indicated with a vertical stroke.

5.1. CONSTRUCTIONS

Phrases are exocentric and endocentric.

Exocentric phrases are predicative or directive. For example, húddul [d0] ibána 'he is sitting' is predicative: maṇàllaṇ gád0ṇ 'eat yam' is directive.

Endocentric phrases are attributive or coordinate. For example, matcài ulí 'very pretty' is attributive; tà0 tóba 'Lake Toba' is co-ordinate.

Many phrases contain a marker. For example, d0 in the predicative phrase lá0 [d0] ibána 'he went'; manəṇ 'or' in the co-ordinate phrase lóbbu [manəṇ] hódə 'cow or horse'.

5.1.1. Endocentric Constructions

5.1.1.1. Attributive Constructions

Attributive constructions fall into several constructions according to the head. This classification intersects the classification according to the relative position of head and attribute.

An attribute phrase will be called after the head; thus, jàbu [na] bálgə 'big house' is a noun phrase; mardàlan pát 'walk on foot' is a verb phrase.

5.1.1.1.1. *Noun Phrase*

The classification below is according to the attribute; the sixth and eighth are distinguished by the head and the relative order of head and attribute. In the first seven classes the relative order is fixed; in the eighth, the order is not fixed.

Head first:

(1) N + N (first noun is head): hòrb0 padáti 'draft water-buffalo' (hórb0 'water-buffalo', padáti 'cart'); hùd0n tán0 'earthen pot' (húd0n 'pot', tán0 'earth'); gùru kapála 'principal' (gúru 'teacher', kapála 'head; leader'); tàrup [ni] jábu 'roof of a house' (tárup 'roof', jábu 'house'); sùhul [ni] ráut 'sheath of a knife' (súhul 'sheath', ráut 'knife').

(2) N + V: kàpal hábaŋ 'airplane' (kápal 'boat', hábaŋ 'to fly'); hòda pátcu 'race horse' (hóda 'horse', pátcu 'overtake; race'); hàlak [na] piníó 'an invited person' (hálak 'person', piníó 'have been invited'); dakdànak [na] módom 'a sleeping child' (dakdànak 'child', módom 'sleep').

(3) N + A: jábu [na] bálga 'a big house' (jabu 'house', balga 'big'); hàlak [na] gógo 'a strong person' (gogo 'strong').

(4) N + Pronoun: jábu í 'that house' (i 'that'); mèja díá 'which table' (díá 'which'); bùkku [na] tòlu (ón) 'all (these) three books' (tòlu 'three'); bàtu [na] dúá 'two stones (a place name)' (dúá 'two'); dèsa [na] uálu 'the eight directions' (dèsa 'compass point', uálu 'eight').

(5) N + Prepositional Phrase: jábu [na] di hùta (í) '(the) house in the village' (di 'at', hùta 'village'); sùrat [na] sìan ibàna (í) '(the) letter from him' (sùrat 'letter', sìan 'from', ibàna 'he; she').

Attribute first:

(6) Pronoun + Noun of Measure: tòlu mákkuk 'three capfuls' (tòlu 'three', mákkuk 'cup'); pìga híraŋ 'how many basketfuls' (pìga 'how many', híraŋ 'basket'); dùa bóŋŋin 'two nights (long)' (dúá 'two', bóŋŋin 'night').

(7) Auxiliary + N: àkka jábu 'houses' (ákka is plural marker); gànup hálak 'every person' (gànup 'every; each'); nàsa jólma 'all men' (nása 'all', jólma 'human being').

Order not fixed:

(8) Noun and Quantifier (i.e. numeral or a phrase of type (6) above): hóda tòlu or tòlu hóda 'three horses'; dùa mákkuk | kópi or kópi | dùa mákkuk 'two cups of coffee'. There seems to be no hard and fast rule determining the order. However, in enumerations, the more common order is numeral + noun; e.g. dúa lóbbu, tòlu hóda, ... 'two cows, three horses, ...'. The head in this construction is a noun

other than a measure noun. The stress pattern in this type of numeral + noun construction is primary-primary; in type (6), it is secondary-primary: tòlu mákkuk 'three cups', tòlu mákkuk 'three cupfuls'.

It is possible to have more than one attribute to one noun head. For example: hùd0n tàn0 [na] bàlga í (pot, earth, big, that) 'the big earthen pot'; sudé hàlak hùta [na] mòra í (all, person, village, rich, that) 'all the rich villagers'. The nesting of the first phrase is

(((((hùd0n) tàn0) na bàlga) í)
1234 4 3 2 1

and that of the second phrase

((sudé (((hàlak) hùta) na mòra) í)).
1 2345 5 4 3 21

Postposed modifiers fall into six position classes: (a) unmarked verb, (b) unmarked noun, (c) ni-attribute, (d) na-attribute, (e) pronoun (demonstrative or interrogative), (f) quantifier as found in type (8).

For example, head + (a) + (b) + (c) + (d) + (e) + (f):

((((((((sèdd0k) mànan) pÈrak) ni si pìtt0r) na ibbàru) í) dúá)
1234567 7 6 5 4 3 2 1

spoon eat silver Pittor (name) new that two
'two of Pittor's new silver tablepoons'.

It is possible to have more than one na-attribute to one head. For example, dakdànak na mòkm0k na mòdom í (child, fat, sleep, that) 'the fat child who is sleeping'.

We also find two or more ni-phrases in succession; e.g. pàt [ni] sara0al [ni] si pìtt0r 'the leg of Pittor's trousers'. It is sometimes difficult to tell to which noun the second (third, etc.) ni-phrase is attributive; e.g. làppak [ni] bùkku [ni] si pìtt0r (cover, book, Pittor) can mean 'Pittor's book cover' or 'the cover of Pittor's book'. However, in most cases, a ni-phrase is attributive to the noun in the preceding ni-phrase; e.g. in àrga [ni] tàrup [ni] jàbu [ni] si pìttor (price, roof, house, Pittor) 'the price of the roof of Pittor's house'.

5.1.1.1.2. Verb Phrases

Of the classes of verb phrases below, the order is fixed in (1), (2), (4), and (7). The constructions in (1) and (2) have a highly limited productivity; phrases built by them can be regarded as idiosyncratic. In the other classes, the order may be reversed depending on the structure of the sentence in which they occur; the order given in the examples is that found in direct order (see §5.2.) when the verb and its attribute are contiguous.

Head first:

(1) V + N: mardàlan pát 'walk on foot' (walk, foot); marbàda bólak 'have a big quarrel; feud' (quarrel, vastness); mallàŋE púdi 'swim on one's back' (swim, back).

(2) V + Pronoun: marròha áu 'be selfish' (have mind, I); marròha sáda 'be stubborn' (have mind, one); mallàkka dúá 'to gallop' (to step, two).

(3) V + Adjunct (Pro-adjunct): lá0 marsógot 'go tomorrow'; pEák disán 'lie there (yonder)'; tarúh0n tusón 'bring here'; mallánE di súŋE 'swim in the river'.

(4) V + Auxiliary: là0 hián 'went previously'; là0 bé '(each) goes individually'; mòdom musÉ 'sleep again'.

Attribute first:

(5) A + V: (hàlak na) hatóp mardálan '(a man who) walks fast'; (dàn dengán) godán marísap '(it is not good) to smoke much'.

(6) Adjunct + V: saúr matúa 'reach old age' (successful, be old); sáut marsáhit 'got sick as expected' (as expected, be sick); t0rús módom 'sleep continuously; keep sleeping'.

(7) Auxiliary + V: nàEŋ lá0 (d0 ibána) '(he) wants to go'; pìtt0r ró (d0 ibána) '(he) came immediately'; sòr mátE (ma ibána) '(he) almost died'.

A verb head can have more than one attribute; see §5.2. below.

5.1.1.1.3. Adjective Phrases

The order in adjective phrases is fixed. The auxiliaries which modify an adjective thus fall into preposed and postposed classes.

Head first:

(1) A + Auxiliary: dengàn saháli 'very good'; dà0 situtú 'truly far'; àrga hían 'extremely expensive'.

(2) Auxiliary + A: matcài ulí 'very pretty'; tàr murá 'somewhat easy'; làm maól 'more and more difficult'.

In colloquial speech, we often find more than one auxiliary (rarely more than two) modifying an adjective; e.g. matcài ulí saháli 'very very pretty'; tùŋ matcài balgá 'very extremely big'.

5.1.1.1.4. Adjunct Phrases

The relative order in these phrases is fixed.

Head first:

(1) Adjunct + Numeral: natt0àri sadá 'the day before yesterday' (natt0ári 'yesterday'; sáda 'one'). This is the only one in this class.

Attribute first:

- (2) Auxiliary + Adjunct: marEak b0tari 'toward evening'.

5.1.1.1.5. Substitute Phrases

In the three classes of substitute phrases below, the order is fixed.

Head first:

- (1) Numeral + Auxiliary: dùa nári 'two more'; lìma nári 'five more'.

Attribute first:

- (2) Auxiliary + Pronoun: mànaŋ ísE 'whoever; anybody' (ísE 'who');
agìà díá 'whichever' (díá 'which').

- (3) Auxiliary + Numeral: nàng0 tólu 'at least three'; nàng0 sáda
'at least one'.

5.1.1.2. Co-ordinate Constructions

A co-ordinate construction is additive, alternative, or appositive.

5.1.1.2.1. Additive Constructions

An additive construction often involves a marker (connective).

- (1) N + N: jábu [d0h0t] haúma 'house and rice field'; más [d0h0t]
pináhan (d0 bolína) '(her bride price consists of) gold and live stock'.

- (2) N + Pronoun: (gárar ma) bùkku í d0h0t ón '(please pay for)
the book and this'.

- (3) Pronoun + Pronoun: áu [d0h0t] ibána 'I and he (= he and I)';
ón [d0h0t] án 'this and that (yonder)'.

- (4) A + A: hipás [jala] m0kmók (d0 ibána) '(he is) healthy and
robust'.

- (5) V + V: (di jabunámi d0 ibána) máŋan [d0h0t] módom '(he) eats
and sleeps (in our house)'; mÈkkEI táŋis (d0 dakdànak í) '(the child)
is laughing and crying'.

- (6) Prepositional Phrases: (lulúi ma) di gìdjaŋ [d0h0t] di t0ru
(ni mèja í) '(look) on and under (the table)'; tu dólok | tu torúan
'upward and downwards'.

- (7) Numerals: sappùlu pítu 'seventeen' (ten, seven); saràtus | dùa
púlu 'one hundred and twenty'; tólu [tàbba] ópat 'three plus four'.

5.1.1.2.2. Alternative Constructions

An alternative construction involves the marker mánaŋ 'or'.

- (1) Nouns: hóda [mànaŋ] lóbbu 'a horse or a cow'.

- (2) Noun and Pronoun: híta [manaŋ] musútta 'we or our enemies'.

- (3) Substitutes: hó [manaŋ] ibána 'you or he'; tólu [manaŋ] ópat
'three or four'; disóá [manaŋ] disán 'here or there (yonder)'.

- (4) Verbs: mardálan [manaŋ] marmót0r 'go on foot or go by car'.
 (5) Adjectives: argá [manaŋ] murá 'expensive or cheap'.
 (6) Prepositional Phrases: di gídjaŋ [manaŋ] di tòru (ni mèja í) 'on or under (the table)'.

5.1.1.2.3. Appositive Constructions

Examples: tà0 tóbá 'Lake Toba'; pùlo samÓsir 'Samosir Island'; àEk sigEá0n 'Sigeaon River'; sahála diapári 'two personal names'; sáhat sihÓbbin 'personal name and family name'; si pítt0r | ànak ni amadjáluk 'Pittor, the son of Amadjaluk'.

5.1.2. Exocentric Constructions

5.1.2.1. Directive Constructions

There are two types of directive constructions. In the first the director is a particle; we will call it *marked*. The second consists of a verb and its agent or object; we will call it *verbal*.

5.1.2.1.1. Marked Directive Constructions

These fall into two classes according to the form-class of the director: *prepositional* and *conjunctive*.

(1) Prepositional Constructions.

The axis in a prepositional construction is a non-clause, mostly a nominal (noun, noun phrase, or pronoun). Examples: di huta 'in the village'; (tarup) ni jabu '(roof) of a house'; sòŋ0n más 'like gold'; (hàu) na tíbbo 'a tall (tree)'.

(2) Conjunctive Constructions.

The axis in a conjunctive construction is a clause. Examples: (ró ma hó,) mòl0 | s0 adóŋ ulaómmu '(please come) when you have nothing to do'; (tibú ma hó ró,) àsa | ráp híta lá0 '(please come early,) so that we go together'; (únaŋ gádis bukkùna í,) sòtuŋ | dimurùhi ibána hó '(don't sell his book) lest thou be rebuked by him'.

5.1.2.1.2. Verbal Directive Constructions

There are two types of verbal directive constructions: *agentive* (consisting of a passive transitive verb and its agent) and *objective* (consisting of an active transitive verb and its object). In both the verb comes first.

(1) Agentive Constructions: hujòu hámi (d0 ibána) '(he) was called by us'; ditòŋ0s nasída (ma hÈpEŋ í) '(so the money) was sent by them'; tarhiddat ibána (d0 pùti na bōrat í) '(the heavy chest) can be lifted

by him'; garàr0n [ni] nat0rásna (d0 utàŋŋa í sudé) '(all his debts) will be paid by his parents'.

(2) Objective Constructions: maŋùla haúma (do ibána) '(he) is tilling the rice field'; manùan ÉmÉ (d0 inátta) '(our mother) is planting rice'; (ringás d0 ibána) manàpu jábu '(he) swept the house (diligently)'.

5.1.2.2. Predicative Constructions

A predicative construction consists of a **predicator** and a **subject**, which may be linked by a predicative particle. When the predicator consists of, or contains, a verb, we will call the construction **verbal predicative**: módom [d0] ibána 'he is sleeping' (sleep, he). Otherwise, we will call the construction **equational**: gúru [d0] ibána 'he is a teacher' (teacher, he). In the examples below, only the direct word order is represented; other types of word order are illustrated in §5.2.1.

5.1.2.2.1. Verbal Predicative Constructions

Verbal predicative constructions fall into four types according to the predicator.

(1) The predicator is an intransitive verb with or without attributes.

Without attribute: tánis [d0] ibána 'he is crying'; húddul [ma] nasída 'they sat down'; marídi [pE] áu '(I promise) I will take a bath'; (únaŋ pío ibána, mòl0) batúh0n ibána '(don't invite him if) he has a cold'; (láo ma hámi dùŋ) módom ibána '(we left after) he was asleep'.

With attribute: nàEŋ módom [d0] ibána 'he wants to sleep'; mardàlan pát [d0] hámi 'we (excl.) walked on foot'; mòdom bé [ma] nasída 'they all went to bed' (each sleeps, they).

(2) The predicator is an agentive constitute (without attribute): hujòu hámi [d0] ibána 'he was called by us (excl.)'; ditòŋ0s ibána [ma] sùrat í 'then the letter was sent by him'; tartùh0r si jóhan [d0] í 'it can be bought by John'.

In this type of predicative construction, the subject may also be a verb, a verb phrase, or an objective constitute: dimulài nasída [ma] marÉddÉ 'to sing was begun by them (= they began to sing)'; diulàhi ibána [ma] tánis 'to cry was repeated again by him (= he again started to cry)'; dit0rùss0n ibána [d0] manùrat suràkna í 'to write his letter was continued by him (= he continued writing his letter)'. This type of construction is found only with a small set of transitive verbs which express some temporal distribution of an event; e.g. mulái

'begin'; pasó 'cease'; t0rúss0n 'continue'; dakdátí 'do repeatedly'.

In these examples, the verb is a passive verb. We also find passive transitive verb predicators without a separate actor; e.g. hujóu [d0] nasída 'they were called by me'; pinío [d0] ibána 'he was invited'.

(3) The predicator is an objective constitute (without attribute): madjàha búkku [d0] ibána 'he was reading a book'; sumòls0l bági [d0] ibána 'he has regretted his fate'; masigàrar utánna [d0] nasída 'they each paid his debt'.

(4) The predicator is a reciprocal verb (without attribute): masisukkúnan [ma] nasída '(then) they asked each other'; (dàn hubót0 na) masih0s0man nasída '(I did not know that) they hated each other'.

5.1.2.2.2. Equational Constructions

There are four types of equational constructions according to their constituents.

(1) Both the predicator and the subject are nouns or noun phrases; the first constituent is the predicate: gúru [d0] túlan '(my) maternal uncle is a teacher'; ib0tókku [d0] si tío 'Tio (name) is my sister'; túkkan na málo [d0] si pítt0r 'Pittor (name) is a skilled carpenter'.

(2) The predicator is a noun or noun phrase, the subject a pronoun: sùpir mótt0r [d0] ibána 'he is a chauffeur'; háu [d0] í 'it is a tree'; parÉdde [d0] áu 'I am a singer'.

(3) The predicator is a numeral, the subject a noun or pronoun: tólu [d0] hodána 'his horses are three (= he has three horses)'; ílma [d0] ón 'this is five'; sappúlu [d0] nasída 'they are ten (= there are ten of them)'.

(4) The predicator is an adjective, a prepositional phrase, or a pro-adjunct of location, the subject a noun or a pronoun: tibbó [d0] hàu í 'the tree is tall'; jukkát [d0] ibána 'he is naughty'; di bílut [d0] nasída 'they are in the room'; disí [d0] amátta 'our father is there'.

The examples given in §5.1.2.2.1.-2. are the occurrent types when there is no complement in the clause (see below).

5.2. CLAUSES

A predicative constitute is a clause. A clause may also contain one or more complements. In marpÉsta d0 nasída natt0ári 'they had a feast yesterday' (have a feast, particle, they, yesterday), natt0ári is a complement.

Clauses are independent or dependent.

An independent clause contains a predicative particle, except when

it is interrogative containing an interrogative pronoun or proverb. For example: marpÉsta d0 nasída 'they had a feast'; mardálan pE áu 'I will walk'; ísE mabbùat bukkukki 'who took my book?'; áha dibàEn ibána 'what did he do?'; marhúa ibána 'what is he doing' (do what, he).

A dependent clause is introduced by a subordinating conjunction or a nominalising particle. For example: (hujóu pE hó) mò10 ró ibána '(I will call thee) when he comes'; (lá0 d0 ibána) àng0 s0 lEanómmu hEpèŋga í '(he is going) if his money is not given by thee'; (dàn hubót0) na maŋápil ibána '(I did not know) that he was studying'.

Other types of clauses are of the minor type (see §5.2.2. below).

Limiting ourselves to the predicator-subject variety of verbal predicative clauses, we have four types according to the verb:

(1) intransitive: marpÉsta d0 nasída natt0ári 'they had a feast yesterday'; batúh0n d0 ibána 'he has a cold'; cf. §5.1.2.2.1. (1).

(2) active transitive: madjàha búkku d0 ibána di parpustakáan 'he was reading a book in the library' (read book, particle, he, in the library); maŋuláhi ma ibána marmÉam 'he started playing again' (start again, particle, he, play); cf. §5.1.2.2.1. (3).

(3) passive transitive: digàrar ibána d0 utàŋga í tu nasída 'his debt was paid by him to them' (be paid by him, his debt, to them); ísE dipì0 ibána 'who was invited by him'.

(4) reciprocal: masileléan d0 dakdànak í di póllak 'the children were chasing each other in the garden'; masibErÉŋan ma nasída '(then) they looked at each other'.

5.2.1. Word Order

The examples in §5.1.2.1.-2. and §5.2. above are in direct order. A clause which is not in direct order is in inverted order.

5.2.1.1. Direct Order

The most common pattern of a clause in direct order is predicator (P) + modal particle (Mp) + Subject (S) + complement (C). For example: módom d0 ibána di bílut 'sleep (Mp) he in-room' = 'he is sleeping in the room'; hatóp d0 ibána mardálan 'fast (Mp) he walk' = 'he walks fast'; marhúa ibána disí 'do-what he there' = 'what is he doing there?'.

When the subject contains an interrogative pronoun, the word order is S + Mp + P + C. For example: ísE manarùh0n ón 'who brought this?'; ànak ni ísE ón 'whose son is this?'; áha ditùh0r hó di tòkk0 í 'what was bought by thee in the shop?'.

When the clause contains an adverb, the adverb is the predicator; e.g. t0któn d0 ibána ringás 'he is always diligent' (always Mp he

diligent); *sáut d0 ibána marsikkóla 'he went to school as planned' (as planned, Mp, he, go to school).*

When the clause contains an adjective as a constituent (and no adverb), the adjective is the predicator; e.g. *hatóp d0 ibána mardálan 'he walks fast' (fast, Mp, he, walk)*; *ringás d0 hálak í marsiájar 'that person studies diligently' (diligent, Mp, the person, study).*

When the predicator is an adjective or an adverb and the complement is an agentive constitute, the latter comes before the subject, i.e. *P + Mp + C + S*; e.g. *hatóp d0 digàrar ibána utàḡḡa í 'quick, Mp, be-paid-by he, his-debt that' = 'his debt was paid by him quickly'.*

The negative and completive modal particles come before the predicator in the direct order; e.g. *dàḡ hatóp ibána mardálan 'he does not walk fast' (not, fast, he, walk)*; *nùḡḡa lá0 ibána tu sikkóla 'he has gone to school' (already, go, he, to school)*; *dàḡ karáni ibána 'he is not a clerk' (not, clerk, he).*

5.2.1.2. Inverted Order

In a clause in inverted order, some element other than those in direct order occupies the initial position. When this happens, that element is emphasised. For example:

Direct Order: *marpÉsta d0 nasída natt0ári 'they had a feast yesterday'.*

Inverted Order:

(1) *nasída d0 marpÉsta natt0ári 'THEY had a feast yesterday';*

(2) *natt0ári d0 nasída marpÉsta 'YESTERDAY they had a feast';*

In (3) *natt0ári d0 marpÉsta nasída*, the verb *marpÉsta 'have a feast'* also gets some emphasis, but to a much lesser degree. Likewise, *natt0ári 'yesterday'* in (4) *marpÉsta d0 natt0ári nasída*.

Inversions like (1) and (2) are more common; the emphatic effect is definite. We will call these normal emphatic order; (3) and (4) secondary emphatic order.

Another set of normal emphatic order is as follows:

Direct Order: *hatóp d0 ibána mardálan tu húta 'he walked home fast'.*

Normal Emphatic Order:

(1) *ibána d0 hatóp mardálan tu húta 'HE walked home fast';*

(2) *mardálan d0 ibána hatóp tu húta 'he WALKED home fast';*

(3) *tu húta d0 ibána hatóp mardálan 'he walked HOME fast'.*

A special inverted order of the above clause is sometimes found: *ibána hatóp d0 mardálan tu húta* which will normally be rendered with a sustained T.C. between *ibána 'he'* and *hatóp 'fast'*; it seems to be best regarded as two minor (elliptical) clauses (*àḡḡ0*) *ibána, hatóp d0 (ibána) mardálan tu húta '(as for) him, (he) walked home fast'.*

It is hard to tell whether there is a difference in degree of emphasis from (1) above; what is clear is that this is a special style of speech.

In a noun-noun equational clause (see §5.1.2.2.2. (1)), the subject in direct order is a noun that is more 'definite'. The scale of 'definiteness' from high to low seems to be: noun with a demonstrative attribute, proper noun, noun of relationship, a more specific noun, a more generic noun. In a normal emphatic order, the roles are reversed, the more definite noun is the predicator. Examples:

Direct order:

- (1) si pítt0r d0 hálak í *'that person is Pittor'*;
- (2) ib0t0kku d0 si tío *'Tio is my sister'*;
- (3) gúru d0 túlaŋ *'(my maternal) uncle is a teacher'*;
- (4) hálak na sábar d0 paráwat *'a nurse is a patient person'*.

Normal Emphatic Order:

- (1) hálak í d0 si pítt0r *'Pittor is THAT PERSON'*;
- (2) si tío d0 ib0t0kku *'my sister is TIO'*;
- (3) túlaŋ d0 gúru *'UNCLE is a teacher'*;
- (4) paráwat d0 hálak na sábar *'a NURSE is a patient person'*.

These clauses could be translated in this way: *'it is that person who is Pittor'*; *'it is Tio who is my sister'*, etc.

5.2.2. Clause Types

Clauses are of the favourite or minor type.

5.2.2.1. Favourite Types

The favourite clause-type centres on the predicative constitute with or without complements. It falls into three classes: declarative, interrogative, and intransitive imperative.

Examples:

Declarative: marsikk0la d0 ibána gànup ári *'he goes to school every day'*; maŋal0ppa iddáhan d0 inátta di dápúr *'our (excl.) mother is cooking rice in the kitchen'*.

Interrogative: ísE mamEàkk0n bùkku 0n tus0n *'who put this book here?'*; pistár d0 ibána marÉt0ŋ? *'does he do arithmetic cleverly? (= is he good at arithmetic?)'*. From the last example we see that a yes-or-no question has the same word order as a declarative; the difference lies in intonation:

2 3 2 2 1+
/pistár d0 ibána marÉt0ŋ / *'he is good at arithmetic'*

2 3 2 3, 3+
/pistár d0 ibána marÉt0ŋ / *'is he good at arithmetic?'*

Intransitive Imperative: ró hò 'come (thou)'; húddul ma hamú disón 'please (you) be seated here'. When uttered in great emotion (impatience), an intransitive imperative may be of the minor type, i.e. with no subject: húddul 'SIT DOWN!'.

5.2.2.2. Minor Types

Minor clause-types fall into several classes: transitive imperative, vocative, aphoristic, and fragments.

A transitive imperative clause consists of an objective constitute with or without a modal particle, or of a verb only with or without a modal particle: búat ì 'take it'; sápu jòlò ón 'will (you) sweep this'; pasákkòñ ulùna í 'do beat his head'; tákkup 'catch (it)'; páñan 'eat (it)' or páñan ma 'please eat (it)'. In very polite speech, a transitive imperative clause may be of the favourite type; in this case, hamú (second person plural) is used to refer to the addressee: tòñòs hamú ma hèpèñ í 'would you send the money'.

Vocative type: inán 'mother!'; laé 'cousin!' (also used commonly between males to a stranger of about the same age as the speaker); linE 'girl's name'; alé àkka dónán 'friends!'. The strongest stress always falls on the last syllable in a vocative.

Clauses of the aphoristic type are numerous. It is common practice among adult speakers to intersperse their speech with aphorisms, which usually come in balanced pairs. In this type of sentence, subordinating conjunctions and modal particles are often omitted; when these are supplied, the clause are of the favourite type. In the examples below, we will first give a literal translation (additions in the translation are put in parentheses), and then the meaning in normal English: máta gúru, róha siséan 'eyes (are) teachers, (and) the mind (is) the pupil' = 'one should take a lesson from what one sees or experiences'; ñalí áEk diiñàni dÉkkE, dolgí dákka diiñàni bódat 'water (is) cold, (yet it) is inhabited by fish; a branch (is) hard, (yet it) is inhabited by monkeys' = 'home sweet home'; hábañ sò marháboñ, sÓngOp sò marÉátan 'flying without wings, perching without a perch' (said of a person who brags without any foundation).

Often, only the first part is uttered; e.g. madábu járum tu na póppOt 'a needle falls in dense (grass)', omitting dàñ diìda máta, diìda róha 'though (it) is not seen by the eyes, (it) is seen by the mind' (said when it is obvious to a thinking person what the real situation is).

Fragment clauses fall into two types: completive and exclamatory. Examples:

Completives: tu Ónan 'to the market' (e.g. in answer to the question

tu dia h0 'to where thou (= where are you going)'; marmÉam 'playing' (e.g. in answer to 'what are you doing there?'); dá0η 'no!'; dan áha 'not what (= nothing)' (e.g. in answer to 'what are you doing there?'); módom d0 'sleeping' (e.g. in answer to 'is he sleeping?'; 'yes, he is'); m0l0 ád0η 'if exist (= if there is something)'.

Exclamatory clauses: hÓras 'greeting'; máttEh 'that serves you right!'; áyamaη 'astonishment or disappointment'; b0háí 'really?' (surprise or disbelief); tÓlE 'come on!'.

5.3. SENTENCES

Sentences are simple, complex, or compound. A simple sentence consists of a single clause; a complex clause consists of one clause with a word, a phrase, or a clause attributive to it; a compound sentence consists of two or more clauses in co-ordinate construction, with or without attributes.

Examples:

Simple Sentences: mársak d0 ibána 'he is sad'; ringás d0 nasída marsiájar gànup ári 'they study diligently every day'; masiurúpan d0 àkka dakdanàkna í di ulá0n d0h0t parhEpÉη0n 'his children help each other in word and finances'.

Complex Sentences: túh0r ma bùkku í, m0l0 sáE hEpÉηmu 'please, buy the book, if thy money is sufficient (i.e. if thou hast enough money)'; ása, dipì0 ibána ma sudé d0ηánna í tu pEstàna í 'so he invited all his friends to his feast'; tibú pE áu rÓ, m0l0 bóí 'I will come early, if possible'.

Compound Sentences: dipì0 nasída d0 ibána, alàì dan rÓ ibána 'he was invited by them, but he did not come'; parjùji lánis d0 ibána, jàla parjùji táddan 'he was an inveterate gambler, and (he was) an itinerant gambler (i.e. a gambler who goes from one gambling place to another)'; másuk ma ibána tu bílut, là0s módom ma ibána 'he went into the room, and he slept (i.e. and he went to bed)'; jádi, lá0 ma ibána tu dápur, jàla dipabboláh0n ma pìngan na disiní 'and so he went to the kitchen, and broke the dishes that were there'.

APPENDIX 1
SYNTACTIC ANALYSIS OF A TEXT

1. INTRODUCTION

The text is given in the usual transcription. Commas, periods, and question marks represent terminal contours. A word-for-word translation is given interlinearly. A free translation is given in §3; the analysis in §4.

The text is taken from T.M. Sihombing *Pabidangkon dohot Pabagashon*, Pematang Siantar (Sumatra), 1959(?), p.47. It is a fragment (the beginning) of a story which narrates how Jonaha, a legendary figure who could always escape from predicaments through his resourcefulness, freed himself from his debts to a certain Amaddorap (Father of Punching), so called for his phenomenal strength.

2. TEXT

- S1 ùltop ni si jOnáha.
 blow gun of Jonaha (name)
- S2 parjùji lánis d0 si jOnáha, jàla parjùji táddan, alài perjùji
 gambler addict Jonaha and gambler wandering but gambler
 tálu.
 lose
- S3 ása, dən na hÓlan ibána marútan.
 therefore not intermittent he have debt
- S4 s0ŋ0n í ma adÓŋ saháli parutanánna, na marg0ar
 like that exist once his creditor have name
 amaddÓrap.
 Father of Punching
- S5 nùŋŋa piga háli ditùŋgu amadd0rap í
 already how many time be demanded by Father of Punching that
 siŋirna í, alài hòlan bagabága d0 d0h0t jádji dibàEn
 his credit that but only promise and pledge be made by
 si jOnáha.
 Jonaha

- S6 di na sadári, ditòp0t amaddórap ma musÉ si jOnáha
 at one day be visited by Father of Punching again Jonaha
 manùngu siñirna í
 demand his credit that
- S7 àng0 sahàli ón, dàñ adóñ bE idáOn hasabárOn di
 as for once this not exist any more will be seen patience at
 di bohína.
 at his face
- S8 makkúliñ ma ibána, didók ma tu si jOnáha, alé laÉ.
 speak he be said by him to Jonaha I say cousin
- S9 nùñña malelekkú s0 digàrar hó utañmí.
 already too long not be paid by thou that they debt
- S10 jàla nùñña lója áu pabegebège bagabàga s0 máhap
 and already be tired I hear repeatedly promise not stopping
 sìan hó.
 from thou
- S11 óppe, gárar utáñmu sa0nnári.
 therefore pay thy debt now
- S12 m0l0 s0 í, diaómmu ma pohulpohulhón.
 if not that will be tasted by thee this my fist
- S13 dibòto hó d0 alána, ubbàEn marìbban gòar
 be known by thou its reason (has made) have rival name
 amaddórap áu.
 Amaddorap I
- S14 dáda nañ maramaniáha áu, ài s0 marhas0hótan d0pE áu.
 not already have teknonymy I for not be married still I
- S15 alái àla ni g0st0ñ ni b0t0hókku d0.
 but because of strength of my arm
- S16 àng0 póla d0 hudórap hálak, dàñ na bóí bE
 if just be punched by me person not capable any more
 hirìmOn ni í bErEñónna mataniári.
 will be hoped by that will be seen by him sun

3. FREE TRANSLATION

Jonaha's Blow Gun

Jonaha was an inveterate gambler, who went gambling from one place to another, but who always lost. Therefore he was continuously in debt.

So it was that he once had a creditor whose name was Amaddorap (Father of Punching). Amaddorap had already demanded his credit several times, but Jonaha had only made promises and pledges. One day Amaddorap again went to Jonaha to demand his credit. This time there was no longer any patience on his face. He spoke and said to Jonaha: "I say, cousin. It's already too long that you have not paid your debts. And I am already tired of hearing promises and pledges from you again and again. Therefore, pay your debt now. If not, you are going to feel my fist. You know the reason why I am called Father of

Punching. It's not because I have a child already, for I am not married yet. But it's because of the strength of my arm. If I do punch a person, it will no longer be possible for him to hope to see the sun."

Continuation of the story: *Jonaha who was no mean fighter himself, knew very well that he was no match for Amaddorap. He talked Amaddorap into staying for dinner with him, and asked him to go bird hunting while his mother, whom Jonaha had told to kill six small chickens, would prepare the dinner. Jonaha took his blow gun with him. Every time he saw a big bird, he aimed his blow gun at it, blew on the blow gun and said: "Bird, fly to my mother's kitchen." After having shot enough birds they went back to Jonaha's house to find that the food was ready and the birds cooked. Amaddorap was so impressed by the magic of the blow gun, that he begged Jonaha to give him the blow gun in payment of all his debt. Pretending at first to be reluctant, Jonaha agreed to do so, but warned Amaddorap that the possession of the blow gun entailed the observance of a taboo, namely, that he should see to it that an ant would not enter its place. When later Amaddorap came furiously to Jonaha, the latter proved before the council of chiefs that Amaddorap had misunderstood the taboo. Amaddorap had taken good care that no ant would enter the container of the glow gun, whereas the taboo's provision was that an ant should not enter the ant's nest.*

4. ANALYSIS

Numbers in parentheses refer to section numbers.

S1. *ùltop ni si jOnáha 'Jonaha's blow gun'* is a noun phrase (5.1.1.1.). The head is *ùltop 'blow gun'*; *ni 'of'* is the possessive preposition (4.4.1.1.) marking the following noun as the possessor; *si* is an auxiliary (4.4.5.) which marks the following word as a personal proper noun; *jOnáha 'personal proper noun'* (4.4.1.).

S2. *parjùji lánis d0 si jOnáha, jàla parjùji táddan, alàì parjùji tálu 'Jonaha is an inveterate gambler, and an itinerant gambler, but a loser (gambler)'* is a compound sentence (5.3.) consisting of one subject, *si jOnáha*, standing in equational constructions with the three coordinate predicators (5.1.2.2.) *parjùji lánis 'inveterate gambler'*, *parjùji táddan 'itinerant gambler'*, and *parjùji tálu 'loser gambler'*, which are linked by the connectives (4.4.4.) *jála 'and'* and *alái 'but'*.

parjùji lánis 'inveterate gambler' is a noun phrase. The head is *parjúji 'gambler'*, formed from *par-* (actor noun formative; 4.6.3.) + the noun base *júji 'gambling game'*; *lánis 'addict; one who knows no bounds'* is a noun attributive to *parjúji*.

d0 is a predicative particle (4.4.3.1.).

si j0náha 'name'; (see S1).

parjùji táddan 'itinerant gambler' is a noun phrase. The head is parjúji (see above); the attribute is the noun táddan 'wandering'. (The noun táddan is only found in attributive position or as a base in secondary derivations, e.g. in the intransitive verb martáddan 'visit another place').

parjùji tálu 'loser gambler' is a noun phrase. The head is parjúji (see above); the attribute is the intransitive verb tálu 'lose'.

S3. ása, dàn na hólán ibána marútan 'therefore he continuously had debts' is a complex sentence (5.3.) with attribute ása 'so; therefore' (a pro-clause; 4.4.1.5.). The head clause consists of

dán 'not', a negative predicate particle (4.4.3.1. and 5.2.1.);

na hólán 'intermittent', the predicator (5.1.2.2.); na is a preposition that makes the following predicative adjective emphatic (4.4.2.1.): hólán 'distant; intermittent';

ibána 'he' (third person singular pronoun; 4.4.1.1.), the subject;

marútan 'have a debt', the complement (5.1.2.2.); it is formed from mar- 'intransitive verb formative'; 4.6.4.2.) + noun base útan 'debt'.

S4. sòŋ0n í ma adóŋ saháli parutanánna, na margòar amaddórap 'so it was that he once had a creditor who was called Father of Punching' is an equational clause (5.1.2.2.). The predicator is the prepositional phrase (5.2.1.1. and 5.1.2.2.) sòŋ0n í 'like that' consisting of the preposition sòŋ0n 'like' + the demonstrative pronoun í 'that'; ma is the narrative predicative particle (4.4.3.1.); the subject is the included clause adóŋ saháli parutanánna, na margòar amaddórap 'he once had a creditor who was called Father of Punching'. This included clause consists of:

adóŋ saháli 'exist once', the predicator; it consists of the intransitive verb adóŋ and the attribute saháli 'once', an adjunct formed from sa- (bound form of sáda 'one') + the measure noun háli 'time (as in three times)';

parutanánna na margòar amaddórap 'his creditor who was called Father of Punching', the subject; it is a noun phrase with head parutanánna 'his creditor' and attribute the rest; parutanánna is made up of par-...-an (noun formative of location; 4.6.3.3.) + útan 'debt' (parután 'person to whom one has a debt') + -na 'his' (third person possessive suffix; 4.4.1.1.); na is the attributive preposition (4.4.2.1.); margóar 'have name' is made up of mar- (intransitive verb formative; 4.6.4.3.) + góar 'name' (noun); amaddórap 'Father of Punching' is a teknonymous proper noun make up of aman- 'father of'

(4.5.1.) + *dórap* (a proper noun from a verb with the same shape, meaning 'punch; strike forcefully').

S5. *nùŋga pìga háli ditùngu amaddòrap í siŋìrna í, alái hòlan bagabága d0 d0h0t jádji dibàEn si j0náha* is a compound sentence (5.5.) consisting of two clauses linked by the connective *alái* 'but'.

The first clause *nùŋga pìga háli ditùngu amaddòrap í siŋìrna í* 'his credit has been demanded several times already by Amaddorap' is passive transitive (5.2.). It consists of:

nùŋga 'already', the completive predicative particle (4.4.3.1.);

pìga háli 'how many times; (here) several times', predicator; it is made up of *píga* 'how many' (interrogative pronoun; 4.4.1.1.) + *háli* 'time' (see S4); it is attributive to:

ditùngu amaddòrap í 'be demanded by (that) Amaddorap', complement; it is an agentive constitute (5.1.2.1.2.) made up of the passive simple mode non-first actor (*di-*; 4.2.4.1.) of *túngu* 'demand', the actor *amaddòrap*, and *i* 'that' (attributive to *amaddòrap*);

siŋìrna í '(that) his credit', the subject; it is a noun phrase with head *siŋìrna* 'his credit', made up of *síŋir* 'credit' + *-na* 'his', and attribute *i* 'that'.

The second clause *hòlan bagabága d0 d0h0t jádji dibàEn si j0náha* 'only promises and pledges were made by Jonaha' is passive transitive (5.2.) in normal emphatic order (5.2.1.2.). It consists of:

hòlan bagabága (...) *d0h0t jádji* 'only promises and pledges', the subject; this is a noun phrase with attribute *hòlan* 'only' (an auxiliary 4.4.5.) and co-ordinate head *bagabága d0h0t jádji* which consists of *bagabága* 'promise' (from *√baga*; 4.6.2.2.) and *jádji* 'pledge', linked by the connective (4.4.4.) *d0h0t* 'and';

d0 (affirmative predicative particle; 4.4.3.1.); it is placed after the first member of the co-ordinate phrase; this often happens. A more common (and more emphatic) order is to place *d0* after the whole subject.

dibàEn si j0náha 'be made by Jonaha', the predicator; it is an agentive constitute made up of the passive simple mode non-first person actor *báEn* 'make' and actor *si j0náha*.

S6. *di na sadári, ditòp0t amaddòrap ma musÉ si j0náha manùngu siŋìrna í* 'one day Jonaha was again visited by Amaddorap to demand his credit' is a complex sentence consisting of attribute *di na sadári* 'one day' and the head clause (the rest)..

di na sadári '(at) one day' is a common expression in storytelling; it consists of the locative preposition *di* 'at' + the attributive preposition *na*, occurring here without head (4.4.2.1.) + *sadári* 'one day', a quantifier made up of *sad-* (an irregular bound form of *sáda*

'one'; cf. 4.3.1.) + *ari* 'day'.

ditòp0t amaddórap ma musÉ si j0náha manùngu siñrna í 'Jonaha was again visited by Amaddorap to demand his credit' is a passive transitive clause. It consists of:

ditòp0t amaddórap 'be visited by Amaddorap', the predicator; it is an agentive constitute consisting of the passive simple mode non-first actor of *tóp0t* 'visit' and actor *amaddórap*;

ma (narrative predicative particle; 4.4.3.1.);

musÉ 'again', a predicator attribute; it is an auxiliary (4.4.5.);

si j0náha, the subject;

manùngu siñrna í 'demand his credit', the complement; it is an objective constitute consisting of the active simple mode (prefix *mañ-*; 4.2.4.1.) of the transitive verb *túngu* 'demand' (see also 3.3.2, IIS1) and the object *siñrna í* '(that) his debt' (see S5).

This clause is in inverted order (of the secondary emphatic order; 5.2.1.2.); that is, the predicator attribute *musÉ* 'again' precedes the subject, thus making *musÉ* emphatic. The direct order is *ditòp0t amaddórap ma si j0náha musÉ manùngu siñrna í*.

Note: Passive clauses are found very frequently in Toba Batak. In a free translation, many of those are translated into the corresponding active clauses in English.

S7. *àng0 sahàli ón, dàñ adóñ bÉ idá0n hasabár0n di bohína* 'as for this time, there was no longer any patience to be seen on his face' is a complex clause consisting of the minor clause (5.2.2.2.) *àng0 sahàli ón* 'as for this time' as attribute and the head clause the rest.

àng0 'if; as for' is a subordinating conjunction (4.4.2.2.).

sahàli ón 'this (one) time' is a noun phrase (5.1.1.1.) consisting of head *saháli* 'one time; once' (see S4) and attribute *0n* 'this' (a demonstrative pronoun; 4.4.1.1.).

The head clause consists of:

dàñ 'not' (negative predicative particle; 4.4.3.1.) + auxiliary (4.4.5.) *bÉ* function as the modal particle; the combination means 'no longer';

adóñ 'exist', the predicator; it is an intransitive verb;

idá0n 'will be seen; to be seen', the complement; it is the promissory mode of *ída* 'see';

hasabár0n 'patience', the subject, a noun made up of *ha*-...-*0n* (noun formative; 4.6.3.3.) + adjective base *sábar* 'patient';

di bohína 'on his face'; attribute to the complement *idá0n*; it is a prepositional phrase consisting of the preposition *di* 'at' + the noun *bóhi* 'face' + *-na* 'his' (third person singular possessive suffix).

A more literal translation of this clause is 'patience no longer

exists to be seen on his face'.

S8. makkúliŋ ma ibána, didók ma tu si jónáha, alé laé 'he spoke; he said to Jonaha: "I say, cousin".' is a compound sentence consisting of makkúliŋ ma ibána 'he spoke' as first member and the rest the second member.

The first clause consists of:

makkúliŋ 'speak', the predicator, an intransitive verb made up of maŋ- (intransitive verb formative; 4.6.4.1.2.) + √huliŋ;

ma (narrative predicative particle; 4.4.3.1.);

ibána 'he', the subject, third person singular pronoun.

The second clause consists of:

didók 'said by him', the predicator, the passive simple mode non-first person actor (4.2.4.1.) of dók 'say';

ma (narrative predicative particle; 4.4.3.1.);

tu si jónáha 'to Jonaha', the complement, a prepositional phrase made up of the preposition tu 'to' + si jónáha (see S1); it is attributive to the predicator;

alé laé 'I say, cousin', the subject, a vocative minor clause (5.2. (5.2.2.2.) made up of the interjection alé 'I say' + noun láE 'cousin' (for stress alternation, see 3.3.3., SR4).

S9. nùŋŋa malelekkú s0 digàrar hó utaŋmí 'it's already too long that thy debt has not been paid by thee' is an equational clause (5.1.2.2.). It consists of:

nùŋŋa 'already', completive predicative particle (4.4.3.1.);

malelekkú 'too long', the predicator, the excessive degree of the adjective lelén 'long (time)' made up of ma-...-hu + lelén (4.2.5.; ma-...-hu is used by some speakers for our pa-...-hu);

s0 digàrar hó utaŋmí 'thy debt has not been paid by thee', the subject, an included clause. This clause consists of:

s0 'not', the negative predicative particle in dependent clauses (4.4.3.1.);

digàrar hó 'be paid by thee', the predicator, an agentive constitute made up of the passive simple mode non-first person actor of gárar 'pay' + actor h0 'thou' (second person singular pronoun; 4.4.1.1.);

utaŋmí '(that) thy debt', the subject; it is made up of noun útaŋ 'debt' + -mu 'thy' (second person singular possessive suffix; 4.4.1.1.) + enclitic demonstrative pronoun i 'that' (4.4.1.1. and 3.3.2., IIS9).

S10. jàla nùŋŋa lója áu pabegebège bagabàga s0 máhap sian hó 'and I am already tired of repeatedly hearing unceasing promises from thee' consists of the connective (4.4.4.) jála 'and' and a clause (the rest). The connective links the clause with the preceding sentence.

The clause is equational; it consists of:

núnna 'already', the completive predicative particle (4.4.3.1.);

lója 'be tired', the predicator, an intransitive verb;

áu 'I', the subject, first person singular pronoun (4.4.1.1.);

pabegebège bagabàga s0 máhap sian hó 'hear unceasing promises from thee', the complement. This phrase consists of the objective constitute pabegebège bagabàga s0 máhap 'hear unceasing promises' as head and the prepositional phrase sian hó 'from thee' as attribute. The objective constitute consists of:

pabegebége 'hear repeatedly', a transitive verb made up of pa- (formative; 4.6.4.2.3., (5)) + doubling of the transitive verb bége 'hear';

bagabàga s0 máhap 'unceasing promises', the object of pabegebége; it is a noun phrase consisting of head bagabàga 'promise (N)' (see S5) and attribute s0 máhap 'unceasing' made up of s0 (negative particle in subordinate phrases; 4.4.3.1.) + máhap 'be satiated; feel satisfaction', an intransitive verb; the meaning 'unceasing' is a transferred meaning: one stops eating when one is full or satisfied.

The attribute sian hó consists of:

sían 'from', a locative preposition (4.4.2.1.);

h0 'thou', second person singular pronoun (4.4.1.1.).

S11. ÓppE, gárar utámmu sa0nnári 'therefore pay they debt now' is a complex sentence (5.3.). It consists of attribute ÓppE 'therefore; so now', a pro-clause (4.1.4.5.) and the head clause (the rest).

This clause is a transitive imperative minor clause (5.2.2.2.).

It consists of:

gárar 'pay', the imperative mode of gárar (4.2.4.1.);

utámmu 'thy debt', the object, made up of noun úta 'debt' + -mu 'thy';

sa0nnári 'now', the complement, an adjunct (4.3.4.).

S12. m010 s0 í, daiómmu ma pohulpohulhón 'if not (that), (this) my fist will be tasted (= felt) by thee' is a complex sentence consisting of attribute m010 s0 í and head clause (the rest).

m010 s0 í 'if not that' is a subordinate minor clause introduced by the subordinating conjunction m010 'when; if' (4.4.2.2.); s0 'not' (negative predicative particle in subordinate phrases; 4.4.3.1.); í 'that', a demonstrative pronoun, the predicator of the minor clause.

daiómmu ma pohulpohulhón 'this my fist will be tasted by thee' consists of:

daiómmu 'will be tasted by thee', the predicator, made up of the passive promissory mode (-0n; 4.2.4.1.) of dáí 'taste' + -mu (second person singular actor suffix; 4.4.1.1.);

ma (narrative predicative particle, 4.4.3.1.); here it makes the threat implied in the promissory mode more imminent;

pohulpohulhón '*this my fist*', the subject; it is made up of the noun pohulpóhul '*fist*' (a noun of instrument formed by doubling from Vt póhul '*make into a ball by the hand*'; 4.6.3.3., (1b)) + -hu '*my*' (first person singular possessive suffix; 4.4.1.1.) + enclitic demonstrative pronoun On '*this*' (4.4.1.1. and 3.3.2., IIS9).

S13. dibòt0 hó d0 alána, ubbàEn margòar amaddórap áu '*the reason is known by thee why I am called Father of Punching*' is a complex sentence consisting of the head clause dibòt0 hó d0 alána '*the reason is known by thee*' and the attributive clause (the rest) introduced by the subordinating conjunction ubbàEn (introducing a clause of result; 4.4.2.2.; morphemically, the active completive-participial mode (UM-) of báEn '*make; do*'; (4.2.4.1.).

dibòt0 hó d0 alána '*the reason is known by thee*' consists of:

dibòt0 hó '*be known by thee*', the predicator, an agentive constitute made up of the passive simple mode non-first person actor (4.2.4.1.) of bóto '*know*' + h0 '*thou*' (see S9);

d0 (affirmative predicative particle; see S2);

alána '*its reason*', the subject; it is made up of noun ála '*reason; cause*' + -na '*its*' (third person singular possessive suffix).

The dependent clause margòar amaddórap áu '*I am called Father of Punching*' consists of:

margòar amaddórap '*have the name Father of Punching*', the predicator; it consists of head margóar '*have name*' (see S4) and attribute amaddórap '*Father of Punching*' (see S4).

S14. dáda nàun maramaniáha áu, ài s0 marhas0hótan d0pE áu '*it's not that I already have a teknonymy (= have a child), for I am not married yet*' is a complex sentence. The head clause is dáda nàun maramaniáha áu '*it's not that I already have a teknonymy*' and attribute the rest introduced by the subordinating conjunction ái '*for; because*' (4.4.2.2.).

The head clause is an equational clause; it consists of:

dada '*not*' (emphatic negative predicative particle; 4.4.3.1.);

nàun maramaniáha '*(one) having a teknonymy*', the predicator; it is a nominalised intransitive verb phrase consisting of náun '*already*' (the completive predicative particle in nominal phrases; 4.4.3.1.) and the intransitive verb maramaniáha '*have a teknonymy*', made up of mar- (intransitive verb formative) + proper noun àmaniáha '*father of So-and-so*' (from ámani '*father of*'; 4.3.1. + áha '*what*', here '*So-and-so*'; 4.4.1.1.);

áu '*I*', the subject. A more literal translation of this clause is

'I am NOT (a person) who has a teknonymy already'.

The dependent clause *sò marhas0hótan d0pE áu 'I am not married yet'* consists of:

s0 'not' (the negative predicative particle in dependent clauses; 4.4.3.1.) + *d0pE 'still'* (progressive predicative particle; 4.4.3.1.); the combination means *'not yet'*;

marhas0hótan 'be married', the predicator; it is an intransitive verb made up of *mar-* (formative) + noun base *has0hótan 'married state'*, which in turn is made up of *ha-...-an* (noun formative; 4.6.3.3., (5c) + intransitive verb base *sóh0t 'be no longer single; stop'*;

áu 'I', the subject.

S15. *alái àla ni gòst0ŋ ni b0t0hókku d0 'but because of the strength of my arm'* is a minor clause introduced by the connective *alái 'but'* (4.4.4.).

àla ni gòst0ŋ ni b0t0hókku 'because of the strength of my arm' is the predicator of the minor clause; it is a prepositional phrase (see below);

d0 (affirmative predicative particle; 4.4.3.1.).

The prepositional phrase above consists of:

àla ni 'because of', a phrase that functions as a preposition (4.4.2.1.8.); it is made up of the noun *àla 'reason; cause'* + *ni* (possessive preposition; 4.4.2.1.);

gòst0ŋ ni b0t0hókku 'the strength of my arm' is a noun phrase consisting of head *gòst0ŋ 'strength'* + *ni* (possessive preposition) + *b0t0hókku 'my arm'*, which in turn is made up of *b0t0h0n 'arm'* + *-hu 'my'* (see S12).

S16. *àng0 póla d0 hudórap hálak, dàŋ na bóí bE hirím0n ni í bErEŋónna mataniári 'if I do punch a person, it WILL no longer be possible for him to expect to see the sun'* is a complex sentence consisting of attribute *àng0 póla d0 hudórap hálak* and head the rest.

àng0 póla d0 hudórap hálak 'if a person IS punched by me' consists of:

àng0 'if', a subordinating conjunction introducing the dependent clause (4.4.4.2.);

póla, predicator; it is an auxiliary which has many meanings (cf. 4.4.5. and 5.2.1.). Here it makes the predicator emphatic; *hudórap 'IS punched by me'* (see below);

hudórap 'be punched by me', the predicator; it is the passive simple mode first person actor (4.2.4.1.) of *dórap 'punch'*;

d0 (affirmative predicative particle; 4.4.3.1.); when it occurs in a dependent clause, it makes the predication emphatic;

hálak 'a person', the subject.

The head clause *dàŋ na bóí bE hirìmOn ni í bErEŋÓnna mataniári* 'it WILL no longer be possible for him to expect to see the sun' consists *dàŋ* (negative predicative particle) + *bE* (auxiliary); the combination means 'no longer';

na bóí 'capable; possible', predicator; *bóí* 'capable; possible' is an adjective made emphatic by the preposed *na* (see S3);

hirìmOn ni í 'will be hoped by him', the complement (for its position see 5.2.1.1.); it is an agentive constitute consisting of the promissory mode of *hírim* 'expect' + *ni* (preposition that marks a separate actor) + *i* 'that' (the actor), a demonstrative pronoun used here to refer to *hálak*; (ordinarily, *i* is not used to refer to a person; the connotation here is the speaker's contempt for the punched person);

bErEŋÓnna mataniári 'the sun will be seen by him', the subject; it is an included clause consisting of:

bErEŋÓnna 'will be seen by him', the predicator; it is made up of the promissory mode (suffix *-On*) of *bÉrEŋ* 'see' + *-na* (third person singular actor suffix);

mataniári 'sun', the subject; it is made up of *máta* 'eye' + *ni* (possessive preposition) + *ári* 'day'.

A more literal translation of the sentence is 'if a person IS punched by me, that the sun will be seen by him CAN no longer be expected by him'.

APPENDIX II

SPECTROGRAPHIC ANALYSIS OF SOME BATAK PHONEMES

The analysis of Batak presented in this study is primarily based on theoretical reasoning. The method employed consists entirely in searching for contrasts. Articulatory characteristics are taken into account in the appropriate places.

Another procedure that can help analysts get the answers is by making use of the methods of acoustic phonetics. One such method is the use of the sound spectrograph, which can make charts of the utterances fed into the machine. This kind of chart is called a spectrogram. The sound spectrogram is especially useful for measurements of pitch and duration. It is also successfully used in the analysis of phoneme qualities, especially of the vowels.

The writer has also made some use of spectrograms, both narrow and broad band, not so much in analyzing his speech, but rather as a means of testing the conclusions he had made. Most of the conclusions he had arrived at have been verified by the spectrographic readings, but some have been contradicted, so that he had to modify his first conclusions in the light of the new evidence, or to search for other evidences to substantiate his earlier findings.

Some of the more useful pertinent spectrograms will be discussed in brief below. Those spectrograms are included in the thesis, but as there is only one copy of each, the spectrograms cannot be included in every copy of this study.

1. VOWEL LENGTH

A comparison of vowel-lengths, in this case the /ɔ/ is made with Spectrograms No. 1-7a. The a series are in broad band representation, and the b series are the corresponding narrow band spectrograms. The durations in this discussion will be in centiseconds (cs). However,

all numberings for duration on the grams themselves are in millimeters (mm).

Spectrogram No. 1a:

| | | | | |
|------------------|--------|---------|------|-------|
| Utterances: | /góar/ | /gótap/ | /sɔ/ | /tɔp/ |
| Duration of /ɔ/: | 13.6 | 12.2 | 14.3 | 8.4 |

Although the difference in the duration of the /ɔ/ of /góar/ and that of /gótap/ is greater than the difference between those of /góar/ and /sɔ/, the first two are grouped together due to their phonotactic similarity, i.e. occurrence in medial open syllables, and the fact that the difference is apparently caused by the difference in the following phonemes, namely a vowel versus a voiceless stop. The main difference between the /ɔ/ in /sɔ/ on the one hand and the /ɔ/ in /góar/ and /gótap/ on the other hand is the fact that as a rule /sɔ/ will be followed by a /+/, which will tend to prolong it, but nevertheless it can be shorter than the /ɔ/ of /gótap/, if the latter gets a stronger stress (see Spectrogram No.2).

Another thing that accounts for the greater length of /sɔ/ is, as observed by Pike (1945:11) for American English, the fact that the entire pitch contour falls on it.

That the /ɔ/ in /tɔp/ is much shorter than the others is obvious from the spectrogram.

Spectrogram No.1b is a narrow band representation of the above utterances.

Spectrogram No. 2: /sɔ+hugótap/

Duration of /ɔ/: 9.9 11.4

See comment above under Spectrogram No.1a.

| | | | | |
|---------------------|---------|------|-------|-------|
| Spectrogram No. 3a: | /gótap/ | /sɔ/ | /tɔp/ | /tɔs/ |
| Duration of /ɔ/: | 9.1 | 15.6 | 8.4 | 9.5 |

This demonstrates the fact that the /ɔ/ in /tɔs/ is longer than that in /tɔp/.

Spectrogram No. 3b: Narrow band representation of No. 3a.

| | | | | |
|---------------------|---------|------|-------|-------|
| Spectrogram No. 4a: | /gótap/ | /sɔ/ | /tɔp/ | /hɔt/ |
| Duration of /ɔ/: | 11.0 | 17.1 | 9.5 | 8.4 |

The greater difference between /tɔp/ and /hɔt/ may be due to the more rapid pitch fall in /hɔt/ (see Spectrogram No. 4b).

| | | | | |
|---------------------|--------|------|-------|-------|
| Spectrogram No. 5a: | /góar/ | /sɔ/ | /tɔp/ | /sɔr/ |
| Duration of /ɔ/: | 14.3 | 15.2 | 9.1 | 10.7 |

Illustration of the greater length of the /ɔ/ in /sɔr/ than in /tɔp/.

Spectrogram No. 6a: /gʃar/ /sɔ/ /tɔl/ /bɔm/
 Duration of /ɔ/: 16.0 20.5 14.3 12.8

Illustrating the fact that the /ɔ/ is longer in /tɔl/ than in /bɔm/.

Spectrogram No. 7a: /ʃ t ɔ/ /ɔ t ʃ/
 Duration of /ɔ/: 12.2 9.9 7.6 15.2
 Duration of /ɔ/: 20.1 14.3 9.9 18.6

This illustrates the fact that the stronger stressed vowel under comparable conditions tends to be longer than a weaker stressed one. The numbers on the first line indicate the durations without the glides. Those on the second line indicate the durations of the vowel /ɔ/ with the on- and off-glides.

Spectrogram 7b: /ʃ t ɔ/ /ɔ t ʃ/
 First formant: 600 550 550 625
 Second formant: 1000 1050 1050 1000

This demonstrates, though pretty faintly, the centering tendency of a weak-stressed vowel. The first formant indicates the height of the tongue position, the greater the formant the lower the position. The second formant indicates the horizontal position of the highest part of the tongue, the greater the formant the more front will be the tongue.

2. VOWEL FORMANTS

The first and second formants of the seven vowels are measured from the spectrograms No. 8 and No. 9. The results are as follows:

Spectrogram No. 8: i e ɛ a
 First formant: 350 425 500 700
 Second formant: 2250 2100 1825 1325

Spectrogram No. 9: a ɔ o u
 First formant: 750 625 500 325
 Second formant: 1325 875 825 825

The vowels are pronounced with a preceding /b-/, and these syllables are pronounced as the first parts of:

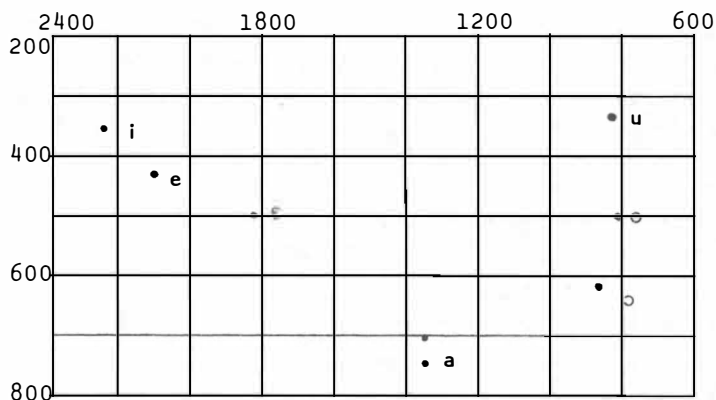
 /bʃsa bɛta bɛrɛŋ bára/
 and /bára bóta bóla bútar/.

From these formants a chart is made below which shows a striking resemblance to the vowel chart on p.12.

DIAGRAM 6
VOWEL CHART

The vertical index is the frequency of the Second Formant.

The horizontal index is the frequency of the First Formant.



3. JUNCTURES

Spectrogram No. 10. Only one gram is made to illustrate the fact that the terminal junctures are tempo phenomena. The blank space on the gram represents /k/ and /||/. Their total duration, 12.2 centi-seconds, is compared here with that of a normal vowel, in this case the following vowel /a/ of /na/, with duration is 4.6. centiseconds. These durations are very low, due to the rapid speed in which the utterance is pronounced so that the whole sentence may be accommodated in the gram, but it does not take too much from the results, as the objective is to get the ratio between the duration of a normal vowel and the /||/.

Spectrogram No. 11. This is intended to illustrate the pitch rise for the /||/, here on the syllable /-lak/. Although it does not show well on the spectrogram, the aural impression is of a definite pitch rise. The pitch fall for the /#/ also does not show well, perhaps due to the speed of the utterance. However, these inflection features of the terminal junctures can be seen very clearly on the pitch spectrograms (see Nos. 14-19).

An attempt to compare the influence of the presence and absence of a /+/- is made with Spectrograms No. 12 and No. 13, although the pitch and stress conditions are not alike. The durations of the phonemes

/a, p, +, i/ on No. 12 are compared with those of /a, p, i/ on No. 13, thus:

| Duration of: | | | | |
|-----------------------------------|-----------------|------|------|-----|
| | total utterance | a | p(+) | i |
| No. 12: /f ^s ap+i.../: | 75.6 | 5.3 | 9.9 | 6.6 |
| No. 13: /i ^s ápi.../: | 80.9 | 12.1 | 8.7 | 4.2 |

Not much can be concluded from these grams, but we see that /p+/ is longer than /p/. Secondly, that the final /p/ is represented by a solid blank, whereas the medial /p/ is indicated by some striations. This seems to suggest that a final /p/ is more voiceless than a medial /p/, or, that a medial /p/ is not so silent. It is difficult for the writer to make conclusions about this in view of his very limited experience with spectrograms and insufficiency of material.

Corresponding narrow band spectrograms are added to both, respectively numbered 12a and 13a.

4. PITCH CONTOURS

The spectrograms dealing with pitch contours will not all be measured as to the fundamental pitch of their syllables, but on each one the height of the tenth harmonic of the pitch points will be indicated by means of vertical lines, from whose relative heights the pitch ratios can be inferred. Phonemic pitch is different from absolute pitch, so that, sometimes, rather different pitches will be regarded as allophones, and therefore classed into one pitch phoneme, and on the other hand slightly differing pitches may belong to different pitch phonemes.

When the fundamentals are indicated, they will only be for the pitch points indicated by stress marks. The numbers indicate cycles per second (cps), and are measured by means of the tenth harmonic.

Not much comment will be given, as the spectrograms will speak for themselves, and when necessary we can always refer back to the appropriate passages in the section on pitch (p.33 ff). The suspended lines on the grams are the phonemic divisions.

Spectrogram No. 14: ^{2 3} /dǐjǎbu+^{2 1} dɔ+ibânǎ #/
 Fundamentals: 145 175 140 95
 Ordinary statement pitch contour. (See p.34.)

Spectrogram No. 15: ^{2 3} /dǐjǎbu + dɔ + ^{2 3} ǐbânǎ || /
 Fundamentals: 175 140 162½ 180
 Ordinary question contour.

Spectrogram No. 16: ^{2 2} /dijábu + dɔ + ^{3 3} ibânǎ || /

Fundamentals: 160 195 200

Question pitch contour with emphasis on /ibana/, see p.36.

Spectrogram No. 17: ^{2 3} /dijábu + dɔ + ^{3 3} ibâna || /

Fundamentals: 230 230

Pitch contour with emphasis on /-jabu/, or if the answer given is repeated in the question.

Spectrogram No. 18: ^{2 3 3} /dɪjábũ/ ^{2 3 1} /dɪjábũ #/

Fundamentals: 175 215 230 190 230 125

The first one is the same type as the one on Spectrogram No. 17, i.e. repeating the answer in the question to verify it or to express mild surprise. The second utterance is the normal contour, here as an answer.

Spectrogram No. 19: ^{2 4 2} /dɪ j á b ũ / ^{2 4 1} /dɪ j á b ũ # /

Fundamentals: 235 300 235 235 312.5 160

The first contour is one of great doubt or irony. The second one is an emphatic answer to this or in general. (see p.37.)

5. PHONOTACTICS

One feature discussed in the section on Phonotactics is illustrated here by a couple of grams, namely, the phenomenon of double consonants. Their phonemic status has been demonstrated on p.44 by means of contrasts. Here the contrasting qualities can be seen as approximately double duration and the shortening influence of the double consonants on the preceding vowels.

| | | | | |
|---|----------|---------|------------------|-------------------|
| Spectrogram No. 20: | /túllaŋ/ | /túlaŋ/ | /pítu/ 'door' | /pítu/ 'seven' |
| Duration in cs. of vowel and following medial consonant | 6.1 13.6 | 8.4 6.8 | 3.8 19.7 | 6.8 10.3 |

| | | |
|---------------------|---------------|-------------|
| Spectrogram No. 21: | /h á d d a ŋ/ | /h á d a ŋ/ |
| Duration in cs. | 6.8 11.4 | 12.5 4.9 |
| and | /t ɔ p p a/ | /t ɔ p a/ |
| Duration in cs. | 6.1 13.6 | 12.2 8.4 |

In conclusion it should be mentioned that the frequencies and durations in different spectrograms are different in absolute numerical values, though they may be classed in the same phonemes. That is a

thing that can not be avoided, as the speech of an individual fluctuates in rapidity and pitch from utterance to utterance. However, the relative connections between the various items can be observed from the grams individually. Keeping this in mind, the sound spectrogram is a useful tool in the hands of language analysts, and much more can still be expected from the specialists in acoustic phonetics that can be of help to the linguists and ultimately to language learning.

APPENDIX III

THE BATAK ALPHABET

As mentioned in the Introduction, the Batak alphabet is one of the forms of the Devanagari alphabet in Indonesia. Other forms are: the Javanese alphabet on Java, the Lampung alphabet in South Sumatra, the Bugis alphabet on the island of Sulawesi.¹






Most of the literature in the Batak alphabet is works on magic, astrology, and mythology. The writing was done by inscribing marks with an iron stylus on a piece of bamboo stem. These bamboo stems, about one foot in length, were used as letters and sent by messenger or given to a passing traveller who would also go by the village of the addressee. More important writings, as magic formulas, legends, and astrological 'science', were inscribed on a long strip of bark, which could be folded and unfolded accordion-like, with a thorn-like part of the bagot palmtree, called tarugi. Black dye was used as ink. This kind of book is called pustaha.

Since the introduction of paper, ink and pen, some books have been written containing narratives, Bible stories and historical accounts, but nowadays writing books in the Batak alphabet has been discontinued, and all writings are done in Latin characters. The Batak alphabet is still taught in Junior High Schools in the Batak speaking areas as a cultural feature only.

The Batak alphabet is syllabic (or biphonemic except the vowels, see below). The characters with the corresponding Latin letters orthographically are listed below. There are twenty different symbols and six diacritics.

¹For geographical names, see Map I.


The vowels are:

 = a
 = i
 = u
 = e
 = o






The letters e and o are formed by using an additional mark or diacritic with the a; for the e a horizontal short line, called *hatadingan*, is placed above it to the left; and, for the o a cross-like mark, called *siala*, is placed to its right. A special variety of u can also be made by placing a hook-like mark, called *haboruan*, to the right of the a, thus:

 or  = u.











A special variety of i can also be formed by placing a small circle, called *haulian*, to the right of the a, thus:

 = i.


The consonants without any additional mark contain the pertinent consonant plus the vowel a. To make a combination of a consonant and another vowel a mark is superimposed or juxtaposed to it, for example:

 = na
 = ne (i.e. na plus *hatadingan*)
 = ni (i.e. na plus *haulian*)
 = no (i.e. na plus *siala*)
 = nu (i.e. na plus *haboruan*)


When a small horizontal line, called *hamisaran*, is placed above the character to its right, the velar nasal /ŋ/, orthographically ng, is added to the whole character, thus:


 = ang  = nang
 = ing  = ning
 = eng  = neng
 = ong  = nong
 = ung  = nung.

A simple consonant symbol can be formed by placing an oblique accent, called *pangolat*, to the right of the character, thus:

 = n

so that  = an

 = in, etc.

and  = nan

 = non, etc.

The consonantal characters are:

| | | | |
|---|--|---|----------------------------|
| Ⓐ | = ba | — | = pa |
| ∧ | = da | ⌘ | = ta |
| ∟ | = ga | ⌚ | = ha or ka ¹ |
| ⌵ | = dja (phonemically: /ja/) | ⌘ | = ma |
| Ⓐ | = na | ⌚ | = ra |
| ∧ | = nga | Ⓐ | = wa |
| ⌘ | = sa | ⌚ | = ja (phonemically: /ya/) |
| ∟ | = la | ⌚ | = tja (phonemically: /ca/) |
| ⌚ | = nja (phonemically: /nya/) ² | | |

In this alphabet there is only one character for the two phonemic entities /e/ and /ɛ/, namely ⌚, etc.; and also one for /o/ and /ɔ/, namely: ⌚x, etc.

The present Batak Latin spelling does not deviate very much from the original Batak spelling, except the changes that had to be made as a consequence of the conversion from a syllabic to a letter alphabet.

¹About the original allophonic distribution of /h/ and /k/, see Chapter I, Section 1, page 4.

²Used only in Angkola-Batak, as this sound does not occur in Toba-Batak.

BIBLIOGRAPHY

- BACH, Emmon and R.T. HARMS
 1968 *Universals in linguistic theory*. New York: Holt, Rinehart and Winston.
- BLOOMFIELD, Leonard
 1955 *Language*. London: George Allen and Unwin.
 ? *Menomini*. (Post-humously edited by C.F. Hockett.)
- BRANDSTETTER, Renward
 1915 *Die Lauterscheinungen in den Indonesischen Sprachen*. Lucerne: E. Haag.
- CHOMSKY, N.
 1965 *Aspects of the theory of syntax*. Cambridge, Mass.: M.I.T. Press.
- CORNELIUS, Edwin T., Jr
 1953 *Language teaching: a guide for teachers of foreign languages*. New York: Thomas Crowell.
- DYEN, Isidore
 1964 *A lexicostatistical classification of the Austronesian languages*. IJAL Memoir 19.
- GLEASON, H.A., Jr
 1955 *An introduction to descriptive linguistics*. New York: Henry Holt.
- HADEN, Ernest F.
 1953 *How to pronounce Spanish*. New York: Henry Holt.
- HALLIDAY, M.A.K.
 1973 *Explorations in the functions of language*. London: Edward Arnold.
- HILL, Archibald A.
 1958 *Introduction to linguistic structures*. New York: Harcourt Brace.
- HOCKETT, C.F.
 1958 *A course in modern linguistics*. New York: Macmillan.
- JOOS, Martin
 1948 *Acoustic phonetics*. Language monograph series, No.23. Linguistic Society of America.
- LADO, Robert
 1957 *Linguistics across cultures: applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- LAMB, Sydney M.
 1964 The sememic approach to structural semantics. *American Anthropologist*, 66/3(2):57-78.
- LANGACKER, R.W.
 1974 *Fundamentals of linguistic analysis*. New York: Harcourt Brace Jovanovich.

- LOEMBANTOBING, A.
1920 *Siadji donda hatahoetan. Weltevreden.*
- LUMBANTOBING, Arsenius
1921 *Rapot ni angka binatang. Lagoeboti.*
- LOPEZ, Cecilio
1939 *Studies on Dempwolff's Vergleichende Lautlehre des Austronesischen Wortschatzes.* Manila: Bureau of Printing.
- MARCKS, O.
1912 *Kurze praktisch-methodische Einführung in die Bataksprache.* Laguboti: Missions-Druckerei.
- MEERWALDT, J.H.
1904 *Handleiding tot de beoefening der Bataksche taal.* Leiden: E.J. Brill.
- NABABAN, P.W.J.
1958 *A phonemic analysis of Batak.* M.A. thesis, University of Texas, Austin.
- PALMER, Harold E.
1921 *The oral method of teaching languages: a monograph on conversational methods...* Cambridge: W. Heffer.
1932 *This language-learning business...* New York: World Book Company. Reprinted by Oxford University Press 1969.
- PIKE, Kenneth L.
1945 *The intonation of American English.* Ann Arbor: University of Michigan Press.
- PIKE, K.L. and E.G. PIKE
1977 *Grammatical analysis.* Dallas: Summer Institute of Linguistics and the University of Texas at Arlington.
- PERCIVAL, W.K.
1964 *A grammar of Toba-Batak.* Ph.D. dissertation, Yale University, New Haven, Connecticut; now published as
1981 *A grammar of the urbanised Toba-Batak of Medan.* PL, B-76.
- PETERSON, Gordon E.
1954 *Phonetics, phonemics and pronunciation: spectrographic analysis.* Monograph series on languages and linguistics, vol.6.
- SIHOMBING, Dr. J.
1956 *Almanak Huria Kristen Batak Protestant.* Tarutung.
- SIHOMBING, T.M.
1959 *Pabidangkön dohot Pabagashon.* Pematang Siantar.
- SITUMORANG, S.
1955 *Pustaka.* Medan.
- TRAGER, George L. and Henry Lee SMITH, Jr
1956 *An outline of English structure.* Washington, D.C.: American Council of Learned Societies.
- TUUK, H.N. van der
1861 *Bataksch-Nederduitsch woordenboek.* Amsterdam: Muller.
1864- *Tobasche spraakkunst, 1. stuk (vol.1), 1964: Klankstelsel.* Amsterdam:
1867 Frederik and Muller. 2. stuk (vol.2), 1967: *De woorden als zindeelen.* Amsterdam: Depot van het Nederlandsch Bijbelgenootschap.
- VOORHOEVE, P.
1955 *Critical survey of studies on the languages of Sumatra.* The Hague: M. Nijhoff.