CHAPTER II
LITERATURE REVIEW

2.0 Introduction

The discussion provided in this chapter is presented in two sections. The first section is a discussion on dialect and dialectology and the second section is on the review of earlier phonological works done in Khasi.

2.1 The notion of dialect

Any language with a reasonably large number of speakers will exhibit dialect variation especially if there are geographical barriers separating groups of people from each other, or if there are divisions of social class (Crystal, 2003). The question is what is a dialect? The key term ‘dialect’ itself has various non-technical meanings. Malmkjer (1995) mentioned that some of the dialects are mutually incompatible and most of them are also implicated in partisan, often negative, attitudes to non-standard speech. These meanings are usually rejected or seriously modified by dialectologists. Most linguists have, however, pointed out that it was not always an easy task to give a definition of the term ‘dialect’ particularly in order to distinguish it from the term ‘language’. In classical Greek, dialects are used to refer to distinct written varieties, each associated with a different area and used for a different kind of literature. The French also used the term ‘dialecte’ to refer to regional varieties which are written and have a literature and the term ‘patois’ to refer to regional varieties which do not have a written form. The meaning of the term ‘dialect’ as used in English is, therefore, quite different from the way it is used in Greek and French. Petyt (1980: 16) stated that in the case of English, another term is often encountered, that is, ‘accent’. Petyt (ibid) further states that ‘many people feel that ‘accent’ is a more restricted term than ‘dialect’. He puts forward the distinction of the terms, ‘accent’ and ‘dialect’ made by linguists in the following form ‘ the term dialect refers, strictly speaking, to differences between kinds of language which are differences of vocabulary and grammar as well as pronunciation. The term ‘accent’, on the other hand, refers solely to ‘differences in pronunciation’ (Petyt 1980:16).
Dialectologists and sociolinguists have, however, attempted to define the terms ‘dialect’, ‘language’ and ‘accent’ from different perspectives.

2.1.1 Definitions of dialect

A dialect is commonly defined as a variety (varieties) of a language. McConnell (1979) states that ‘a dialect is a subvariety of a language, either regional or social. It is distinguished from other subvarieties of the same language by unique combination of language features: pronunciation,(including stress and intonation); grammatical forms; words and expressions; meanings of words and expressions’. Chamber and Trudgill (1980:3) state that in a common usage, dialect is a ‘substandard, Open status, often rustic form of a language, generally associated with the peasantry, the working class, or other groups lacking in prestige’. They further add that dialect is also a ‘term which is often applied to forms of a language, particularly those spoken in more isolated parts of the world, which have no written form’. In such definitions, therefore, ‘dialects are often regarded as some kind of (often erroneous) deviation from a norm- as aberrations of a correct or standard form of a language’ (Chamber and Trudgill,1980:3). It is to be noted that Chambers and Trudgill like most other sociolinguists and dialectologists have, disagreed with these viewpoints. They clearly mentioned that the above viewpoints on dialect will not be adopted in their book Dialectology(1980). They preferred to accept the notion that ‘all speakers are speakers of atleast one dialect’(ibid p 3). They are of the viewed that even Standard English, is a dialect as any other form of English. They rejected the notion that ‘any one dialect is in any way linguistically superior to any other’. In order to avoid complexity over the term ‘dialect’, Chambers and Trudgill (1980) suggest alternative terminology, which they consider, is more neutral. Chambers and Trudgill (1980:5) state that ‘we shall use ‘variety’ as a neutral term to apply to any particular kind of language which we wish, for some purpose, to consider as a single entity’. They further state that ‘the term will be used in an adhoc manner in order to be as specific as we wish for a particular purpose’. The notion of dialect put forward by Chambers and Trudgill is accepted and used in present sociolinguistic and dialectological studies. Petyt (1980:11) defined dialect as ‘different forms of the same language’. To Petyt, this means that using a language necessarily involves using one of its dialects. The
definition of dialect provided by Petyt is similar to Chambers and Trudgill notion of dialect. Crystal (2003) defined dialect as a “regionally or socially distinctive variety of language, identified by a particular set of words and grammatical structures”. According to Malmkjer (1995:123), dialect refers to ‘a geographical variety of a language’. He mentions that dialectologists, however, have increasingly used the term to refer to any user-defined variety, that is, any variety associated with speakers of a given type, whether geographically or otherwise defined, e.g. members of a given social class, males/females, people of shared ethnic background, etc. Bussmann (1998:307) defines dialect as ‘a linguistic system that (a) shows a high degree of similarity to other systems so that at least partial mutual intelligibility is possible; (b) is tied to a specific region in such a way that the regional distribution of the system does not overlap with an area covered by another such system; (c) does not have a written or standardized form, i.e. does not have officially standardized orthographic and grammatical rules’. He further explains that apart from this narrow definition, the term ‘dialect’ is used by linguists in various other senses. For example, the broader use of ‘dialects’ would refer to ‘the various languages that stem from a single ancestral language, such as the ‘Romance dialects’ from Latin’. He also added that in the investigation of the conditions and the origin of the dialectal structure (dialectology), dialects must be defined as individual languages in which extralinguistic aspects like topography (mountains and rivers as natural borders), trade routes, and political and religious centers are taken into account alongside strictly linguistic criteria. Seen from a genetic and historical perspective, dialects must be considered older than standardized languages and can, therefore, in their modern form, be seen as a reflex of a historical development. Since dialects—owing to their oral tradition and lack of standardization—are ‘more natural’ than standardized languages, they are particularly suited for testing linguistic hypotheses about historical processes, as is evident in both neogrammarians (Neogrammarians) and structuralist (structuralism) investigations. Bussman(1998) is of the opinion that more recent investigations of dialect have been increasingly influenced by the sociolinguistic approach. In addition to the above definitions, the Generativists defined dialect as ‘any variety or variety feature not shared by all speakers of a language, whether or not use of such a feature correlates with
any non-linguistic factor; in cases where there is no such correlation, one may speak of randomly distributed dialects’.

Having looked at the various definitions of dialect, the question that still needs to be answered is –what is the distinction between language and dialect?

2.1.2 Language and dialect

The distinction between language and dialect poses a number of difficulties. The problem that arises is how a language can be distinguished from a dialect, and the related problem is how to decide what a language is. Chambers and Trudgill (1980:3) suggest that one way of looking at this is to say that ‘a language is a collection of mutual intelligible dialects’. This definition has the benefit of characterizing dialects as subparts of a language and of providing a criterion of distinguishing between one language and another. Petyt (1980) holds similar view. He opines that mutual intelligibility is the essential criterion to distinguish a language from a dialect. He stated that ‘dialects are different but mutually intelligible forms of speech’. He illustrated the examples of two speakers, inspite of some observable differences in their speech, can understand each other, they are held to be using different dialects. If two speakers cannot understand each other, they are speaking different languages. He further supported his view by giving the example of a Geordie, Cockney and a Cornishman who differ in speech, but can understand each other for the most part. Thus, they are said to be speakers of different dialects of the same language. But an Englishman and a German cannot understand each other, thus, they are said to be speakers of different languages.

Mutual intelligibility as a criterion to distinguish language and dialect is not entirely successful and there are difficulties with this criterion. According to Petyt, mutual intelligibility is not an all-or-none matter. There are degrees of comprehension between speakers. Petyt cited the example of a Lancastrian and a Yorkshireman who understand each other to a very great extent, while a Geordie and Cornishman may at times be in some difficulty. Moreover, the intelligibility may not be the same in each direction. For example a Geordie may understand a speaker of Standard English perfectly, whereas, the latter might frequently have problems in following the former. If such variations existed only within the same language, it might not be too damaging to
the criterion for distinguishing language and dialect. But this is not the case. For example, some North German dialects have a greater degree of mutual intelligibility with some Dutch dialects than they have with some South German dialects. Considering the example, it can be said that it is dialect continuum which raises several difficulties for the mutual intelligibility criterion. Dialect continuum refers to ‘a succession of geographically adjacent dialects, say A-B-C-D-E-F-G…, each mutually intelligible with its neighbours’ (Petyt 1980:14). This means that speakers of adjacent dialects, for example, D can be understood by C and E, and perhaps by B and F too, the extremes A and G are not mutually intelligible. The question that arises here is does lack of mutual intelligibility means that there is a difference of language rather than simply if dialect? If A can understand C, but cannot really understand D, does the language division come between C and D? But C and D and understand each other very well, so the mutual intelligibility criterion should classify these as different dialects rather than different languages. Such a dialect continuum extends from Northern France to Southern Italy; speakers of the various dialects can understand the neighbouring dialects, and others within a certain geographical proximity, but the dialects farthest from each other are certainly not mutually intelligible. The common example of dialect continuum in India is "Hindi". Hindi has several dialects. These dialects are spread over the entire Hindi speaking region usually termed as ‘Hindi Belt’ which includes the Indian states of Bihar, Chhattisgarh, Delhi, Haryana, Himachal Pradesh, Jharkhand, Madhya Pradesh Rajasthan, Uttar Pradesh and Uttarakhand.

In addition to the above linguistic criterion for distinguishing language and dialect, there are other non-linguistic factors- concerning common cultural or political allegiances, or the consciousness of the speakers. In the case of the Chinese, the Scandinavians, the Dutch and the Germans, and in similar situations, the fact that there are political units involved seems more important than questions of mutual intelligibility. Cantonese and Mandarin speakers are said to use the dialect of Chinese, partly because they both belong to the Chinese nation and share the Chinese cultural and literary heritage. On the other hand, the Norwegians and Danes, who can understand each other quite well, are said to speak different languages partly because they belong to different nations with different cultural centres and traditions. Further, in certain linguistic
situations, ‘when political factors are not significant, cultural considerations can lead to speakers being regarded as of different languages’ (Petyt 1980:16).

From the above discussion, it can be said that though the criterion of mutual intelligibility may have some relevance, it is not so useful in deciding what is a language and what is not a language.

2.2 Dialectology

Dialectology is a branch of sociolinguistics that focuses on understanding dialects. Crystal (2003:136) defines dialectology as ‘the systematic study of all forms of dialects, but especially regional dialects’. According to Crystal (ibid), dialectology is also known as ‘dialect geography’ or ‘linguistic geography’.

Dialect is divided into two basic types - regional and social dialect. Dialects which identify where a person is from are called ‘regional dialects’ though other terms are used, e.g., ‘local’, ‘territorial’ and ‘geographical’. Dialects which identify where a person is in terms of social scale are called ‘social dialects’ or ‘class dialects’. ‘Caste dialect’ which can be classed under social dialect is found only in India. Recently, the term ‘sociolect’ has also been used for social dialect. Some languages are highly stratified in terms of social divisions, such as class, professional status, age and gender, and here major differences in social dialect are apparent. The study of social dialects is called social dialectology. Social dialectology is the application of dialectological methods to the study of social structure focusing, on group membership as a determinant of dialectal competence (Crystal 2003).

Traditionally dialectology studies commenced in the nineteenth century and have taken the formed of detailed surveys using questionnaire and tape recorded interviews. Regionally distinctive words (distinct in form, sense or pronunciation), were collected, and collection of such words were plotted on the maps and compiled in a dialect atlas (or linguistic atlas). According to Crystal (ibid), if a number of distinctive items all emerged as belonging to a particular area, then this would be evidence for saying that a dialect existed. It was often possible to show where one dialect ended and the next began by plotting the use of such items (isogloss), drawing lines along their limits of use, and, where a bundle of such isoglosses fell together, postulating the evidence of dialect
boundary. On one side of the bundle of isoglosses, a large number of words, forms, senses and pronunciation would be used which were systematically different from the equivalent items used on the other side. Dialectological methods of this kind have more recently been supplemented by the methods of structural dialectology, tries to show the patterns or relationship which link set of forms from different dialects.

Local differences in speech have attracted attention for many centuries. But it was during the nineteenth century that the study of dialect received its greater impetus, from the development of comparative philology. The most important work in dialectology in modern times appeared in Germany. The first grammar which attempted to treat not just one dialect but all the dialects of an area was published in 1821 by Johann Andreas Schmeller. *The Dialects of Bavaria* gave a historical-geographical-grammatical presentation of the German language in this area and included a small map classifying the Bavarian dialects- probable, the first ‘*minilinguistic atlas*’. In between 1821 and 1876, there were many dialect surveys being carried out in Germany. In 1876, Georg Wenker, a young schoolteacher from Düsseldorf, began to work on a survey of the dialects in Düsseldorf to develop the ‘first great dialect survey’. He sent postal questionnaires out over Northern Germany. These postal questionnaires contained a list of sentences written in Standard German. These sentences were then transcribed into the local dialect, reflecting dialectal differences. Many studies proceeded from this, and over the next century, dialect studies were carried out all over the world. Some of the major dialect studies carried out are the Linguistic Atlas of New England (1931) by Jakob Jud and Paul Scheuermeyer (1931), the Atlas Linguistique de la France by Albert Dauzau (1939), Survey of English Dialects (1948), A Word Geography of England by Orton and Nathalia Wright (1974), The Linguistic Atlas of England by Orton, Stewart Sanderson and John Widdowson (1978), Word Maps by Clive Upton with Sanderson and Widdowson (1987), An Atlas of English Dialects by Upton and Widdowson (1996), Survey of Scottish Dialects by Angus McIntosh (1952), Handbook of the Linguistic Geography of New England published by Hans Kurath (1943).

With reference to Khasi language, the first survey on its major dialects was undertaken by Grierson(1903) in his *Linguistic Survey of India*. The Central Institute of Indian Languages, Mysore, had undertaken a project- *A Dialect Survey of Khasi* since
2006, to study the linguistic structure of the Khasi dialects. Bareh (2008) has done an extensive study on the linguistic structure of Pnar, one of the major dialects of Khasi.

2.3 Variation within Languages

The most significant variations or differences within languages occur at the level of the lexicon (vocabulary), phonology (pronunciation), grammar (morphology and syntax) and usage. Moreover, they are not just qualitative, in the sense that dialect A uses one feature and dialect B another, but they may also be quantitative, in the sense that dialect A uses one feature more often than dialect B does. Finally, variation may be regional, social or stylistic in its origins, and the methods that linguists have used to study each type differ slightly. We will now elaborate on these important concepts and provide examples.

2.3.1 Lexical variation

Differences in vocabulary are one aspect of dialect diversity which people notice readily and comment on quite frequently. They are certainly common enough as markers of the differences between geographical areas or regions--for instance the fact that "a carbonated soft drink" might be called pop in the inland North and the West of the United States, soda in the Northeast, tonic in Eastern New England, and cold drink, drink or dope in various parts of the South (Carver 1987:268). Or the fact that a person who was "tired" or exhausted" might describe themselves as being all in, if they were from the North or West, but wore out or give out if they were from the South (ibid.:273). With reference to Khasi language, some dialects of Khasi exhibit lexical variation with regard to certain lexical items. For example, for the word ‘go’, in Standard Khasi is ‘leit’, in Jowai is ‘lai’, in Nongstoin is ‘lie’ and in Nongpoh is ‘lai’. Accordingly, lexical differences play a significant role in regional dialectology (the study of regional dialects). Lexical differences are not as salient in distinguishing the speech of different social or socioeconomic classes, and they have accordingly played a much smaller role in social dialectology, which has concentrated instead on differences in phonology and grammar.

Many scholars have observed that different varieties of the same language often come with their lexical peculiarities. Some words may be restricted to a specific register, while others may have different meanings in different regions. Semantic variation is
the different meanings that particular words have from dialect to dialect, or the different words that are used for the same entity in different dialects. This might more accurately refer to as the study of lexical semantic variation.

2.3.2 Phonological variation

Phonological variation refers to differences in pronunciation within and across dialects. For example, people from New York and New England pronounce "greasy" with an s, while people from Virginia and further South might pronounce it with a z. An example from Khasi dialects is in the pronunciation of the word ‘shkor’ meaning ‘ear’. Other speakers of Khasi dialects pronounce ‘shkor’ with a ‘r’, while people from Laitkor and Laitlynkot pronounce the same word with a ‘l’ instead of ‘r’.

Phonological variants are fairly salient as markers of regional dialect. Rickford (2002:1)5 has given an example of the stereotypical Bostonian pronunciation of "Park your car in Harvard yard" as Pahk yo' car in Hahvahd yahd, which includes not only the r-lessness of Pahk, yo', Hahvahd and yahd (the r in car is retained because the following word begins with a vowel)--a feature shared with many other American dialects, particularly in the South-- but also the more distinctive use in these words of a long maximally Open or Open front vowel [a] where other dialects use a slightly fronter and less Open vowel [Å]. In order to represent the pronunciations with some precision, linguists often use a phonetic alphabet in which each distinguishably different sound is uniquely represented by a different symbol. It is to be mentioned here that phonological variation is central not only to regional variation but also to social variation and stylistic variation.

2.3.3 Grammatical variation

Grammatical variation involves two sub-types: morphology and syntax. Morphology refers to the structure or forms of words, including the morphemes or minimal units of meaning which comprise words, for instance the morphemes un- and happy in ‘unhappy’, or the morphemes cat and -s in ‘cats’. Syntax refers to the

structure of larger units like phrases and sentences, including rules for combining and relating words in sentences, for instance the rule that in English yes/no questions, auxiliaries must occur at the beginning of sentences, before the subject noun phrase (e.g. Can John go? versus the statement John can go).

According to Rickford6, there are many examples of regional variation of both types. For instance, the form of the past tense of catch, climb and draw was sometimes catched, clum and drawed respectively in parts of the Eastern United States but only caught, climbed and drew respectively in the Western United States.

Grammatical variation is much more common as a marker of social dialects and formal/informal styles than it is of regional dialects, with non-standard or vernacular variants sometimes being strongly stigmatized for their associations with limited education or use by the lower working class, but simultaneously being strongly admired and adopted for their connotations of informality, masculinity or non-pretentiousness. Whether positive or negative, grammatical variables tend to have strong social marking. One example at the level of morphology is the absence of third person present tense -s, as in "She like Ø liver". This feature is common in working class in Detroit and elsewhere in the US, but it is also common in other working class English varieties, for instance among English speakers in Norwich.

2.3.4 Language use/Speech events

As mentioned above, dialects can differ at the level of their words, sounds, and grammatical patterns. These are the three components of language that have been investigated in dialectology and linguistics for more than a century. Besides these three components, a fourth level, which has only begun to receive serious attention over the past thirty years, involves what is characterized, with deliberate vagueness, as language use (Rickford, 2002). Language use refers to a community's rules for constructing, participating in and (where relevant) evaluating verbal activities larger than the sentence, including narratives and telephone conversations and verbal routines like lecturing or telling jokes which are often described as "speech events." It also includes the variegated

---

aspects of language use which fall under the "ethnography of speaking," including conventions for speaking loudly, softly, much, a little, or not at all, whether addressees are to remain silent or vocally interactive during a speaker's turn, whether one is expected to broach or avoid certain topics and make extensive use of simile, metaphor and rhyme, and so on (Hymes 1973). Further, it also includes rules for turn-taking as well as the rules for conversational implicature, presupposition, and speech acts (events like commands, requests, promises and threats which are usually accomplished through the use of words) which fall within narrower definitions of "pragmatics".

Although different regions do have different conventions for language use, this is not something that has been systematically investigated by dialect geographers. Most of what is known about variation in language use has come from studies of different social groups, including men versus women and particularly, different national or ethnic groups.

Chambers and Trudgill (1980) stated that dialect difference is one of the most common observations that human beings make from day to day. Observations of dialect difference are so common that studying it became important. Dialect differences have also been observed in the Khasi language but not much work has been done in this particular area and whatever is available has not been studied extensively. Hence, there is a need to study Khasi dialects systematically. This research work focuses on some of the regional dialects of Khasi. Though variations within Khasi commonly occur at the level of the lexicon (vocabulary), phonology (pronunciation), and grammar (morphology and syntax), this research work is limited only to phonological and lexical variations.

2.4 The Phonology of Khasi7: Review of Earlier Works

In this section, an attempt has been made to discuss on some of the works done on the phonetics and phonology of Khasi. Although substantial works have been done on Khasi, the main focus, however, has been on the grammar of the language. Some of the works done in Khasi relating to the sound systems are Pryse (1855), Roberts (1891), Rabel (1961), Henderson (1967, 1976), Abbi (1979, 1987), Nagaraja (1985, 1989), Sten(1996) and War (2004). It is to be mentioned here that in each of these works, the emphasis has been more on segmental phonology and not much appears to have been

---

7 The term Khasi here refers to Standard Khasi
done in the areas of suprasegmental (prosodic) phonology. The discussions on some of the works are given below:

1. **An Introduction to the Khasia language: W.Pryse (1855)**

“The principal sources, from which the compile drew the little information which he has ventured to offer to the public in the following sheets were the following: conversing with the natives and analyzing the few books previously published in the dialect” (Pryse: p. iii). It is evident that few books, including portions of the Bible translated by Thomas Jones I (1941) were available to Pryse.

Pryse’s work is one of earliest works on Khasi language available to us. In his work, Pryse talks about the orthography, also on the consonants, vowels and diphthongs and a detailed discussion on the grammar of the language. Pryse’s missionary zeal extended to a valuable, though sketchy record the speech of the people with whom he interacted.

He observes that ‘the Khasi has no literature, no books, no marks to indicate sound, no standard by which either orthography or their pronunciation might be tested. The same words are pronounced in various ways in different villages. The variety of words used to express the same thoughts is also great’(ibid:v). This actually puts us to think about the phonology of Khasi which remains as an unsolvable problem today.

Pryse (ibid, 1) lists the Khasi alphabets. The alphabets are A B D E G H I J K L M N O P R S T U W Y. In addition to this, Pryse (ibid:1) lists the consonants and vowels separately. The consonants mentioned by him are [b, d, g, h, j, k, l, m, n, p, r, s, t].

Of the consonants mentioned by Pryse, [g] is a consonant which is not found in Khasi. It is to be mentioned here that there are also other consonants found in Khasi which Pryse failed to notice. Some of the consonants are ſ which corresponds to sound [ɲ] and ng which corresponds to sound [ŋ]. Thus, we can say that Pryse’s graphemes do not correspond with phonemes in Khasi.

Pryse mentions that ‘in Khasia word phlang ‘grass’, some of the natives pronounce the word as if written ‘flang’, and other sounds with an aspirated p only’(ibid,p 4). Pryse has clearly stated that [f] has been left out of the Khasi alphabet,
because it sounds very seldom, if ever occurs in Khasia’ (ibid, p 4). From this we note that: i) in Khasi there is a variation between [f] and [ph]. It is important to note here that [f] is not found in Khasi. This means that the introduction of [f] in Khasi is the result of the influence of English or may be neighboring languages. ii) Pryse has rightly identified the presence of aspirated sound in Khasi when he cited the example of phlang ‘grass’. While treating aspiration, Pryse made a sort of contrastive analysis of Khasi aspirated sounds with those of English sounds. He could not find aspiration in English language and as a result he had to resort to Bengali for his explanation. He observes that labials [b], [p], cerebral [t], guttaral [k] and dental [s] occur with aspiration. He has provided example for labial [p] from Khasi as in [phah] ‘to send’. He has not provided any example for voiced counterpart [bh] although it is available in the language. The cerebral\(^8\) [t] as in [thaw] ‘to create’ is provided and observes that the voiced counterpart [d] is never aspirated. This statement of Pryse is disputable for the reason that there are examples for voiced sound [d]. Infact, retroflex sounds are not found in Khasi. Furthermore, the sound which Pryse has noted as cerebral is not a cerebral (retroflex), it is phonetically, an alveolar sound[t] and a post alveolar [d]. The phonetic descriptions of these sounds are given in Chapter 3.2.1.

Pryse has identified seven vowels in Khasi. The vowels are [a, e, i, o, u, w, y]. He also mentions that some vowels in Khasi have both short and long counterparts. These vowels are a,á, e,é, i,í, and u,ú. The remaining four vowels o, w, y does not have any long counterparts.

It is to be noted here that barring [a, á], all the vowels do not show difference in length. Furthermore, the examples given for vowels with length are not available in any data that has come to my knowledge or notice. The example given by him for long [ú] as in [úm] meaning ‘water’ is to my knowledge is not correct. The vowel [u] as in [um] ‘water’ is a high, back, short, rounded vowel. The correct phonetic form is [um] ‘water’ in present day Khasi.

\(^8\) The term cerebral used by Pryse is equivalent to the term ‘retroflex’ which is commonly used today in the field of phonetics and phonology.
He further mentions that [o] ‘has generally the sound of the short English o……… . It never has the long open sound of o in English as in tone, alone etc’ (ibid, p 3). This observation of Pryse is in line with the analysis of this present study.

Pryse (ibid: 1) lists [w]and [y] as vowels. It is to be pointed here that phonetically in Khasi, ‘w’ and ‘y’ are neither vowels nor they are consonants. The preferred description for [w] and [y] are approximants. In the present analysis, they are treated as approximants.

Pryse listed seventeen ‘principal diphthongs’ in Khasi. He also mentions the presence of diphthongs or ‘junction of vowels’. These are [ ai, ei, oi, au, ae, ay, eu, ey, ew, ia, ie, io, iy, oe, ou, oy]. The analysis of diphthongs by Pryse is not different from other contemporary scholars or those in the past. Hence, this is taken up in the treatment of Khasi phonology in Chapter 3.3.4.

2. **A Grammar of the Khasi Language: H.Roberts (1891)**

According to Roberts, his work on *A Grammar of the Khasi Language* is based on the dialect of Cherrapunjee (Sohra). He states that ‘the Khasis have no written language of their own, and therefore no literature of any kind. There are no materials, from which to connect to their present with the past or to trace out a history for them’ (Roberts, 1891:.xvi). From this statement, it is evident that Khasi does not have a script of its own prior to the latter half of the nineteenth century. Thus, nothing can be said much about the history of the people as well as the language.

In his work, Roberts posited a question “What stage in linguistic development should we assign to the Khassi language?” (ibid:xvii). The answer given to this question is that ‘it is not so monosyllabic, like Chinese, and it ‘is too far removed from the so-called inflectional’. These observations of Roberts are true so far as the phonological and morphological characteristics of Khasi is concerned.

Roberts’s work contains a detailed description of the grammar of Khasi. He divided his work into three main parts, viz., (i) orthography, (ii) etymology and (iii) syntax.

In his discussion on the orthography of Khasi, Roberts also makes an attempt to describe the sound system in terms of the pedagogical structures available by him during
those times. He mentions that the alphabets in Khasi consist of 21 letters only. He divided these letters into three classes of sounds i) vowels, ii) semivowels and iii) consonants.

Roberts recognizes ‘twelve simple vowels’: five long vowels and seven short and we fail to understand the term ‘simple’ here. When he lists an equally long list of ‘thirteen proper diphthongs’, we realize that the term ‘simple’ is opposed to diphthongs. The ‘twelve simple vowels’ are as follows:

[a,á, e,é, i,í, o,ó, u,ú, y,ý ]

Roberts has identified twelve simple vowels: a,á, e,é, i,í, o,ó, u,ú, y (as in English word *fun*), y (as in French *lune*). For him, all the vowels have long counterparts. But it is to be mentioned here that barring [a,á], the other vowels do not show difference in quantity at least in the phonemic level. The examples given for vowels which Roberts has identified as long are not acceptable for native speakers of Khasi. The examples given for the short [u] as [lum] meaning ‘to gather’ and long [ú] as [lúm] meaning ‘hill’ are unacceptable. The difference between the two vowels is infact, not on the quantity but rather on the quality of the vowel. The vowel [u] as in [lum] ‘to gather’ is a high, back, rounded vowel, but the vowel [u] as in [lum] meaning ‘hill’ is a lower high, more centralized, less rounded vowel. The correct phonetic forms are [lum] ‘to gather’ and [lʊːm] ‘hill.

The ‘thirteen proper diphthongs’ mentioned by Roberts are [ ai,ái, aw,áw,ei, ew, iw, íw, ie, oi, ói, ui úi]. Diphthongs such as aw,áw, iw, íw, ew as mentioned by Roberts are phonetically not diphthongs. In the articulation of such sounds, the second member [w] is more of a semi-consonant than a vowel. The detailed description and analysis of diphthongs is discussed in Chapter 3.2.3 where explanation is given as to why the present researcher differs in her treatment of diphthongs in Khasi.

Roberts identifies two semivowels in Khasi: [i] and [w]. These semivowels phonetically are realized as palatal approximant [j] and biblabial approximant [w]. The confusion created by Roberts is that he makes use of the phonetic symbol [i] both of high front unrounded vowel and also palatal semivowel. The nomenclature problem can be solved, but the phonetic problem remains.

Roberts has also identified the presence of vowel quantity in Khasi. Interestingly, he uses an acute accent [´]to mark a vowel which is long, leaving short vowel as
unmarked. Occasionally, he uses grave accent [´] to mark shortness. It is to be noted that the method adopted by him is not acceptable to present day phonetic analysis as it creates more confusion in the minds of the reader. The notation use for vowel with length is the IPA [:] immediately after the vowel segment.

Regarding the consonant sounds, Roberts has mentioned fourteen of them. But in his work, he has not listed the consonantal sounds separately as he has done in case of vowels. What could be guessed is that he has included all the consonants in orthography, which he carefully lists. This chart includes 14 consonants (ibid:1) as listed below:

\[ b, k, d, g, ng, h, j, l, m, n, p, r, s, t, \]

He has further listed a set of nine aspirated consonants \[ bh, kh, dh, jh, ph, rh, ngh, th, sh \].

The aspirated letters listed by him are bh, kh, dh, jh, ph, rh, ngh, th, sh. Except for rh, ngh, and sh, all the aspirated letters mentioned by him are aspirated sounds in Khasi. The examples given by him for [sh] is [shong] which means ‘to stay’, [rh] is [rhem] meaning ‘warm’ and [ngh] is [rynhang] meaning ‘ajar’. But it is to be noted here that [sh] is phonetically a \[ʃ\], [rh] is a cluster of two consonants occurring within a syllable and [ngh] is just a sequence of consonants and they fall under separate syllables.

He (ibid:5) remarks that ‘the final \(h\) when preceded by a vowel, its power is that of a slight guttural, as [rah] ‘to bear’; [pah] ‘to utter a sound’; [lah] ‘to be able’. It is interesting to note here that Roberts has identified the presence of a guttural sound. What Roberts has identified as a guttural sound is phonetically a glottal stop, a sound common in Monkhmer languages, which occurrence is very frequent in word final position.

It is to be noted that Roberts’s work is one of the good introductory works on the sounds of Khasi. It is to be pointed here that as early as 1891, with limited sources available to him and with undeveloped language studies of those times, he made every effort to bring out the sound structure of Khasi and his attempt is laudable. The influence of Roberts’s work can be seen in the work of later scholars who also knowing or unknowingly followed him by suggesting more or less similar phonetic statements. With drawbacks in his work, Roberts has paved the way for further researches in Khasi studies.
3. **Khasi, a Language of Assam: Rabel (1961)**

In her work on Khasi, Rabel (1961) attempted to provide an extensive analysis of Khasi.

Under the section titled 'Phonology', Rabel (1961:2) lists 21 consonantal phonemes in Khasi /p, pʰ, b, t, tʰ, d, k, kʰ, ?, j, s, ʃ, h, m, n, ɬ, ɮ, l, r, w, j/. Unlike other linguists working on Khasi phonology, Rabel excludes the inclusion of [bh, dh, jh] at the phonemic level. This present research is in accordance with Rabel’s analysis. There are many reasons for the exclusion of these sounds at the phonemic level. The detailed discussion of these sounds in this present work is given in **Chapter 3.1.3**.

According to Rabel(ibid:15), the minimum constituent of a syllable is CV (consonant - vowel). But in this research work, it has been found that the minimal constituent of a syllable can also be VC, though, CV is more common than VC.

Rabel (ibid: 15) points out that there is an occurrence of a triple consonant cluster [spdok] 'short and stout' which is in free variation with [suupdok]. This is incorrect. In the present analysis, it has been found out that three members consonant cluster is not permissible in Khasi.

Rabel (ibid:17) mentions the occurrence of /s/ and /l/ in the final position of a word. However, according to the data of this study, /s/ and /l/ cannot occur in the final position of a native word. Occurrence of /s/ and /l/ in the final position is seen only in borrowed words / loans words as in skul 'school' and klas 'class'.

Rabel (ibid:13) lists five phonemic monophthongs /i, e, u, o, a/. All these monophthongs have length counterparts. Rabel (ibid) has given a detailed description of Khasi vowels along with phonetic transcriptions and a chart showing vowels distributions. Henderson(1967) commented that Rabel has noted many fine distinctions of vowel quality and vowel length in Khasi. Henderson (ibid) further mentions that her analysis of Khasi vowels is very close to Rabel’s, though there are minor variations. It is to be noted that Rabel’s work on Khasi vowels, is accepted by most Khasi linguists and scholars. The analysis of Khasi vowels presented in this present research is similar to Rabel’s, though it differs only in the treatment of length and diphthongs.
Rabel also makes a distinction between minor and major syllables. A syllable having a sonant as peak is called a 'minor syllable' and a syllable having a vowel as peak is called a 'major syllable'. Rabel (ibid:17) states that minor syllables can only occur in combination with major syllables i.e., Cs - may constitute the first syllable in a two syllable word. Two minor syllables may succeed each other as first and second syllables in a three syllable word. She gave the following examples:

[hn - nin] 'yesterday' minor + major syllable
[pn - kn-ma:u] 'to remind' minor+ minor + major syllable

This observation made by Rabel is true and is in accordance with the phonological feature of Monkhmer languages, Khasi, being one of them. It is to be noted that in current phonology, for syllable structure having a composition of minor + major syllable, the term sesquisyllable is used.

There are many instances in Rabel's work where she has taken sonants as vocalic. A few examples need to be mentioned:

[plleŋ] 'egg'
[trk^aŋ] 'fern'

According to Rabel, all sonants except /ɲ/ may serve as vocalic. From the above examples, it has been found that sonants, in Khasi, can function as vocalic, but they are phonologically conditioned (Refer Chapter 3.4.1.1).


Henderson(1967: 564) observes that ‘one of the most striking characteristics of languages of the Austro-Asiatic family such as Mon, Khmer and Vietnamese is the great variety of vowel qualities that some of them differ from each other slightly that it is hard for the observer to believe that they can be functionally distinct in practice’. Khasi, an Austro-Asiatic language belonging to the Monkhmer branch is not an exception to this. Henderson pointed out that the problem of vowel quantity of Khasi vowels has perplexed ever since Schmidt wrote his paper on Khasi vowels in 1903,1904. Henderson (ibid) noted that in her analysis of Khasi vowels, she found out some of the confusion that has arisen may be ascribed to the fact that what was interpreted by Roberts (1891) and
Schmidt (1903) as a difference in the final consonant with an accompanying difference in vowel quantity, was in her informant’s speech a difference in vowel quantity only.

Henderson (ibid) has attempted to throw some light on Khasi vowels by comparing the analysis given by scholars like Roberts (1891), Schmidt (1904) and Rabel (1961) with her informant’s speech. She mentions that in order to present a clear description and distribution of Khasi vowels, there are a number of conditioning factors that have to be taken into consideration. The factors listed by her are a) lexico-semantic, b) syntactic, c) grammatical, d) phonological and e) orthographic (ibid:565). According to her, under phonological consideration, the most important factor is the syllabic context of vowels and diphthongs. She provided the description of Khasi vowels under phonological considerations by taking six contexts into account. These are i) vowels in Open syllables, ii) diphthongs in Open syllables, iii) vowels in syllables with final [ʔ], iv) vowels and diphthongs before final nasal consonants, v) vowels before final stops except [ʔ], and vi) vowels before final [r].

Her observations of Khasi vowels are as follows:

i) Contexts (i), (ii), (iii): Vowels and diphthongs in Open syllables and before glottal stop.

Henderson stated that what is marked by Roberts in these contexts as a difference in quantity corresponds to a difference in quality in some of her informant’s pronunciation. She cited the example of Roberts for the vowel [e] in [ne] ‘or’. Her informant’s pronunciation for the same word given by Roberts is [neˑ] ‘or’. From the data, it can be observed that the difference in the representations given by Roberts and Henderson is in vowel quantity. Roberts has a short vowel [e], whereas, Henderson has a half-long vowel [eˑ]. In this present analysis, all vowels in the open syllable are found to be half-long. Thus, the representation given by Henderson is in line with the present researcher.

According to Henderson, Schmidt mentioned that the principal vowels [a], [i], and [u] are generally short finally in open syllables. He further mentioned that there is an exception to this, and the exception is the occurrences of these vowels in the following words:
Henderson stated that there is no significant difference in vowel length in her informant’s pronunciation. She further mentioned that Schmidt’s [i] in the word [di] ‘to sell’ corresponds to her informant’s [eˑ]. This reflects a difference in vowel quality.

The analysis presented in the present research differs from Schmidt’s observation. In this present analysis, vowels in open syllables are treated as half–long, and it is marked by diacritic ‘ˑ’ as in words [kbaˑ] ‘paddy’, [jriˑŋ] ‘rubber’, [deˑ] ‘sell’.

Another example which Henderson has given based on Robert’s work is a vowel before a glottal stop. The word is ‘líh’ ‘white’. It is to be noted here that the ‘h’ in the word given by Roberts is phonetically transcribed as [ʔ] and the i in ‘líh’ meaning ‘white’ represents a long vowel [iː]. The phonetic representation of Henderson informant’s for the word is [leˑʔ]. For the word ‘white’, the present analysis differs from both Roberts and Henderson. For Roberts, the vowel in the word is a long vowel [iː] and for Henderson, it is a half-long [eˑ]. Their representations as mentioned show the difference of vowel, not only in the quality but also in quantity. In this present analysis, it has been found that the vowel [e] before a glottal stop will always be a ‘full long vowel’, irrespective of the environments, i.e., VC or CVC. Based on this finding, the phonetic form of the word ‘white’ is [leːʔ].

Henderson (1967) mentions that Schmidt’s (1904) is of the view that [a] and [o] in Khasi are always short before [ʔ], [i] and [u] can be both long and short, and [e] before [ʔ] is always long. The examples cited by Henderson as given by Schmidt for short [i] before glottal stop are ‘tih’ ‘dig’, mih ‘ksih’ ‘otter’ and ‘khih’ ‘move’ and for the long vowel, the examples are ‘íh ‘leave’, ríh ‘hide’ and sníh ‘skin’. Based on these examples, Henderson observes that Schmidt’s short vowel [i] correspends to [iˑ] and his long vowel corresponds to [eˑ] in her informant’s speech.
Further Henderson cited the examples from Schmidt for short [u] before [ʔ]. These examples are kuh ‘crop’, śuh ‘away’ and pyduh ‘rotten’. Henderson’s observation on Schmidt’s [u] is that there is a problem in indentifying the forms used by her informant. For the word kuh, her informant pronunciation is [khuˑ], and for the word pyduh it is [pdʊˑʔ]. Henderson stated that her informant’s data for the word [pdʊˑʔ] ‘to strike’ appears not to be relevant comparison with Schmidt’s pyduh.

For long vowel u, Schmidt provided these examples buh ‘to put’, duh ‘to lose’, ruh ‘cage’. Henderson observes that these words are pronounced with a vowel which should be transcribed as [ʊ]. The quality of the vowel in these words is rather more centralized than the closed [o] found in open syllable.

With reference to long vowel e, Schmidt provided these examples; eh ‘hard’, heh ‘big’ kseh ‘pine’, theh ‘pour’ and teh ‘bind’. Henderson observes that the vowel in all the above words is [ɛʔ] in her informant’s speech.

ii) Contexts (iv) and (v): Vowels and diphthongs before final nasal and stop consonants.

Henderson (ibid) observes that Roberts’s í and ú before nasals corresponds to centering diphthongs, iˑə and uˑə respectively in her informant’s pronunciation. She supported her analysis by presenting the acoustic analysis of [iˑə] in the word [diˑəŋ] meaning ‘tree’. She further gives the example of the word [deŋkseːʔ] meaning ‘pine tree’. Considering the acoustic analysis of the word ‘pine tree’, she notes that there is a change in duration and formant structure of the vocalic element [iˑə], [iˑə] in the word ‘pine tree’ is perceived as shortening and monopthongization accompanied by change of quality from [iˑə] to [e].

With reference to vowels and diphthongs before final nasal consonants, Henderson views that Schmidt’s rule is that before [-ń] (probably a velar nasal), [i] occurs both long and short, [a] and [u] only short. With other nasals, [a], [i] and [u] occur both long and short and [e] and [o] are always short.
According to Henderson, Roberts (ibid) noted that diphthongs ai, ái, au, áu, ei, úi can occur before final t, d, or n. She cited the examples given by him in words like kait ‘plantain’, pait ‘break’, leit ‘go’, buit ‘skill’, ngeit ‘believe’, and tuid ‘flow’. Considering the examples given by him, Henderson (ibid, 569) remarks that Roberts ‘appears to have taken the spellings –it, -id, and –in at their face value’. She pointed out that the Khasi spellings for final palatal consonants are written as –in [ɲ] and -it, -id [c]. Henderson (ibid, 575) notes that ‘Schmidt had no means of checking at first hand but his knowledge of cognates in other languages enabled him to recognize that he was here dealing with the reflexes of original final palatals with an i-like on glide, and he consequently declined to regard the preceding sequence of vowel plus on-glide as a diphthong as Roberts had done’. Rabel (1961) also observes that these spellings represent the final palatal consonants [c and -ɲ]. Henderson agrees with Schmidt (1904) and Rabel (ibid) in treating the final sound as palatal consonants based on the analysis of her informant’s speech.

With reference to this context, Henderson (1967:575) states that ‘the whole issue of the length of vowels occurring before final stops was, however, clouded for Schmidt by yet another Khasi spelling convention, presumably, ascribable to the link between vowel length and voicing of final consonants in Welsh, whereby final [b] and [d] are regularly used to indicate a preceding long vowel, final [p] and [t] the corresponding short vowel’. In the absence of evidence, Schmidt had no alternative but to accept these spellings at their face value and to postulate both voiced and voiceless final stops for Khasi. Schmidt, however, pointed out that this dependence of vowel quantity upon the nature of the final consonant must be regarded as Khasi peculiarity. His rule is that final voiced plosives occur only after long vowels, final voiceless only after short vowels. Schmidt’s view on the quantity of vowel and the occurrence of final voiceless and voiced stops needs to be carefully examined.

vi) Vowels before final [-r].

According to Henderson, Schmidt recognizes vowels [a], [i] and [u] as occurring both with long and short before [-r], and vowels [e] and [o] as a general rule always occur with length before [-r]. The exception to this rule is ‘a few onomatopoeic adjectives’ as in
words [her]‘sharp’, [ker ker] ‘dangling’, [ter ter] ‘one after the other’ etc., Based on
her own analysis, Henderson agrees with Schmidt in this context. In the present study, the
description of vowels and their allophononic realizations are dealt with in detail (Refer
Chapter 3.2.1).

With reference to Rabel’s observations of Khasi vowels, Henderson(ibid,578)
remarks that ‘Rabel notes many fine distinctions of vowel quality and length which are
always referred to the phonetic context in which they occur’. Henderson clearly
mentioned that her informant’s speech is similar to the analysis presented by Rabel,
though there are minor variations. Some of the variations between the analysis of Rabel’s
and Henderson’s are in words ‘sdie’ meaning axe, and particles ‘mo’ and ‘to’. Rabel
represented ‘sdie’ as [sdii], and the particles ‘mo’ and ‘to’ as [mo] and [to] respectively,
whereas Henderson’s represented ‘sdie’ as [sde] and ‘mo’ and ‘to’ as [mɔˑ] and [tɔˑ]
respectively. It is to be noted here that Henderson’s representation is similar to the
analysis presented in the present study. In addition, Henderson(ibid, 580) remarks that
her ‘reservations about Rabel’s phonemic solution do not arise from her phonemic
findings or from objections to her methodological approach or to logical processes by
which the solution is reached, but from the fear that is likely to prove misleading
to many scholars and readers’.

5. The Sound System of Khasi (1979), Palatals or Lamino Dentals? A Probe
into a Feature Theory (1987): Anvita Abbi

Abbi’s article, The Sound System of Khasi (1979), can be said to be a first work
of its kind. She is the first one to look at the sound system of Khasi from a generative
approach. Earlier works on Khasi phonology are mostly descriptive. Her paper is divided
into three parts: Part I consists of the sound system and representations of phonological
rules of Khasi, part II deals with the discussion of completely new sound in Khasi, and
part III deals with the probable areas of difficulty in learning Hindi sounds for Khasi
speakers. Abbi(1979: 221) mentions that there are 24 phonemic consonants in Khasi.
She stated that not all stops are prevalent in all positions. She clearly mentions that it is
difficult to distinguish between voiceless and voiced stops in the final position. She notes
that [g] and [c] consonants are not found in the language. She is of the view that in
speech of some speakers, there exist a variation between dentals and alveolars, while there are others who varied between alveolars and retroflexes. Abbi (ibid: 222) cites the examples of words with retroflexes and alveolars. Some of the examples given by her relating to the presence of retroflex sounds are [dlpey] ‘ashes’, [kʰʌndə] ‘earth’.

According to Abbi, Khasi exhibits the presence of a ‘completely novel sound’, which is being ignored by many linguists (Rabel, Henderson) who have worked on the phonology of Khasi. She described this sound as fronto dentals (1979) and as ‘lamino-dentals’ (1987). According to Abbi (1987: 102), lamino-dentals are very interesting and intriguing sounds as they can neither fall into the category of co-articulation nor can be specified as unitary articulation. According to her, this sound is produced when ‘the front of the tongue is pressed against the Lower teeth, in the process of which the body of the tongue rises high towards the hard palate, though not high enough to produce a pure palatal sound [c]’ (Abbi,1979:228). Some of the examples of fronto dentals/ lamino dentals given by her are:

[tʰɪllet] ‘tongue’
[toyt] ‘flow’
[blet] ‘mad/dull’

Abbi’s identification of Khasi consonants is in accordance with the analysis of this present research work, though, there are which are not borne out by this study. Abbi noted that [g] and [c] are not found in Khasi. In this present analysis, it has been found that [c] is present in Khasi, but only at the phonetic level. Its occurrence is predictable. It has been treated as an allophone of /ʃ/. The absence of [g] in Khasi has been traced taken into account historical evidences. The discussion of [c] and [g] is given in Chapter 3.1.3.

Another important point which the present researcher differs from Abbi’s observation is regarding the presence of retroflex sound in the language. The presence of retroflex sound in Monkhmer languages is not a common phonological feature. The examples given by her for retroflex sounds are incorrect. For example [d] in [dlpey] ‘ash’ which Abbi described as a retroflex is in fact, a post alveolar [d]. In the present analysis, [d] phonetically is being described as a post alveolar (not even alveolar) sound.
The articulation of it, is not the same as [t]. It is to be mentioned here that the examples provided by Abbi to support the presence of retroflex sound can be a result of idiolect variation.

In her description of Khasi vowels, Abbi (ibid: 224) lists seven phonemic vowels. Out of these /ə/ is one of them. But in the present analysis, the central vowel /ə/ is not found to be present either phonetically or phonemically. One of the examples provided by her to illustrate the presence of [ə] is [pʰəŋ]‘grass’. It is to be noted here that the correct phonetic representation of this word which is [pʰaŋ].

Abbi in her work did not discuss the diphthongs in detail. She listed only the phonetic diphthongs and these diphthongs are aU, eI, oU, Iu, eU and Ia. With the exception of /ia/, in this present analysis, all the others are treated as –w and –j endings. (Refer Chapter 3.2.3)

The claim made by Abbi regarding the presence of Lamino-dentals in Khasi is debatable. The analysis in the present research work for such sounds is in accordance with the analysis given by Rabel and Henderson. Rabel and Henderson posit a [c] in all those words which Abbi has as voiceless or voiced lamino-dentals.


Nagaraja(1985) gives a brief outline of Khasi phonology. The detailed description of Khasi phonetics and phonology is given in his work in 1989. Under the section phonemic inventory, Nagaraja(1989:25) lists 11 phonemic vowels in Khasi- /i i:, e e:, a a:, o o:, u u:, ɨ/ and 24 consonantal phonemes. These are /p, pʰ, b, bʰ, t, tʰ, d, dʰ, k, kʰ, ʔ, j, jʰ, s, ʃ, h, m, n, n̥Josh, l, r, w, y/. The phonemes presented by Nagaraja are in line with the analysis presented in this present study. However, there is a difference in the treatment of vowel length. In the present analysis, with the exception of /aː/, other vowels with length are treated as an allophonic variants of the short vowels. The treatment of consonantal phonemes is similar
with the exception of voiced aspirated stops [b̩h, d̩h, j̩h]. The detailed discussion of the voiced aspirated stops in this analysis is given in Chapter 3.1.2.

Nagaraja (ibid: 26,27) lists certain contrastive pairs to illustrate vowel phonemes. The examples for /e/ and /e:/ are in words /em/ ‘well to do’ and /e:m/ ‘no’ and /u/ and /u:/ as in words /lun/ ‘tadpole’ and /lu:n/ ‘throw’. The above contrastive pairs to illustrate the contrast in length, between /e/ and /e:/ and /u/ and /u:/ are found to be incorrect. The correct forms are /ɛm/ ‘well to do’, /ɛ:m/ ‘no’ and /lʊn/ ‘tadpole’ and /lʊn/ ‘throw’.

In his description of the vowel system of Khasi, Nagaraja adds the central vowel /ɨ/ to the list of vowels already present in the language. The presence of central vowel /ɨ/ has never been mentioned in the earlier work done by Rabel (1961). Nagaraja has also pointed out that /ɨ/ can occur in the initial as in word /ɨm/ 'to live' (ibid: 27). This is found to be incorrect. The correct form which is accepted to native speakers of Khasi is [im] 'live'. In the data given by Nagaraja, it has been found out that the central /ɨ/ occurs frequently when followed by /m, n, ŋ, r, l/ in disyllabic words. It is to be noted that in this present study, the central vowel [ɨ] is treated as an allophone of /i/ as the occurrence of this vowel is predictable (Refer Chapter 3.2.1).

One of the most striking phonological features of Khasi is its richness in initial consonant clusters. According to Nagaraja, consonant clusters in Khasi occur mostly in the initial position. Nagaraja (1989: 31) mentions that two member clusters are more frequent than three member clusters. In the final position, occurrence of consonant clusters is very rare, though he pointed out that there is a possibility where two consonants can occur. This statement pointed out by Nagaraja needs to be further clarified.

Nagaraja (1989: 36) also gives examples where three member clusters are found in the initial position. These are:

- pdy- [pdyaŋ] ‘receive’
- bny- [bnyat] ‘tooth’
- kʰry- [khryat] ‘cold’
- kʰɲy- [kʰɲyaŋ] ‘insect’
From the above examples, it is observed that in words with three consonant clusters, the third mostly is a frictionless continuant /j/. This particular observation of three member cluster by Nagaraja differs with the present analysis. In the present study, the third member should be a vowel /i/ or specifically diphthong /ia/ in case of the above examples. Thus, it can be said that in Khasi, only to two consonant clusters is permitted in the initial position.

Regarding, the occurrence of consonant clusters in the final position, Nagaraja is of the viewed that two consonants can occur. He (ibid: 48) provided the following examples:

- yt [payt] 'to break' [payd] 'crowd'
  [dayt] 'to bite'

The presence of two consonant clusters in the final position is not accepted. As per the data obtained by me, and as per my intuitions as a native speaker, it is to be pointed here that in Khasi, consonant clusters in the final position do not exist at all. In fact, in the above words, there should be a vowel /i/ instead of the frictionless continuant /j/.

7. **Shaphang Ka Ktien Khasi (About the Khasi Language) : H.W.Sten (1996)**

In his book *Shaphang Ka Ktien Khasi*, Sten (1996) provides an outline of Khasi literature and language, the Khasi writing system or the Khasi alphabets, the Khasi spellings, word juncture, word formation, word level categories, phrases and clauses, and on functional structure.

Under the section of Khasi Alphabets, ‘Ki Dak Thoh’, Sten discusses the major classification of sounds in Khasi- vowels and consonants. According to Sten, in the writing system, six vowels are used. These vowels are: a, e, i, o, u, y. He also mentions that besides y, other vowels have alternative realizations. He gives the example of a which has three realizations. The examples cited by him (ibid:19) are:

a as in ai ‘give’
á as in wád ‘search’

à as in jāka ‘place’

He further explains that á in Khasi is similar to phonetics [ʌ] and that á is long vowel [a:].

Sten (ibid:20) has given the description of vowels based on the way these vowels are articulated i.e., the position of the tongue including the height and body use in the production of the vowel and also the position of the lips. He mentions that front vowels like [i], [í], [a], [á], [e], [é] are produced with the front of the tongue raised towards the hard palate and back vowels like [o], [ó], [u], [ú] are produced with the back of the tongue and in the articulation of such sounds the tongue is moving backward. It is to be noted here that Sten (ibid, 20) did not give any examples for front vowels in Khasi but he did provide examples for back vowels. The words with back vowels cited by him are [súm] ‘spear’, [tlór] ‘fell’ etc. He further states [y] is a central vowel in Khasi because in the production of it, the body of the tongue use is in between front and back. He adds that phoneticians have also added another criterion for description of the above vowels depending on the height of the tongue i.e., the raising of the tongue from its normal position. According to him, in the production of vowels like [i], [y], [u], the tongue is raised from its normal position and the lips are slightly narrowed so as to prevent the movement of the tongue. In English, he said these vowels are referred to as close vowels.

Vowels [e] and [o] are produced with the tongue slightly lowered than the close vowels; hence they are called half close vowels. Two vowels [é] and [ó] are produced with the lips slightly open and the tongue is lowered than the half close vowels. These vowels are called half open vowels. [a] and [á] are produced with the lips open and the tongue in the Lowest position. These vowels are called Open vowels. It is to be pointed here that Sten did not provide examples of words with close, half close, half Open or Open in Khasi. He further states that the position of the lips is also very important in the description of vowel sounds. Vowels according to him should be described in terms of the lips position- rounding of the lips orifice ‘pynpyllon’ and spreading of the lips ‘pynmadan’. Vowels like [i], [í], [a], [á], [e], [é], [y] are produced with spread lips, hence, they are called spread vowels. The examples provided by him for spread vowels in Khasi are
words like [pin], [phér], [hér], [tat], [pyntrei]. [o] and [ó] are produced with the rounding of the lips orifice, hence they are called round vowels. Unlike spread vowels, no examples are provided for round vowels in Khasi.

The description of Khasi vowels provided by Sten is in terms of phonetic description. It is to be noted that the detailed articulatory phonetic description of the vowels have not been taken into account. Another drawback of his description of vowels is in the use of diacritics for long vowels. His use of diacritics is based on Roberts (1891) which is not acceptable in present day linguistics. He correctly identified the presence of central vowel in the language. But the symbol used by him is incorrect. [ɨ] should be used instead of ‘y’. Another observation of his description is that he did not provide examples for some of the vowels present in the language. Examples are important to support the claim of the presence of the sounds in any language.

In addition to the above mentioned vowels, Sten also has identified that in Khasi there are vowels in Khasi which occur in sequences in a word. These vowels are called diphthongs. Sten (ibid, P.22-23) lists the following diphthongs in Khasi:

[ie], [ia], [io], [iu], [ei], [ai], [oi], [ui], [eu], [au], and [ou].

The diphthongs listed by Sten are too many in number. It has been observed that detailed articulatory and acoustic analysis has not been carried out. In this present work, the number of Khasi diphthong is reduced only to one, and that is /ia/. This has been done taken into consideration the acoustical analysis. For detailed discussion of Khasi diphthongs refer Chapter 3.2.3.

Sten (ibid: 23) mentions the presence of triphthongs in Khasi. I quote

“da kaba niew ia u w uba wan khatduh ha ka kyntien kum u vowel ngi lah ban iohi ha ka Khasi don sa ki triphthong : kata lai tyllli ki vowel ha ka juh ka kyntien: briew, thliew siew kliaw, siaw (ibid: 23). This can be translated as :

“by considering w as a vowel which occurs in the final position of a word, we can say that Khasi has triphthongs: ie., three vowels in the same word. Examples given by him are briew ‘man’, thliew ‘pit’, siew ‘to pay’, siaw ‘to whistle’.”
In the present analysis, no triphthong is found. Another observation of Sten’s work is that though Sten’s tries to provide a phonetic description of the Khasi vowels, yet he fails to give a phonetic representation of the sounds. Without providing phonetic transcription, it is difficult to point whether the words cited by him (pg 23) have a [w] or [u].

Sten also discusses about the consonantal sounds in Khasi. He further states that these consonants used in orthography can be classified into different groups based on their characteristics. The term referred for the description of consonants is in English, a language understood by all. He has given a brief description for these consonants in Section 2.3 of his book.

Describing sounds in any language (for example Khasi) based on sounds of another language (English) presents major problems. For example, Sten describes [t] as an alveolar sound, which is true in English. But in Khasi [t] is more dental. Its articulation is different from the English [t]. Thus, it can be said that his description poses problems to a language learner or to a reader.


*Ki Sawa Bad Ki Dur Kyntien Jong ka Ktien Khasi* is a book written in Khasi on the Khasi language by War(2004). In this book, War (2004) talks about linguistics as a discipline- it is more a science than an art. She also discusses about the formal and functional aspects of language. The formal aspect of language deals with the structure of a language and the functional aspect deal with the functions or uses of a language. She discusses the linguistic classification of Khasi- the genealogical classification as well as the typological classification. In chapter IV and V, a discussion on the sound systems of Khasi is provided. The detail discussion on the sound systems by War is given in the following paragraph. Chapter VI is the discussion on the phonetic alphabets/symbols and the process of assimilation in Khasi. Chapter VII is about morphology of Khasi. The discussion under morphology includes affix and clitics, compound words, reduplication etc. Chapter IX is on borrowed words in Khasi and Chapter X deals with coining of new terms and akin words. According to War (2004), one of the features which distinct Khasi from other languages is the used of akin words. Akin words play a very important
role not only at the morphological or lexical levels but also at the semantic level. Meanings which are context bound can be expressed through the use of akin words. Thus, War rightly said that this area needs to be studied extensively.

In chapter IV, War (2004) discusses in detail about the consonantal system of Khasi. She describes them in terms of their places and manners of articulation, and also whether they are voiced ‘khia’ or voiceless ‘sting’. According to War (2004:34-35) there are 27 consonants in Khasi. These are /p, b^h, t, d, t^h, k, , c, k^h, , dz, dz^h, s, f, h, m, n, n^h, l, , r, , w, j/.

She mentions (ibid:34) that some of these consonants (i.e., /p, b, t, k, s, f, m, n, n^h, l, r/) can occur in the initial, medial and final position of a word. There are also some consonants which can occur only in the initial (i.e., /dzh, lh/) and others only in the final position of a word (i.e., /c/).

According to War (op.cit), there are 27 phonemic consonants in Khasi. Out of these consonantal sounds, aspirated sounds like /b^h, d^h, dz^h, l^h, r^h, / are treated as one phonemic unit. The analysis presented in this research work differs with War in the treatment of these sounds as one phonemic unit. In this research work, the above aspirated sounds are analyzed as combination of consonants (consonant cluster) rather than one phonemic unit. The detailed discussion of these sounds is given in Chapter 3.2.3.

Considering the occurrences of [c] in the final position in Khasi, in this analysis it is treated as an allophone of /Ʉ/ and not as a separate phoneme. War has represented this sound as a phoneme. Another observation is that she did not provide minimal pairs/contrastive words to establish the phonemic status of this sound.

War (ibid: 35) mentions that /b/ and /d/ can occur in the final position of a word. The examples provided are ‘tab’ and ‘sad’ ‘comb’. In the analysis of these voiced stops in Khasi, their occurrence are found only in the initial and medial position of a word. In the final position, only the voiceless counterparts /p/ and /t/ can occur. One of the evidences to support their absence in the final position is by taking words in English with these consonants in the final position of a word. It has been found that most of the Khasi
speakers were unable to pronounce the sound in the final position correctly. Where these sounds occur, it has been found that they produced it with a voiceless sound.

War (ibid) discusses about the phonotactics of Khasi. Here she talks about consonant cluster, consonant sequences, and consonant cluster in terms of their distributional restriction and syllable types. It is to be mentioned here that her discussion on the phonotactics of the language is in accordance with the analysis represented in this research work.

In Chapter V, War discusses the vowels in Khasi. In her discussion of Khasi vowels, she includes the description of the vowels as given by Roberts (1891), Schmidt (1904), Rabel (1961), and Henderson (1967). War (ibid, p48) listed the vowels present in Khasi. These vowels are /i, I, e, ε, ɔ, ɔ:/, aː, o, u/. She explains that /I/ in Khasi is a vowel which is shorter than /i/ as in words /Im/ ‘live’, /tIp/ ‘know’ etc. According to War (ibid:49), Khasi has 13 diphthongs- /-ia, -eu, εu, -au, -aːu, -ɛi, -aːi, -ɔi, -ɔːi, -iu, -ui/. Besides diphthongs, she also mentions the presence of triphthong in Khasi. The sound is /-iau/ in words like ‘siaw’ meaning ‘to whistle’ and ‘duration’ meaning ‘sea’. She further notes that there are triphthongs which are formed when two morphemes are joined together. The examples cited are /u/ + ɛi/ becomes /uɛi/ ‘Mas+Q.mkr’ and /ki/ + /ɛi/ becomes /kiei/ ‘Pl + Q.mkr’.

According to War, there are eleven phonemic vowels in Khasi. But in this analysis, the phonemic vowels are seven in number. In the present analysis, contrastive pairs of the vowels are provided to show their phonemic status. War (ibid) has identified the presence of phonemic vowel /I/ in Khasi, but she fails to provide contrastive pairs to show its status as phoneme.

War (ibid) mentions the presence of triphthong /-iau/ in Khasi in words like miaw meaning ‘cat’ and siaw meaning ‘to whistle’. This observation given by War differs with the present analysis. In this present analysis, [w] in miaw and siaw is treated as a consonant occurring in the coda position rather than a vowel occurring in the nucleus position of a syllable. This can be further supported by the evidences provided for diphthongs. Based on these evidences, it can be said that Khasi has no triphthong.
War (ibid) mentions that there has been a proposal from some of the scholars like Nagaraja regarding the presence of long vowel /i:/ and /u:/ in Khasi. She further said that on careful examination, it has been found out that these vowels are not found in the language. This can be supported with evidences in borrowed words with long vowels. Some of the examples cited by her are

<table>
<thead>
<tr>
<th>English</th>
<th>Khasi</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i:/</td>
<td>‘beet’</td>
</tr>
<tr>
<td>/u:/</td>
<td>‘boot’</td>
</tr>
</tbody>
</table>

It is to noted that the observation made by War regarding the absence of long vowels /i:/ and /u:/ is in line with the analysis made in this present research work.

A review on the earlier works done in Khasi phonology reveals that the analysis of sounds presented in the present study is similar to some extent with earlier scholars. The main difference in the analysis presented in the present study is in the treatment of vowel sounds which include monopthongs, diphthongs and triphthongs. The analysis for consonantal sounds presented in the present study is similar, with the exception of voiced aspirated stops. Though scholars/linguists have mentioned about the asymmetrical patterning of stops in Khasi, no one has ever taken up this unresolved issue. In this present research, the absence of [g] and the phonemic status of [c] in (Standard Khasi) has been carefully dealt from historical perspective. In all of the earlier works, the study of syllable structure and stress in Khasi has not been discussed in detail. These two important areas in Khasi phonology has been taken up in this research.