Chapter II

LITERATURE REVIEW

2.0 Introduction:

This chapter attempts to review the concept of Linguistic Area as given by different linguists. The concept of linguistic area will be provided in two sections – Early Views and Current Views.

2.1 Linguistic Area:

The term linguistic area is used to denote a group of geographical contiguous languages characterized by a number of specific structure isoglosses. These shared areal features having acquired as the result of contact and not inherited. A linguistic area can, therefore, be seen as the natural counterpart of the language family, and this was certainly Trubetzkoy’s intention when he first proposed the concept in 1928, coining the term ‘Sprachbund’ or ‘Language Association’, to describe it. The name ‘linguistic area’ in English comes from (1943) translation of Sprachbund (literally ‘language union’), made widely known by Emeneau (1956). Thus, while languages may be grouped into language families on the basis of shared inheritance from a common ancestor, they may also be grouped areally on the basis of shared features which they have acquired as a result of contact.

2.2 Areal Feature:

An areal feature is the feature shared by languages within the same geographical area as a consequence of linguistic diffusion. Sharing of features between two or more languages can be due to contact of the languages belonging to different language families in the past. Due to borrowing, languages of different genetic origins have come to share certain
borrowed features – not only vocabulary, but also phonological, grammatical, or syntactic features.

2.3. Concept of Linguistic Area

2.3.1. Early Views

The notion of areal linguistic have long existed. There were discussions of borrowed structural traits and disputes about ‘mixed languages’ from the beginning of comparative linguistics (Girard, 1747; Schleicher, 1850:143; Ebel 1856 [cited by Kuhn, 1861: 75 and Haarmann, 1976: 20]; Müller,1861: 90; Schuchart, 1866–68; Whitney, 1868: 197 and 1979[1875]: 119; Schmidt, 1872; Powell 1891: 216–77; Meillet, 1921[1921]: 82; 1967: 102; Bloomfield, 1933: 468; Weinreich, 1953; Vendryes, 1968: 308, 319–20).

Before explicit definitions of ‘linguistic area’ were sought, numerous traits were identified of the Balkan area (Kopitar, 1829 and 1857; Schleicher, 1850; Miklosich, 1861; Sandfeld, 1902, 1912, 1930, 1934 and 1938; for others, Schaller, 1975: 37–48) and the South Asian (or Indian) linguistic area (Konow, 1906; Bloch, 1919, 1925, 1930 and 1934; Vendryes 1968: 305).

Boas (1917): The origins of modern areal linguistics can be traced back to Franz Boas’s (1917, 1920, 1929) work with American Indian languages. Boas identifies examples of shared structural traits which did not seem to fit the genetic classifications (Boas, 1920: 211). He spoke of ‘acculturation’ and ‘absorption’, and raised the question of the difficulty in some instances of distinguishing what was inherited from what was diffused. Boas’ ‘areal-typological’ approach was influential (Campbell, 1997a: 62–6). He compared the structural traits of languages in a particular region with their neighbours to determine whether they might be due to diffusion or be inherited, representing genetic relationships. Boas’s thinking influenced the

**Trubetzkoy (1923):** It has been argued that the concept of linguistic area or *Sprachbund* was first presented in Trubetzkoy (1923), as it happens that several languages in a region defined in terms of geography and cultural history acquire features of a particular congruence irrespective of whether this congruence is determined by common origin or only by a prolonged proximity in time and parallel development. They proposed the term language union for such groups which are not based on the genetic principle. (Trubetzkoy, 1923: 116, quoted in Toman, 1995: 204). Key elements are geographical region; and any shared features (whether from common origin or parallel development). Not stated: anything about borrowing or diffusion (though perhaps implied in ‘prolonged proximity’).

Trubetzkoy (1928) is better known; his ‘Proposition 16’ in the first International Congress of Linguists is generally cited as the origin of the concept. He spoke of the need for the notion ‘language group’ (*Sprachgruppe*) – a collection of languages bound to one another by a number of systematic agreements.

(Trubetzkoy, 1928: 18) divided *Sprachgruppe* into two types, families of genetically related languages and *Sprachbünde*. The latter were defined as: Groups composed of languages which show a high degree of similarity with respect to syntax, a similarity in the principles of morphological construction, and which offer a large number of common culture words, sometimes also an outward similarity in the phonological inventories, – but which possess neither systematic sound correspondences, nor has any correspondences in the phonological make up of units nor any common basic lexical items – such languages groups we call *Sprachbünde* (Trubetzkoy, 1928:18). A ‘language group’ with syntactic, morphological and often phonological similarities, lacking systematic sound correspondences, no common basic vocabulary. Not mentioned: borrowing,
possibly inferred from ‘lacking systematic sound correspondences, and no common basic vocabulary’. There is nothing especially ‘areal’ about this view, apart from the fact that languages that share such traits tend, by inference, to be near one another. Trubetzkoy (1931: 233–4; 1931: 350–1), in a paper largely about dialect geography, associated phonological areal traits with isoglosses of dialect geography, but which extend beyond language boundaries into other languages. It has become common to associate areal traits with dialect isoglosses (Jakobson, 1931, 1938; Weinreich, 1953; Jacobs, 1954; Martinet, 1956; Emeneau, 1980[1956]: 111, 1971, 1974 and 1980; Ramanujan and Masica, 1969; Katz, 1975; Masica, 1976, 1992: 111). Trubetzkoy credited Jakobson for examples and pointed out others of his own for example, glottalized consonants in the Caucasus region shared by North and South Caucasian, Indo-European and Turkic languages.

According to Trubetzkoy’s renditions, it is clear that the structures that are shared by ‘neighbouring languages’ are not a result of a common origin. Under a Sprachbund it is understood that a group of languages through common fate in the same culture area and through reciprocal influence has approximated one another so strongly that in any of them roughly the same thing can be said in roughly the same way. A group of languages, in the same cultural area, mutual influence, converge to say the same thing in the same way.

**Bloch (1934):** The most important early description of South Asia areal linguistic features is Bloch (1934). Although Bloch was primarily concerned with discussing the adoption of external linguistic features by Indo-Aryan, he nevertheless discussed the sharing of features within all of South Asia. Bloch in general sought to minimize the overall importance of the borrowing of non-Aryanisms for the history of that language family. Still his inventory of areal features provided an excellent starting place or the later establishment of an Indian linguistic area. It included both features that
are of a wide variety of South Asian languages as well as some limited to a few languages. Some of the most important areal features noted by Bloch are as follows:

1. The existence in Sanskrit of names of tribes of ancient peoples that formed rhymic pairs by a process of reduplication plus initial consonant change (e.g., Pulina - Kulina, Kosala - Tosala, Kalinga-Tralinga). This process seems to be Austro-Asiatic in origin.

2. The existence of Dravidian loan words in Sanskrit, dating as far back as Rigveda, e.g., RV *ulukhala*—‘mortar’, AV *musala*—‘pestle’.

3. The sharing by Santali, a north Munda language, of a number of vocabulary items with certain dialects of Hindi, as well as with Oriya and Bengali. In addition, there are several lexical items in Sanskrit that are probably of Munda origin: *tambula* ‘betel’, *kadala* ‘banana’, *bana* ‘bamboo arrow’.

4. The adoption of numerous lexical items from Indo-Aryan (particularly Sanskrit) into all Dravidian languages.

5. The presence in Indo-Aryan, Dravidian, and Munda (except Sora) of a series of retroflex consonants that contrast with dentals (a contrast not shared by Indo-Aryan’s other Indo-European relatives).

6. The development in Sanskrit’s phonological system of short *e* and *o*, in addition to long forms of these vowels. This is in accordance with Dravidian and Munda patterns, both of which have short and long *e* and *o*. 
Bloch also points out the following areas of morphological similarity among various Indian languages:

1. The consistent use in both Dravidian and Indo-Aryan of suffixes, and deemphasizing of the use of prefixes and infixes.

2. The absence in both Dravidian and Indo-Aryan of proverbs and prepositions as such.

3. The absence of a dual number (originally present in Sanskrit, and lost in Middle Indic).

4. Double nominative stems of nouns, the oblique stem admitting of the force of a genitive and of being followed by words more or less emptied of their proper sense.

5. Personal pronouns having two stems, that of the nominative and that of the direct and indirect object (e.g., H. māi ‘I’; mujh ko ‘to me’; G. hu ‘I’; m-/mar-1st sg. Obl.; Ta. naan ‘I’, en-akku ‘to me’).

6. The existence in the verb of third persons in the form of nouns, i.e. participles or participial stems (e.g., Ta. irukki-var “he-who is”)

7. Participles varying in gender.

8. Presence of an absolute construction. In this construction two independent clauses are linked together by adding a special non-finite verb form after the verbal stem of the main verb in the first of the clauses. Usually the clauses so conjoined have co-referential subjects, and stands in any of a limited number of logical relations to each other, i.e. temporal subordination, cause and effect, adverbial plus main predication.

10. The independent adoption in Marathi, Oriya and Sinhalese of the Dravidian relative participle to their syntax, an invariable adjective admitting a subject in the nominative in any construction.

This data needs to be interpreted carefully. Although he pointed out many areas in which Indo-Aryan has been influenced by non-Indo-Aryan sources, Bloch considered these areas to be essentially outside of the core of features that defines that language family. He states that “remarkable, and in certain cases conclusive as these concordances may be, the evolution of Indo-Aryan has not resulted in denaturalization” (Bloch, 1934:328); that is to say, even though there have undoubtedly been areas of mutual influence between Dravidian and Indo-Aryan, these areas have not affected the essential linguistic composition of those language families.

**Emeneau (1956):** M. B. Emeneau first postulated the existence of a South Asian linguistic area in 1956, but many of the criteria that he cited in setting up the area were pointed out earlier. A ‘linguistic area’ has been defined by Emeneau as “an area which includes languages belonging to more than one family but showing traits in common which are not found to belong to other members of [at least] one of the families (Emeneau, 1956). In other words, there are resemblances between languages which are the result of contact rather than common origin, such as Numeral classifies in Indo-Aryan Languages which clearly originated in the Eastern end of the sub-continent and typically associated with the languages of the Sino-Tibetan groups. There are number of linguistic features that have been cited in defining a South Asian linguistic area. The literature in which these features are discussed often focuses on the historical processes that lead to their dissemination over a wide range of languages. Many of the features identified as being areal have been observed in grammars dating back at least as far as Caldwell’s Comparative Grammar of the Dravidian or South-Indian family of languages (Caldwell, 1856). Almost all of these data are drawn from Indo-Aryan and Dravidian, with fewer taken from Munda and
from Tibeto-Burman. This concept provided scholarly substance for explaining the underlying Indian-ness of apparently divergent cultural and linguistic patterns. With his further contributions (published as a book in 1980 by Stanford University under the title language and linguistic area Essay by M. B. Emeneau), this area has now become a major field of research in language contact and convergence. Consequently, South Asia is now recognized not only as a ‘linguistic area’ but also as a ‘socio-linguistic area’, a ‘cultural area’ and also as a ‘translation area’.

M. B. Emeneau, in his important 1956 article “India as a Linguistic Area” added new data to the phenomena described by Bloch and postulated the existence of a distinct Indian linguistic area. Emeneau discussed the historical processes of borrowing that must have been involved the creation of the area and noted that “the end result of the borrowings is that the languages of the two families, Indo-Aryan and Dravidian, seem in many respects more akin to one another than Indo-Aryan does to the other Indo-European languages” (Emeneau, 1956). The bulk of Emeneau’s data were taken from Dravidian and Indo-Aryan, although he did cite information from Munda where available and relevant.

After recapitulating areal features scattered throughout earlier literature, the two main sets of new data, that Emeneau gave in this article involved what he called the “echo-word” construction, “classifiers” or “quantifiers”. Emeneau defined the “echo-word” construction as one in which a basic word formulated as CVX is followed by an echo-word in which CV is replaced by a morpheme gi- or u- or the like (or C replaced by m- or the like) and X echoes the X (or VX echoes the VX) of the basic word. The meaning of the echo-word is ‘and the like’, e.g. puli gili (tigers and the like) (Emeneau, 1956:10). According to Emeneau, this construction is found in all Dravidian languages, is widespread in Indo-Aryan, and is attested in Munda in at least Sora. Emeneau concluded that it is clear already that echo-
words are a Pan-Indic trait and that Indo-Aryan probably received it from Indo-Aryan.

Another areal feature proposed in this article is the use of “classifiers” or “quantifiers”. In constructions marked by these, when a noun is numerated by means of a numeral or a similar word, the construction contains also one of the smallish class of words or morphemes which we can call by either of these terms. The term ‘classifiers’ indicates that there are as many classes of nouns as there are classifiers; the term ‘quantifier’ indicates that in numeration of nouns there is always specifications of land or discrete entities as classed by various criteria (e.g. human vs animal, animate vs non-animate, long and thin vs flat and thin vs spherical). Such quantifiers are, to be sure, used in probably all languages; English has a ton of coal, two acres of land, three heads of cattle, etc.

According to Emeneau, the construction is of Indo-Aryan origin and has spread to Dravidian and Munda. It has been reported in Bengali, Assamese, Oriya, Maithili, and Emeneau added citations from Marathi, village Hindi and Nepali. In Dravidian, Emeneau showed forms from Kolami (Wardha dialect), Parji, Kui-Kuvi, Kurukh and Malto. He observed that in Munda, Korowa, Santali and Mundari, the forms are used, but in Sora they are not. He concluded that there is a large area of India, especially eastern and central India, with this feature the construction is originally Indo-Aryan. It spread thence to the other languages as a total construction consisting of numeral and classifier, and then was elaborated in some of the languages with native material, the native numerals, native morphemes as additional classifiers, etc.

A recent article by Emeneau (Emeneau, 1969) provides an extensive discussion of onomatopoetic forms as a Pan-Indian areal feature. A large portion of his data comes from Kota, and the Kota data forms an excellent basis for examining the feature throughout South Asia. According to
Emeneau, onomatopoetics occur in Kota in two basic syntactic constructions:

A. as a direct quotation followed by a form of a quotative verb in-
   (second stem id-) ‘to say so and so’.

B. as an expander (adverb) preceding the verb part of a predication.

A. avn dop i□r ki mu virti ko. ‘He fell down with the noise of
   falling’ (dop i□r ‘saying dop’;)

B. avn a□tltr dopn neta’lk vi□teyt … ‘He having fallen from the
   attic to the ground with the noise of falling (dopn).

A. Cadm kordr krdr vad□ guc i□i.ko ‘His voice having gone one
   becoming reduced gradually, he whispered (sain guc guc)’

B. a’n guc gucn a□dge.n ‘I will tell you secretly’ (guc gucn
   ‘whisperingly’)

In the examples cited above, the purely onomatopoetic elements are
dop ‘the sound made by a falling object’ and guc guc ‘the sound made while
whispering’. These constructions represent the formal means by which such
onomatopoetic words are integrated into Kota sentences. The onomatopoetic
words are a special class of lexical items having describable phonological
properties which have a restricted range of meaning. It has been pointed out
in scholarly literature (Caldwell, 1856:554; Emeneau, (1944:15-29;
Hoffman, 1952) that onomatopoetic forms such as those cited above occur in
a wide range of South Asian languages. Although the exact phonological
realization of onomatopoetic forms varies with the phonological rules of the
particular languages, and in the language-specific morphological or syntactic
constructions in which they are used, the existence of such forms seems to
be an areal trait.

---

8 Shapiro M and Schiffman Language and Society in Asia, 1983, p 118-120
Two further important South Asian areal features have been pointed out by Emeneau in his 1974 article, “The Indian Linguistic Area Revisited”. The first of these is the fact that both Indo-Aryan and Dravidian have morphemes with common sets of semantic and grammatical properties while the phonological realizations of the morphemes differ greatly across the language families. Their uses are remarkably similar. The second feature involves sets of lexical items in various South Asian languages that are semantically interrelated in common ways and that interrelate with caste structures in ways too similar to be due to chance.

**Andronov (1964):** Andronov (1964) listed additional areal features that are considered to have been borrowed from Dravidian to Indo-Aryan and features that were borrowed in the other direction. Of the former type are the simplification of consonant clusters in the Indo-Aryan in accordance with Dravidian phonological patterns; the frequent voicing, spirantization or deletion of intervocalic stops, also in accordance with older Dravidian patterns; the presence of a large number of onomatopoetic terms whose formation accords with older Dravidian patterns; the modification of Indo-Aryan syntactic patterns to those of the Dravidian languages; and the presence of “chains of participles and compound verbs”. Andronov also cited a number of grammatical features of Dravidian that he considered to be of Indo-Aryan origin; including the lost of short e and o in Brahui; the development of nasal vowels and diphthongs of an Indo-Aryan type in Brahui, Kurukh and several other languages; the development of aspirated consonants in some of the modern Dravidian languages; the lost of sentence types in which a synthetic negative form of the verb is used, and the adoption of constructions using a special negative word of the Indo-Aryan type; the lost of personal nouns, widely used in early Dravidian texts, in many modern Dravidian languages; the development of adjectives and adverbs in most modern Dravidian languages; and the development of compound and complex sentences.
Andronov attributed a greater degree of linguistic importance to the notion of linguistic area than those of Emeneau, going so far as to suggest that convergence of this sort may actually eradicate genetic boundaries between language families. He saw areal convergence as able to gradually weaken genetic lines of demarcation between language families and ultimately render them irrecoverable and meaningless as typology markers.

.....the so-called ‘genetic’ relationship of languages within one family, a remnant of naturalistic conceptions of language, is not primordial and perpetual. It is historic in its nature: it is formed gradually and gradually can it weaken and disappear. In this sense the development of the typological similarity of the modern Indo-Aryan and Dravidian languages can be regarded as a prerequisite or an initial stage in the formation of a new linguistic family. If the direction of their development does not change in the future, the now observed tendency to develop the formal similarity may gain strength and result in the formation of new relationship ties and of a new language family, which will be neither Indo-European, nor Dravidian (Andronov, 1964b:13).

Masica (1976): Masica (1976) in an article “The Order Typology as a Definer of an Indian Linguistic Area” and in a paper written jointly with A. K. Ramanujan “Toward a Phonological Typology of the Indian Linguistic Area” (Ramanujan and Masica, 1969). Unlike earlier investigators of the South Asian linguistic area, Masica was interested in determining the extent to which the known bundle of shared linguistic features is unique to South Asia. Past literature on Sprachbund phenomenon had been confined to demonstrating shared linguistic features and to describe the processes of convergence which led to the sharing. Masica’s concern is rather with determining the extent to which these features are purely South Asian ones, and whether such a linguistic area might be part of some larger “Asian Sprachbund.”

---

9. Language and Society in Asia, Shapiro, M and Schiffman, 1983, p121-122
The linguistic traits that Masica examined in dealing with these questions are not all the same as the features that had traditionally been examined in the Indian Linguistic area literature. Rather he examined a number of variables involving word order that Joseph Greenberg had advocated as basic indicators of linguistic typology, such as the relative positioning of subjects, verbs, and object in surface structures. Some languages of the world regularly place the verb at the end of clauses and after subjects and objects, while other languages generally place the verb between the subject and the object. Masica, in his paper (1974) examines this and other syntactic variables in five Indian languages (Hindi-Urdu, Bengali, Telugu, Malayalam and Santali), and compares the results obtained with a large sampling of non-South Asian languages extended from Europe to East Asia, marking the limits of the range of these syntactic features isoglosses as in classical dialect geography.

Masica (1974; 159-64), the syntactic traits ostensibly shared by much of the Indian Sprachbund are as follows:

1. The presence of verbs that follow their objects rather than precede them. When subjects are present, they will generally precede the objects.

2. The placing in position of goals of verbs of motion, predicate nominals and adjectives, adverbial complements, and infinitival complements of catenatives.

3. The use of verbal auxiliaries that follow rather than precede the main verbs with which they are used.

4. The marking of the syntactic functions of nominal elements by means of postpositions, as opposed to prepositions.

5. Patterns of word order in which adjectives, genitive phrases, demonstratives, and numerals generally the nouns that they modify,
6. Word order in which qualifiers precede adjectives. Included under “qualifiers” are intensifiers- ‘very’, ‘somewhat’, etc.

Masica considered the presence of the above syntactic features to collectively constitute an Indian syntactic norm. By comparing the dissemination of these features with their distribution in a sampling of adjoining languages, he is able to conclude that there is indeed such a thing as an Indian linguistic area. He cited four specific ways in which such an area can be substantiated:

A. First, the basic OV isogloss defines a massive middle segment of the world, a largely left-branching syntactic zone in which India is the main southern anchor. In keeping with its basically north-south trend, they named it as macrozone INDO-ALTAIC….

B. Subsequent isoglosses, especially those involving the structure of a noun phrase, cut the OV belt almost in two… and isolate the subcontinent as a separate sub-zone.

C. These subsequent isoglosses… define three transitional subzones, where the basic OV syntax began to give way to opposing phenomena: the Iranian, the Tibeto-Burman, and the Abyssinian….

D. A thick bundle of isoglosses separates India from Southeast Asia beyond Burma, from Arabic, and from Africa beyond Ethiopia (Masica, 1974:172)

Masica concludes that the area so defined is “not an area of transition (formed by the intersection of isoglosses), but a trait-core area, surrounded by concentric isoglosses” (Masica 1974:172).
2.3.2. Current Views

Though different linguists have provided a number of definitions on linguistic area, yet there are no acceptable definitions of the term. Linguists who are currently working in this field are of the view that linguistic area is ill-defined, and they have tried to provide various definitions of this term. Stolz (2002, p. 33) states that ‘the number of definitions on linguistic area is almost coextensive with the number of linguists working in the field’. He further adds, ‘working with a notion, which is yet to be defined conclusively can be exhilarating, since by definition there must be work to be done; but it can also be frustrating, since it can sometimes feel as though meta theoretical discussions are impeding progress on the all too necessary description and analysis of particular languages and linguistic areas’ (Stolz, 2002). Below are some of the current views on ‘linguistic area’:

**Trask (1996: 315):** Trask suggests that centuries of contact between languages can lead to a particularly striking result. Several neighbouring but unrelated languages can come to share a number of structural properties with one another, properties which they do not share with their closest genetic relatives elsewhere. A group of languages in which this situation obtains is called a Linguistic Area, or, using the German term, a *Sprachbund*.

**Campbell (1998):** Campbell views that, *areal linguistics*, related to borrowing is concerned with the diffusion of structural features across language boundaries within a geographic area*. Campbell (2002) attempted to give a definition that would represent the field generally. A linguistic area is a geographical area in which, due to language contact and borrowing, languages of a region come to share certain structural features. Structural features shared among languages of a geographical area (where usually some of the languages are unrelated or at least no all close relatives). It is assumed that the reason the languages of the area share these traits is because they have borrowed from one another.
Both Trask and Campbell also introduced the term ‘convergence’ for the process of progressive structural assimilation that is characteristic of linguistic areas. However, this appearance of definitional harmony is short-lived, and Campbell in particular goes on to raise a number of salient and persistent difficulties with the notion of linguistic area. Although both Campbell and Trask characterized linguistic areas as involving structural borrowing, Campbell (1998: 300) proceeds almost immediately to invoke ‘not only borrowed words, but also shared elements of phonological, morphological or syntactic structure’.

Thomason (2001): Thomason observes that linguists have struggled to define the concept linguistic area ever since Trubetzkoy (1923) coined the term. This is mainly because it is not always easy to decide whether a particular region constitutes a linguistic area or not. In spite of prolonged efforts to define ‘linguistic area’, there is no general agreement as to its definition. For example, even for the most widely accepted linguistic areas, such as the Balkans, scholars do not agree wholly on which languages belong to the area, which linguistic traits characterize the area, and even its precise geographical extent. In Stolz’s (2002: 260) words, ‘these terms Sprachbund, linguistic area, and areal type seem to invite as many meanings and readings as there are linguistic minds to contemplate them’.

According to Thomason, whether the number of languages involved one of the more salient features characterized a linguistic area nor implied to the definition of linguistic areas more a negative one. Thomason (2001: 99) further suggests that ‘perhaps the most prominent similarity is the lack of an explanation for many or most of the areal features’. It is notable also that Trask, Campbell and Thomason all work mainly by illustration rather than by definition in their outlines of linguistic areas. In other words, all work towards what a linguistic area is by giving examples of what a linguistic area has been said to be, rather than by giving a fairly watertight definition and some examples of situations the definition excludes. Even Heine and Kuteva
(2005), provide a set of characteristics they see as indicative of linguistic areas. They state that ‘this characterization is fairly general, it is not meant to be a definition; but it as a convenient heuristic for identifying possible instances of sprachbunds.’ In turn, the illustrations given tend to show how the concept is elastic and problematic, since it is debatable how far the candidate examples share unifying characteristics.

Thomason (2001:99) does not only have a confusion over definitions, or perhaps a proliferation of partially overlapping definitions, but also a proliferation of partially overlapping terms. Campbell (1998) talks not only about linguistic areas, but also about Sprachbunds, diffusion areas, adstratum relationships and convergence areas. While these are seen to a large extent as being inter translatable terms, there are attempts elsewhere to subdivide linguistic areas.

Heine and Kuteva (2005): Heine and Kuteva distinguish Sprachbunds, which are ‘defined by the presence of a limited set of linguistic properties’; metatype, where ‘the languages concerned exhibit a high degree of mutual inter translatability’; and grammaticalization areas, which ‘are the result of one and the same historical process, more specifically, of the same process of grammaticalization, even if there may be other properties in addition’ (though Heine and Kuteva also concede that these types may not be mutually exclusive). This foregrounding of grammaticalization echoes Campell’s (1998: 306) suggestion that different features may be weighted differently in the establishment of a linguistic area, though he does not single out a particular feature or feature type in ideas of convergence. Heine and Kuteva (2005:182) suggest that ‘the contribution that the study of grammaticalization can make to defining areal relationship is a modest one’ (whereas, on the other hand, they see cases of metatypy, like Kupwar in India or the East Tucanoan-Tariana contact region in north-west Amazonia, as almost prototypical cases of convergence). These definitional questions, of course, are not peculiar to discussions of
linguistic areas. In the literature on many linguistic topics, there have been lengthy discussions of what particular constellation of features is needed to identify in order to ‘approve’ a given label. On the other hand, it can be asked whether the very existence of those contributory features means there is no need to recognize the super ordinate category of grammaticalization or convergence in their own right. Both these rather mechanical, feature-counting and checking approaches can be challenged¹⁰.

Based on the discussion of the literature relating to the concept of Linguistic area, it can be said that different scholars, view the ‘concept’ of linguistic area differently. Some of them agree with the presence of a Linguistic area but there are others who disagree with this concept. This study sets out to examine the concept of linguistic area which cannot be ignored when one discusses about shared linguistic features. For example, Santhali, a Munda language, exhibits the presence of Ergative case marker. This morpho-syntactic feature exhibited by Santhali, is not a feature of Austro-Asiatic languages. Its presence in the language, possibly can be due to the influence of a language, particularly Indo-Aryan languages which have ergativity. This can be illustrated by the following examples:

Santhali: uni –do poisa –i ko –ked –a
He Erg money Abs Ask Past Agr
‘He asked for money.’

The above example supports the fact that ergativity is a shared feature in Santhali because of the contact with Austro Asiatic languages in a linguistic area. Another notable feature which can support the presence of linguistic area, is the presence of gender marker ‘-a’ which marks masculine and ‘-i’ which marks feminine in Santhali, as in the following example:

k□□a ‘boy’ ku□i ‘girl’

¹⁰ Thomason, Sarah Grey, Linguistic areas and language history
The example provided above, shows the presence of gender markers ‘-a’ and ‘-i’ marking masculine and feminine in Santhali. It is to be noted here that the gender marking system is not a typical feature of Austro Asiatic family. The presence of gender marker in Santhali can be attributed to the influence of Indo-Aryan languages (like Hindi, Assamese, Bengali), which exhibit the presence of this type of morphosyntactic feature.

2.4. North East India as a Linguistic area:

With reference to North East India, Moral (1994), claims that North East India can rightly be called a linguistic area. To support his claim, Moral (ibid) discusses certain linguistic features shared by languages of the North East. Some of the features listed by him are phonological characteristics, grammatical characteristics and shared lexical features.

2.4.1. Phonological Characteristics:

(a) The north eastern part of India is totally out of this bloc because not a single language, irrespective of family, has contrasts between dental and retroflex sounds, a unique feature of the language here is the total absence of any retroflex sounds. Instead the languages here have a whole series of dental or alveolar sounds which include oral or nasal stops, fricatives, laterals, approximants, flaps and trills. This absence of retroflex sounds and the presence of alveolar or dental sounds are features of typical northeast Indian languages.

(b) Another striking phonological feature shared by all the languages of the northeast is the use of velar nasal /ŋ/. Extensive use of this sound at all positions in all positions of a word is seen in all the languages of the north east and the way /ŋ/ is pronounced here clearly marked this area different from the rest of the mainland. It is interesting to observe that in all the languages of north eastern India /ŋ/ is always pronounced singly but in the rest of the country speakers of any language attach a homorganic sound /g/
immediately after /ŋ/, and it is pronounced simultaneously. Thus /ŋ/ is realized as /ŋg/. For examples, certain important place names of the north east like Rangiya, Pengeri, Dhing, are pronounced by speakers here as /roŋja/, /peŋeri/ and /dhiŋ/, respectively, but contrary to this language speaker from the mainland will pronounce then as /roŋ gia/, /peŋgeri/ and /dhiŋg/ etc. This is a common hurdle faced by all language speakers learning any language belonging to the mainland and vice versa. Moreover, there is a limitation of occurrences of /ŋ/ in a majority of Indian languages. This is a distinct Sino-Tibetan feature creeping into all the languages of the area because it is observed that for almost any Sino-Tibetan language in South-East Asia /ŋ/ is a very common sound.

(c) The palatal sounds of underlying Sanskrit have attained different sound values in different Indo-Aryan varieties. For instance, in Bangla they have become palatal affricates, in Marathi they have become dento-palatal affricates, but in Asamiya these sounds were lost due to the impact of neighbouring languages where the alveolars or the dentals are the most dominant.

Regarding the fricatives, one can postulate that /s, z, h/ occur in all languages irrespective of any family in the northeast. However, in Asamiya only an extra fricative /x/ is noted.

2.4.2. Grammatical Features

(a) Personal deictics or markers: The use of personal markers in regard to the use of various kinship terms in reference to the age and rank of both the speaker and listener is a feature which separates the north eastern languages of India from all other languages in the rest of the country. However, a few languages belonging to the Munda group in Central India, particularly Santhali have somewhat similar system. G. A. Grierson in modern Indo-Aryan Vernaculars (p 75) has mentioned this aspect, particularly in Boro and a few other Tibeto-Burman languages. According to Dr. Banikanta Kakati,
this is an Austric feature which has influenced all the other languages of this area. This seems possible because the distant Munda language of Santhali may have preserved this system, while the others might have lost it due to the Indo-Aryan and Dravidian impact where no traces of this system are seen.

(b) Negativization process: The process of negation of verbs in Asamiya is another feature which clearly demarcates it from the rest of the sister new Indo-Aryan languages and other Dravidian languages. In Asamiya /n/ is attached to the verb followed by a vowel which is the exact copy of the first syllable of the verb.

A similar system is also observed in Tai-Ahom where the negative marker is /m/ and it is prefixed to various verbs, as in the verbs /kin/ in Tai-Ahom which means ‘to eat’, and /mikin/ which means ‘not eat’.

Khasi also has a more or less similar system where the negative marker is /em/. Whenever it is prefixed to a verb negativization occurs.

This feature of prefixing the negative marker seems to be an influence from Khasi but it has undergone a slight change in its use as is observed in the examples from Asamiya and Ahom. Even Rabhamese, a variety of Rabha spoken in West Assam bordering the Khasi Hills, is reported to have a similar system.

(c) The use of plural suffixes in all the languages spoken in the northeast is another feature which marks them as different from a majority of Pan-Indian languages. In Asamiya, for instance, all the bound forms such as /hot/, /bur/, /bilak/, /mokha/, /zak/, /xokol/ etc. denote plurality and are suffixed to a noun or a pronoun. Boro and Garo have /bur/, /bilak/, /mokha/, etc. as plural markers.

Many Tibeto-Burman languages such as Rabha, Tiwa, Mising, Karbi, etc., use a variety of such types of plural markers. In Khasi also a plural
system is observed, as in /ieŋ/ ‘house’ but /sikitieŋ/ ‘house’, where the plural marker /sikit/ is prefixed. In Tai-Ahom, the plural markers are /nam/, /khOn/, /cheo/, etc., as in /konnam/ ‘man many’, /lucheo/ ‘many horses’, etc.

Moral (1994) mentions that the Khasi has plural markers. He cites the examples of a singular form /ieŋ/ ‘house’ and the plural form is /sikitieŋ/ ‘houses’ in which the plural marker is marked as /sikit/. It is to be mentioned here that from the data collected and based on the intuition the present researcher as a native speaker of Khasi, the plural marker given by Moral is found to be incorrect. The correct form of plural marker in Khasi is ‘ki’. Hence the plural for house as given by Moral should be /ki ieŋ/ ‘houses’ (Refer Chapter V, Section 5.1.)

(d) The derivation of nouns from verbs through suffixation is a common feature in all the languages of this area. For example, the verb /kha/ in Asamiya meaning ‘to eat’ changes to the noun /kha On/ meaning ‘good eating’. Almost all verbs can be transformed to nouns by the addition of the suffix /On/ in the language. In Garo, the suffixes /ani/, /gipa/ and /gimin/ attach to a verb to change it to a noun, as in the verb /dak/ ‘do’ which changes to a noun /dakani/ ‘being done’, /dakgipa/ ‘one who does’ and /dakgimin/ ‘already done’ respectively.

(e) The extensive use of classifiers in all the languages in this area is another feature which is a peculiarity of the northeast. However, in certain Indian languages, limited use of a few classifiers is noticed. Among the new Indo-Aryan languages, the most notable being Bangla, Oriya, Maithili and Marathi, a few classifiers are used. Similarly, other languages like Santhali, Kurukh and Malto of the Munda family make limited use of classifiers. But almost all the languages of the north east (Asamiya, Khasi or any other Sino-Tibetan language) use a huge number of classifiers. For almost everything or every shape these languages use a different classifier.
2.4.3. Lexical Features

(a) According to Moral, all the Tibeto-Burman languages (except Angami in Nagaland) spoken in the Brahmaputra valley show a uniformly similar percentage score and the rate of Asamiya (Indo-Aryan) influence is more in comparison to Khasi (Austric) and Angami (Tibeto-Burman). This may be attributed to the fact that Asamiya and the Tibeto-Burman languages, barring Angami are directly exposed to each other and the process of diffusion was rapid and extensive. Since Angami and Khasi are slightly isolated, therefore the rate of influence be it Indo-Aryan, Tibeto-Burman or Austric, is relatively less in comparison to the other languages. Another important factor to be considered is that a majority of the population of the different Tibeto-Burman groups in the Brahmaputra valley adopted Hinduism, and as a result the process of Aryanisation was accelerated, which is not true in the case of Khasi or Angami. All the language families have influence each other and the more the languages are exposed to one another the higher are the rate of influence. However, this minor study on lexical acculturation was to be taken up in an extensive way where all the Tibeto-Burman languages Khasi and Asamiya are taken into account in a much broader canvas.

(b) Another interesting lexical phenomenon is the commonality of place and river names of all the languages spoken in the north east, which is not shared by any languages in the entire Indian subcontinent. In the derivation of such names many Sino-Tibetan and Austric features have distinctly crept into, as in:

(i) The Boro word ‘di’ for water is noticed in innumerable river names in the north east, as in dihiŋ, diborU, dibonŋ, digaru, dikhou, dikreŋ, disang, digboi (boi meaning ‘flow of water’ in Asamiya.)

(ii) The Tai-Ahom influence is also evident in such names where the Tai-Ahom equivalent for water is nam. The mighty Brahmaputra is nam-ti-
lao, other such names are nam-daŋ, nam-zin, nam-khun, nam-saŋ, nam-
khe, nam-shao, nam-rup, nam-ti, nam-phuk, nam-sai, nam-choom, etc.

(iii) The Khasi equivalent for water is um and occurs in many names found in these parts as in umtru, umsuŋ, umiam, umraŋsu, umpling, um-khrnah, etc.

(iv) –ti– is typical Tai-Ahom classifier which has crept into all the languages in the north east. In Ahom, it is particularly used with place names which eventually became names of many important places in the north east, such as tiphuk, tipam, tiru, tirap, tifai, tiŋkhoŋ, tiŋrai, tiok, tiho, etc.\(^{11}\)

2.5. Micro-Linguistic Area

Masica (1992: 38) observes that, “South Asia shows many sub-areas which have features that place them within larger areal configurations”. Within a large linguistic area with a complex history, South Asia shows many sub-linguistic areas or micro-linguistic areas, for example, Nepal, Sri Lanka, etc.

There are a number of phenomena that clearly should be sought in identifying small scale linguistic areas. Such an area should include speakers of genetically unrelated languages living in proximity over a protracted period of time, where there are regular social contacts between the speakers of the different languages, possible leading to widespread bilingualism and/or code switching. In addition, there must be evidence of linguistic borrowing between the adjoining codes. In South Asia, a number of areas with such a concentration of circumstances exist. The Himalayas particularly in areas of contact between Indo-Aryan and Tibeto-Burman, Sri Lanka, with Tamil-Sinhala contact the Northwest India. With regard to Indo-Aryan, Iranian and Dardic, tribal areas in Eastern and Central India where there is

extensive contact between speakers of Dravidian and Munda Languages and Dravidian and Indo-Aryan border areas\textsuperscript{12}.

According to Abbi \textit{(2001)}\textsuperscript{13}, “A part from areal features there are also those features that identify a micro-linguistic area, and it will perhaps be of use to know about these features of working in the particular area”. Those sub-areal features can be enumerated by Abbi as Nasalisation (Northern India), Aspiration northern India), Gender Agreement (Western India), Right hand is ‘eating hand’ (Central India), Quotative verb ‘Say’ as complimentiser (Southern and North Eastern India), Relative-correlative pronoun (Northern India) and Classifiers (Eastern and Central India).

This study is perhaps the first to examine morphosyntactic features to establish micro-linguistic area. The following section is necessary to define morphology and morphosyntax.

\textbf{2.6. Morphology and Morphosyntax:}

The term \textit{morphology} was first coined by August Schleicher in 1859. Morphology is the part of linguistics which deals with the study of words their internal structure and partially their meanings. It is also interested in how the users of a given language understand complex words and invented new lexical items. As Morphology is concerned with word forms it is akin to phonology, it is also related to lexical studies as the pattern examined by Morphology are used to create new words. Furthermore it is also linked with semantic as it deals with the meaning of words. Scholars differentiate between derivational morphology and inflectional morphology. The former is concerned with the relationship of different words, and with the way in which vocabulary can be built from some elements, as in English, un-speakable, while the latter deals with the forms of one word that it takes up depending on its grammatical functions in a sentence. So in English it

\begin{footnotesize}
\textsuperscript{12} M Shapiro and Schiffman, \textit{Language and Society in Asia}, 1983, p 135
\textsuperscript{13} Abbi Anvita, \textit{A Manual of Linguistic Fieldwork And Structures of Indian Languages}, 2001, p 45
\end{footnotesize}
appears that it rather takes advantage of derivational morphemes than inflectional ones.

In Morphology morphemes are the smallest minimum meaningful units that carry meaning or fulfill some grammatical function. It is the identification, analysis and description of the structure of a given language's morphemes and other linguistic units, such as words, affixes, parts of speech, intonation/stress, or implied context. Morphological typology represents a method for classifying languages according to the ways by which morphemes are used in a language — from the analytic that use only isolated morphemes, through the agglutinative and fusional languages that use bound morphemes, up to the polysynthetic, which compress lots of separate morphemes into single words.

While words are generally accepted as being the smallest units of syntax, it is clear that in most languages, if not all, words can be related to other words by rules (grammars). For example, English speakers recognize that the words *dog* and *dogs* are closely related — differentiated only by the *plurality morpheme* -s, which is only found bound to nouns, and is never separate. Speakers of English recognize these relations from their tacit knowledge of the rules of word formation in English. They infer intuitively that *dog* is to *dogs* as *cat* is to *cats*; similarly, *dog* is to *dog catcher* as *dish* is to *dishwasher*, in one sense. The rules understood by the speaker reflect specific patterns, or regularities, in the way words are formed from smaller units and how those smaller units interact in speech. In this way, morphology is the branch of linguistics that studies patterns of word formation within and across languages, and attempts to formulate rules that model the knowledge of the speakers of those languages.

“The part of morphology that covers the relationship between syntax and morphology is called morphosyntax, and it concerns itself with inflection and paradigms, but not with word-formation or compounding.”
is the study of grammatical categories or linguistic units that have both morphological and syntactic properties. It deals with set of rules that govern linguistic units whose properties are definable by both morphological and syntactic criteria\textsuperscript{14}.

A linguistic paradigm is the complete set of related word forms associated with a given lexeme. The familiar examples of paradigms are the conjugations of verbs, and the declensions of nouns. Accordingly, the word forms of a lexeme may be arranged conveniently into tables, by classifying them according to shared inflectional categories such as tense, aspect, mood, number, gender or case. For example, the personal pronouns in English can be organized into tables, using the categories of person (first, second, third), number (singular vs. plural), gender (masculine, feminine, neuter), and case (subjective, objective, and possessive).

The inflectional categories used to group word forms into paradigms cannot be chosen arbitrarily; they must be categories that are relevant to stating the syntactic rules of the language. For example, person and number are categories that can be used to define paradigms in English, because English has grammatical agreement rules that require the verb in a sentence to appear in an inflectional form that matches the person and number of the subject. In other words, the syntactic rules of English care about the difference between \textit{dog} and \textit{dogs}, because the choice between these two forms determines which form of the verb is to be used. In contrast, however, no syntactic rule of English cares about the difference between \textit{dog} and \textit{dog catcher}, or \textit{dependent} and \textit{independent}. The first two are just nouns, and the second two just adjectives, and they generally behave like any other noun or adjective behaves.

An important difference between inflection and word formation is that inflected word forms of lexemes are organized into paradigms, which

\textsuperscript{14} http://www.wordnik.com/words/morphosyntax
are defined by the requirements of syntactic rules, whereas the rules of word formation are not restricted by any corresponding requirements of syntax. Inflection is therefore said to be relevant to syntax, and word formation is not. The part of morphology that covers the relationship between syntax and morphology is called morphosyntax, and it concerns itself with inflection and paradigms, but not with word formation or compounding.

Some of the morphological and morphosyntactic features that are taken up for the analysis in this research work are reduplication, affixes, classifiers, quantifiers, causative verbs and adpositions. The morphosyntactic features are case, tense, person, number and gender.

The review on earlier works done relating to linguistic area by Boas (1917), Trubetskoy (1923), Bloch (1934), Emeneau (1956), Andronov (1964) and Masica (1976) shows that in most of these works, the study is limited only to phonological and morphological/grammatical features of the languages belonging to Dravidian language family, Indo-Aryan language family and Munda languages of Austro-Asiatic family. In this present research, the study of linguistic area includes Garo and Mizo of Tibeto-Burman language family and Khasi of Austro-Asiatic language family spoken in Northeast India. This research focuses not only the morphological features but also the morphosyntactic features are subjected for the analysis.