Chapter Two
Theoretical Framework

2.1 Morphology down the Ages

Etymologically, morphology refers to the ‘study of shape’ of anything across the branches of knowledge. Morphology, though the term itself is only a 19th century coinage, as it relates to language study, has drawn active interest of grammarians and linguists down the ages starting from 1600 B.C. circa. The ancient Babylonian grammarians’ inscription on clay-plates on Akkadian language is one of the earliest recordings of morphological study. The organisation of Sumerian word categories like adverbs, verbs, pronouns etc. resembles the modern Paradigmatic study reflecting on person, number, tense, mood etc. The study of Morphology in the real sense, however, started with Panini’s (500 B.C) ‘Astadhyayi’. His work contains formal rules for units below the level of word. The terms coined by him are still found relevant and a few of them are used by the Generative Grammarians. Even before this, there was Greek Classical tradition. Plato had gone into the differences of Nouns, Verbs and Adjectives. Aristotle’s *Kratylos* first raised the question of arbitrary relation of meaning and sound. The Greek grammarians (216-90 B.C.) were more interested in the semantic rather than the formal aspects. The Stoics like Diogenes, Laertes and Appolonius took grammar as a major discipline within the study of philosophy. Even the Alexandrian grammarians separated it from philosophy and gave it the status of a specialized study.
Marcus Varro’s categorization of different ‘parts of speech’ based on their reflected case and tense in *De Lingua Latina* (47-45 B.C.) resembles the [±] mechanism of the Generative school.

With a renewed interest in Sanskrit and Hindi and the focus on Comparative study of Indo-European languages, a scientific and precise study of Morphology ensued towards the turn of the twentieth century. Before the Structuralists, there was a classical phase starting the Word Paradigm model – Hocket (1947), Nida (1949) etc. The early Generative period gave rise to Structuralism, with the ground-breaking work by Saussure (1916), Bloomfield (1933), Jakobson (1956, 1966) etc. The entire focus, however, was diverted to Phonology and Syntax ignoring the intermediate stage i.e. Morphology. As a result, there was a virtual halt in further development in Morphology. Therefore, a lull prevailed during the early 1960s to early 1980s. A renewed interest in Morphology starts after Chomsky’s treatise on the Lexicalist Hypothesis. Thereafter, Morphology has come a long way and is still going strong. However, Phonology seems to outrank Morphology, Semantics and Syntax most of the times. Even now with the Optimality Theory popularly known as ‘OT’ (Prince and Smolensky, 1993) the Phonological aspects have gained more prominence over Semantics and Syntax though a few works on them in this model of ‘out-put oriented grammatical theory’ is also going on. This OT framework has particularly ‘supplanted rule-based framework’ of input-oriented grammar.
2.2 The Contemporary Approaches to Morphology

It has been widely accepted that for any linguistic expression, be it a morpheme, a word, a sentence or even a discourse, it must have three distinct and interrelated properties viz. pronunciation, meaning and syntax. As such, a grammar of a language can be organized in any of the two schemes Fig. 2.2 (1) or Fig. 2.2 (2) proposed by Mohanan and Mohanan (1997).

Fig. 2.2 (1)

morpheme syntax phonology semantics

word syntax phonology semantics

sentence syntax phonology semantics

Fig. 2.2 (2)

discourse syntax phonology semantics

Word grammar=Morphology

Sentence grammar

Discourse grammar
In the first organizational method (Fig. 2.2 (1)) all the three properties – syntax, phonology, and semantics – are assumed to be integrated, whereas in the second one (Fig. 2.2 (2)) they are considered interrelated. In other words, phonology and semantics are assumed to interact with syntax from outside. Thus in the present-day study of Morphology, there is a gradual shift from the conventional approach to an approach leading towards convergence. The conventional meaning of Morphology means the study of (a) the internal structure of words and (b) the relationships between words. The approach of Selkirk (1982) and others assign the “word internal syntactic structure” to Morphology and the “sentence internal syntactic structure” to Syntax as schematized in Fig. 2.2 (3) - Mohanan and Mohanan (1997).
(Those outside the boxed items belong to the Sentence Syntax and the items in the boxes belong to the Word Syntax. The dotted box within Box 1 above is concerned with derivational processes, whereas the rest part of Box 1 and Box 2 comprise the inflectional processes. Box 1 is lexical and Box 2 post-lexical.)

With the inclusion of "word internal semantic structure" Lakoff (1993) under the domain of Morphology, there is a further shift. So now, Morphology includes the study of

(a) "word internal syntactic structure"
(b) "word internal semantic structure"
(c) "word internal phonological structure" and yet one more
(d) "word internal formative structure"

The inclusion of (d) above is, of course, disputed first by the school of Word and Paradigm Morphology – Matthews (1972), Anderson (1984), Zwicky (1988) – who hold that there are instances when morphological operations do not involve formatives and secondly by Amorphous Morphology (Anderson 1992) and Projection Morphology (Ford and Singh 1985) who outright reject that any morphological operation involves formatives.

As Morphology now includes the above-mentioned dimensions, there seems to be a shift in the definition of terminology, which
has been in use for decades. One such example relevant to this study is 'morpheme'. ‘Morpheme’ traditionally means the minimal meaningful unit of grammar. This is presently inadequate because it does not speak of the phonological, syntactic and semantic realizations. The inclusion of phonological element is, however, not a viable proposition in view of the distinction between morphemic compositionality and phonological compositionality as shown in the examples below:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>worked</td>
</tr>
<tr>
<td>Morpheme</td>
<td>work+ed</td>
</tr>
<tr>
<td>Formative</td>
<td>w3:k+t</td>
</tr>
</tbody>
</table>

The word in A as well as in B has two morphemes each, but the word in A has two 'formatives' whereas the word in B has only one 'formative'. Therefore, 'formatives' belong to the phonological string rather than to the morphemic one as believed earlier. As such, 'affixes' which were considered to be bound morphemes are now considered, given the difference between 'morpheme' and 'formative', bound formatives. Similarly, a stem is a unit that contains at least one formative. Consequently, stems and affixes are units of "formative structure" of words. It justifies the inclusion of (d) i.e. "word internal formative structure" in the study of Morphology.
As a corollary to the two models of grammar shown in Fig. 2.2 (1) and Fig. 2.2 (2), two other models are as shown in Fig. 2.2 (4) and Fig. 2.2 (5) below.

![Diagram](image)

As per the scheme in Fig. 2.2 (5), both syntactic structures and formative structures have access to phonological structure as well as semantic structure. However, the scheme in Fig. 2.2 (4) allows only formative structure to have access to phonological and semantic structure. The first model (Fig. 2.2 (4)) goes well with the Affixal Morphology of Lieber (1980), Marantz (1981, 1982) and Kiparsky (1983) and the second
one (Fig. 2.2 (5)) with the Word and Paradigm Morphology of Matthews (1972) and Anderson (1982).

Depending on what dimensions are accepted and what are left, there can be three models of Morphology:

Morphology I: It includes all the four dimensions (a), (b), (c) and (d)
Morphology II: It includes only (a) and (d)
Morphology III: It includes only (d)

2.3 Theoretical Model Adopted

Of the theoretical bases of grammar discussed in § 2.2, the models suggested in Fig. 2.2 (2) and Fig. 2.2 (4) affiliating to Morphology II (that is, the most widely accepted one) is adopted for the present work. Moreover, of the three streams mentioned earlier in § 1.2, the model primarily adopted for the analysis of Sambalpuri Compounds is that of Lieber (1980, 1983). For affixational Word Formation Processes, insights from a host of them like Selkirk (1982), Kiparsky (1982), Mohanan (1983) along with the works of Bauer (1983), Madhavan (1983), Maheswari (1983) Botha (1984), Sailaja (1990), Katamba (1993) and Mohanan and Mohanan (1997) are followed. However, Lieber's (1980, 1983) scheme remains the basic model throughout, except for Reduplication Formation, for which primarily Abbi (1992) is referred to. The major assumptions and
formulations of the later two streams of Lexicalist Morphology, (i) Lexical Phonology and Morphology which integrates phonology and morphology and (ii) Lexical Functional Grammar which dissociates syntactic structure from predicate argument structure, are applied in the analysis of Sambalpuri.

Some of the basic theoretical assumptions referred to in various literatures of Lexical module of Morphology are:

(a) That all lexical entries of un-derived forms, words and stems, before passing through the word formation processes are stored in the Lexicon.

(b) New words can be added only to the open classes of words i.e. the major lexical categories

(c) Certain DRs are blocked due to phonological, morphological, semantic or aesthetic factors. The Lexicon has a set of phonotactic constraints to filter out entry, which are not well formed according to the acceptable rules of the language.

(d) That a Word Formation Rule (hereafter written WFR) can take only the major lexical categories as its base does not involve any phrasal or sentential element in the process.

(e) DRs change the syntactic category of the base and consequently the derivational affixes are not sensitive to syntax. No syntactic rule is relevant in reference to word internal structure.
(f) All morphologically complex words have a 'head'. The head is invariably the right hand most element, be it a bound formative like a suffix or a free morpheme i.e. a word as in Compound formation.

(g) Affixes are arranged in the lexicon in certain ordered levels. At least two levels are generally accepted.

(h) The output of each layer of affixation input results in a well-formed structure or in other words a 'word'.

(i) There are combinatorial restrictions as a morpheme can attach to words of a limited lexical class.

(j) WFRs attach only one affix to the stem at a time. By no rule, two affixes can attach at the same time. In other words, the processes of affixation are done through a set of steps.

(k) The Affixation processes are Cyclic. An affix introduced at a later cycle presupposes the availability of information that only becomes available when rules of an earlier cycle have applied.

It may be mentioned here that the assumptions (a) to (c) are empirical in nature and are universally applicable. The rest of the assumptions (d) to (k) are applied to Sambalpuri/Koshli to find out their applicability or non-applicability as the case may be.

It is pertinent at this point to explain some of the terms used according to the definitions accepted in this thesis, because terms often get newer shades of meaning or even completely new meanings.
A 'root', whenever the term is referred to in the thesis, means 'a form that is not further analysable in terms of derivational or inflectional morphology' or in other words, a 'word-form' that remains when all inflectional and derivational affixes have been removed (Bauer 1983). It is the 'irreducible core of a word' with nothing else attached to it (Katamba 1993).

A 'stem' is the part of the word form that remains when only all inflectional affixes have been removed. It may contain one or more derivational affixes and be in a position to take further input of only inflectional affixes. In other words, the form is in existence before any inflectional affixes are added to it (Bauer 1983). This definition is of course not consistent with Lieber's, who seems to have used 'stem' to mean a 'base', which is ready to take further affixation, inflectional or derivational. To her a 'stem' is a morpheme that does not sub-categorise other morphemes.

A 'base' refers to any form to which affixes of any kind, derivational or inflectional can be added. However, a base to be so referred to may already have derivational affixes and it must be ready to take further input of derivational or inflectional affixes (Bauer 1983).
Both 'stems' and 'affixes' are taken as 'formatives' rather than as morphemes in consistent with the grammatical theory referred to in 2.2.

Affixes are added to a base to get an output word.

Input base $\rightarrow$ Affixation $\rightarrow$ Output word

Derivational affixes may cause any or all of the following changes viz. a) Phonological changes, b) Category changes, c) Semantic changes and d) Pragmatic changes. (Akmajian et al 1995)

As for the notational conventions, there are many such conventions of which the three relating to the lexical module are: (a) the convention of Boundary Distinction (Aronoff 1976, Siegel 1979), (b) the Feature Percolation Conventions (Lieber 1980), and (c) the Bracket Erasure Conventions (Pesetsky 1979, Mohanan 1982).

According to (a) formative boundary is indicated by + and the word boundary by #. The # boundary is automatically inserted at the beginning and end of every string dominated by one of the lexical categories like N, V, Adj. or Adv. The assignment of # boundary replaces or supersedes any + boundary that might already be in that position. The pre-cyclic or Class I affixes are attached with the + boundary and the post-cyclic or Class II affixes are attached with the # boundary. Class I formative +
boundary may attach to stems but Class II word # boundary can attach to words only. It clearly suggests that + boundary comes inside #boundary and after the attachment of #boundary, no further derivation is possible.

\[ [\text{adjectiv}]_N + \text{al]}_{\text{Adj.}} \rightarrow [\text{adjectiv} + \text{al]}_{\text{Adj.}} \# \text{ness}]_N \]

This convention of boundary markers becomes superfluous because of Level Ordering Hypothesis. According to this, affixes of different Levels are arranged in a fixed order; hence, the convention of Boundary Distinction is replaced by (c) the Bracket Erasure Convention. It eliminates the need of any boundary marker, and employs the universal Bracketing method allowing erasure of brackets at the end of a layer with the assumption that information available in the earlier brackets becomes irrelevant to subsequent layers of bracketing.

\[ [\text{defend}]_V \text{ant]}_N \rightarrow [ [\text{defendant}]_N \text{s}]_N \rightarrow [\text{defendants}]_N \]

There are, however, differences of opinion amongst linguists regarding the Levels of Affixation. While Siegel (1979), Booij and Rubach (1987) and others are satisfied with only two levels, Kiparsky (1982) and others speak of three levels, where as Halle and Mohanan (1985) speak of even four. Despite their differences, there seems to be some agreement as to the location of certain affixes. For example, all of them keep Irregular Inflection in Level 1, Derivation in both, Level 1 and Level 2, Regular
Inflection always in the last of the Levels (2 or 3 or 4 depending upon their classifications) (Fig. 2.3 (1), (2) and (3)). Halle and Mohanan’s inclusion of an extra Level between Level 2 and Level 3 is purely of phonological relevance for which they suggest a loop to link Level 3 (in this Level they keep ‘Compounding’) and Level 2, and does not affect the Morphological aspects in any way. Sailaja’s (1990) approach, however, seems to support the Level ordering of Booij and Rubach (1987). She suggests the keeping of Compounding, Derivation and Regular Inflections in a single Level.

Kiparsky (1982):

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Irregular Inflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class I Derivation</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level 2</th>
<th>Class II Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compounding</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Level 3</th>
<th>Regular Inflection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2.3 (1)
Halle and Mohanan (1985):

Level 1  Irregular Inflection
         Class I Derivation

         \[
         \]

Level 2  Class II Derivation  \[
         \]

Level 3  Compounding  \[
         \]

Level 4  Regular Inflection  \[

         Syntax

Fig. 2.3 (2)

Booij and Rubach (1987)

Level 1  Irregular Inflection
         Class I Derivation

         \[
         \]

Level 2  Class II Derivation

         Compounding

         Regular Inflection

         \[

         Syntax

Fig. 2.3 (3)
The Cyclicity of affixation processes is no longer considered a valid principle because of the fact that not all phonological rules, nor all lexical rules are cyclic. Even Szpyra (1989) argues that there are many dual class affixes, which attach very productively and that level ordering is inappropriate for such forms. It is widely accepted by the generative linguists that the lexicon is hierarchically organised by strata based on the properties of affixes, and affixes unambiguously belong to one stratum. Any phonological alteration in the roots is imposed on the roots from outside by affixes. The Stratum 1 affixes trigger changes in the base to which they are attached while Stratum 2 affixes are supposed to be neutral. Goldsmith (1983), however, argues that setting of strata based on properties of affixes is not tenable because the same affix can simultaneously belong to two strata. For the analysis of Sambalpuri, the stratum ordering seems to be of little use, as many of the affixes are not only both neutral and non-neutral but also the same affix, unlike that of English, is often used in many classes of roots. Strauss (1982) feels that the ordering restrictions are absolutely language specific. He speaks of another property of affixes that Class 1 affixes attach to stems that may not be words but Class 2 affixes can only attach to stems that are words.

The Feature Percolation Convention (hereafter written FPC) of (b) claims to be a simple method of capturing the total processes of word formation with the minimal rules and theoretical instruments utilizing a single, instead of a set of, Context Free Rewrite rules. As per FPC, both
inflectional and derivational affixes enjoy the same status in the lexicon. The features of a morpheme percolate or become a part of the node dominating that morpheme. Lieber (1980, 1983) speaks of four of such conventions, the first three for processes of affixation and the last one for the processes of compounding. Her theories have the following assumptions: that all morphemes and formatives, stems and affixes, have lexical entries and they all carry information about their categories and subcategorisation features (the lexical items they must attach to), their semantic representation, argument structure, diacritic specifications etc.

Convention I: All features of stem morpheme, including category features percolate to the non-branching node dominating the morpheme.

Convention II: All features of an affix morpheme, including category features, percolate to the first branching node dominating that morpheme.
Convention III: If branching node fails to obtain features by Convention II, features from the next lowest labelled node automatically percolate up to the un-labelled branching node.

Conventions IV: If two stems are sisters (they form a compound), features from the right hand stem percolate up to the branching node dominating the stems.
The above-mentioned notation is done in four steps, viz.

Step 1: unlabelled binary branching tree is drawn
Step 2: morphemes are inserted into the tree
Step 3: the non-branching nodes are labelled according to the features of the morphemes
Step 4: the branching nodes are labelled according to the stems they dominate

Fig. 2.3 (7)
Fig. 2.3 (8)

Such branching pattern is so broad that it even includes the derivational processes that may not be logically acceptable to a particular language especially in the processes of compounding. Therefore, she makes use of the Argument Linking Principles as an interpretative filter to block out unacceptable compound formations. Others like Williams (1981) have also used that mechanism. However, the LFG School of Linguists use the method of Theta (θ) role in encoding the semantic relation between the predicate and its argument. In the LFG model, the functional categories as [Subject], [Object] etc. are also included whereas in the model adopted by Lieber (1983) and Williams (1981) such functional categories have no place for feature percolation. There are minor differences amongst the individual
linguists regarding argument structures. For example, the Internal Argument (Lieber, Williams) is also called the Subject Argument (Bresnan, Kaplan), the External Argument the Predicate Argument. While Jackendoff (1972) and Gruber (1976) include 'Location' as an argument, Selkirk (1981) leaves it. However, the differences have been ignored for the present study. A lot of work is being done these days on the various aspects of Argument Structures such as in Construction grammar (Goldberg 1995) or in Grimshaw (1990, 1997) or Rosen (1990) but considering the limited accessibility to the resources, the present work confines itself to the Argument Linking Principles of Lieber (1983) alone. However, Allen (1978), Williams (1981) and others are referred to along with Lieber (1983) for the analysis of primary compounds.

Following Selkirk (1978), Lieber (1980) gives the following two lexical structure rewrite rules for compounds and derivational morphology which presumably precedes the feature percolation conventions as per the organizational scheme mentioned above:

**Compound formation:** \( X_s \rightarrow Y_s Z_s \)

**Derivational Morphology:**

i) \( X_s \rightarrow (Af) Y_s \)
   \( X_s \rightarrow Y_s (Af) \)

ii) \( X_s \rightarrow Y_r \)

iii) \( X_r \rightarrow (Af) Y_r \)
    \( X_r \rightarrow Y_r (Af) \)
The X, Y and Z stand for major lexical categories and the subscript s and r respectively stand for stem and root as per Lieber's definition. The above rewrite rules predict, "Since derivational affixes attach to roots or stems, they will never attach outside of inflectional affixes" – (Lieber 1980: 45). This prediction, however, has already been disproved (Botha 1984) as regards English but it seems to work in Sambalpuri. Lieber's (1980) scheme of the organization of the lexicon is given in Fig. 2.3 (9).

In the following two chapters on Prefixation and Suffixation processes in Sambalpuri, each affix is analysed according to Lieber's (1980) notational scheme giving the following details: a) The Category (conjugation/declension class), b) Phonological representation, c) Semantic representation, d) Sub-categorization feature, e) Diacritic feature and f) Insertion frame.
PERMANENT LEXICON

\[ \Downarrow \]

| Lexical Class (i) Morpholexical Rules | Lexical Class (ii) Morpholexical Rules | CATEGORY \( V \) | CATEGORY \( V \) |

\[ \Downarrow \]

LEXICAL STRUCTURE
Lexical Structure Rewrite Rule
Feature Percolation Conventions

\[ \Downarrow \]

STRING DEPENDENT RULES
Reduplication
Infixed
Umlaut
etc.

Fig. 2.3 (9)