Chapter 1

INTRODUCTION

CONTENTS:

1.1. STATEMENT OF THE PROBLEM
1.2. IDIOMS IN LANGUAGE
1.3. THE INDIAN GRAMMARIANS
1.4. DEFINING IDIOMS
1.5. IDENTIFYING IDIOMS
1.6. PROPERTIES OF IDIOMS
1.7. NATURE OF IDIOMS

1.7.1. TRANSFORMATIONAL BEHAVIOR OF IDIOMS
1.7.2. SCALES OF IDIOMATICITY OR IDIOMATIC CLINE
1.7.3. CRITERIA FOR IDENTIFICATION OF IDIOMS
1.7.4. TYPES OF IDIOMS
1.7.5. DISTINCTION BETWEEN COMPOUNDS AND IDIOMS
1.7.6. DISTINCTION BETWEEN PROVERBS AND IDIOMS
1.7.7. DISTINCTION BETWEEN COLLOCATIONS AND IDIOMS
1.7.8. DISTINCTION BETWEEN DEAD METAPHOR AND IDIOMS

1.8. OBSERVATIONS
1.9. METHODOLOGY
1.1. STATEMENT OF THE PROBLEM

The purpose of this study is two-fold: to develop a tool as an aid for a translator to come up with an appropriate equivalent idioms while translating between Telugu and Hindi. In order to carry out this task one must come up with certain criteria in identifying idioms for the purpose of transfer. It is essential for such a task to study idioms, their defining characteristics, i.e., the identifying criteria to enable a translator albeit a computer program to identify or distinguish an idiom from those which are not in a given text and translate them appropriately. It requires to improve our knowledge of idioms and discriminate between the literal and idiomatic contexts.

1.2. IDIOMS IN LANGUAGE

Idioms are extremely important but the most ubiquitous, and less understood categories of language. However, they are not a vestige of the language that one can choose to avoid altogether. They form an essential part of the general vocabulary of a particular language. A description of how the vocabulary of a language is growing and changing will help to place idioms in a proper perspective. But the problem arises with both an effect and a cause of disagreement over the exact definition of ‘idiom’ and its interpretation. The vast majority of idiomatologists, after more or less reflection, settle for a definition along the lines of ‘a complex expression/phrase whose meaning is not a compositional function of/not made up of the meanings of its parts’ (Cruse 1986): the precision of the wording varies. So too, enormously, does the interpretation.

It seemed more valuable therefore, to work out, thoroughly, a definition of an idiom, its identification, then study its properties, its nature and classification. There are several limitations in this work: the foremost problem is with the distinguishing idioms from compounds, proverbs, collocations and metaphors. Though these language structures have distinguishing terminology, they are not distinguished pragmatically. The second problem we faced is with the collection of idioms from the extensively available data.
The study of idioms in the context of computational application rather itself is a restricted and contrived study.

1.3. THE INDIAN GRAMMARIANS

Previous work on idioms has been sporadic, uneven, and often less than well known. The problem is as old as linguistics itself. The principle of compositionality of meaning that is, central to any consideration of idiom, was debated for some 1300 years by those highly perceptive and methodical precursors of modern linguistics, the Sanskrit grammarians. The Padavadins or Bhatta-school of the later \textit{Mimamsa}

\[
\text{\ldots regarded padas (inflected words) as the significant parts of a sentence and interpreted the sense of a sentence as the composite or united meaning of the padas that go to constitute it (Chakravarti 1933: 12)}
\]

They were routed, however, by the \textit{vakyavadins} (vakya = the sentence) of the prabhakara-school. Bhartrhari, expounding this position, argued that just as letters cannot be divided into smaller parts, so words are not divisible into letters nor sentences into words. Words may be analyzed into stem and formative, and sentences into separate words, according to the principle of apoddhara (disintegration); but this device although useful is unreal. A clear line is drawn

\[
\text{\ldots between the sentence and its so-called constituents (padas) or, in other words, between what is real and what is unreal...the sense conveyed by a sentence is also indivisible. Just as a word (sabdasphota) or a sentence does not really consist of any parts, so the meaning denoted by it does not admit of any division... Indivisibility is thus a peculiar characteristic that equally applies to both the}
\]
A specific example of this, involves the status of verbal prefixes. Upasargas (‘prepositions’) were generally agreed to be indicative (dyotaka) rather than denotative (vacaka), that is, they serve to specialize the more general meanings of verbs and nouns. Panini (5th century BC) recognized them as independently significant (Panini 1.4.93). His view was elaborated by others who claimed that they have particular meanings, and that their main function is to specialize the meanings of the nouns and verbs to which they are attached.

Patanjali (2nd Century BC), however, asserted that verbal roots carry a range of meanings in themselves and that prepositions have no particular meaning to contribute. Bhartrhari (8th Century AD) defended the Paninian position; but eventually Punyaraja resolved the debate by concluding that

…in cases of verbs joined with prepositions the meaning is derived usually from a harmonious combination of ‘dhatu’ [verb stem] and ‘upasarga’ and not from any one of them severally. (Chakravarti 1933:171; see further pp. 167-77).

Compound nominals were also carefully classified, distinguished from juxtaposed separate words, and departures from strict compositionality discussed (Chakravarti 1933:411-2 and 443ff). Thus, although there is nothing in the surviving corpus of the Indian grammarians’ work that corresponds precisely to a discussion of idiom, the crucial problem of compositionality is fully and carefully debated, and many valuable observations are made.
1.4. DEFINING IDIOMS

The miscellaneity of previous work on idioms is inevitable. The wild diversity is an effect and a cause of disagreement over the exact definition of ‘idiom’ and its interpretation.

Idioms are defined in various ways in the literature. Idioms are variously called word-combinations (Zgusta 1971), fixed expressions (Alexander 1987) and phrasal lexemes (Pawley 1985; Lipka 1990). It is loosely defined as a group of words whose meaning cannot be predicted from the meanings of the constituent words, for example: *it is raining cats and dogs.*

The basic working unit in Lexical Semantics is a lexical item. This has a form and a meaning. Generally speaking, lexical items are simply words, but there are interesting exceptions. Normally, the meaning of a grammatically complex expression is built up, in conformity with the principle of compositionality, by combining the meaning of its parts. However, there are some complex expressions that do not behave in this way (Cruse 1986).

The accurate meaning of an idiom is determined by the usual way of determining the meaning of any idiom in any language, that is, *by recognizing that there is a conflict between the literal meaning of the individual parts of the lexical unit and the way that that unit has apparently been used by an author or a speaker.*

Logan Pearsall Smith (1925) defines idioms as:

> The idiosyncrasies of our language, and, above all, those phrases which are verbal anomalies, which transgress, that is to say, either the laws of grammar or the laws of logic’. Thus, allowing that compound words can be idioms. The high incidence of oddity in the use of prepositions is noted;
grammatically distinguished from semantic anomaly; but … hundreds of English ‘idioms’, all phrasal and most figurative, classified according to their original semantic field.

Hockett (1958) is the first of the modern western grammarians to give serious consideration to the definition of idiom and its consequences. His discussion is worth quoting at length:

“Let us momentarily use the term “Y” for any grammatical form the meaning of which is not deducible from its constituent of a larger Y, is an idiom. A vast number of composite forms in any language are idioms. If we are to be consistent in our use of the definition, we are forced also to grant every morpheme idiomatic status, save when it is occurring as a constituent of a larger idiom, since a morpheme has no structure from which its meaning could be deduced…

…The advantage of this feature of our definition, and of the inclusion of morphemes as idioms when they are not parts of larger idioms, is that we can now assert that any utterance consists wholly of an integral part of idioms. Any composite form which is not itself idiomatic consists of smaller forms”.

Idioms will thus range from morphemes to proverbs or even poems, taking in pronouns, proper names, figures of speech, and private family languages. A dictionary should, ideally, contain only and all the idioms of a language, although this is not possible in practice. Idiom formation is a constant process.
Idiomaticity is taken to be completely pervasive of language: generative grammarians of all sorts thought of idioms as a listable set of aberrations, but more recent work such as that of Bolinger (1975), and Mitchell (1971) has swung back to seeing idiomaticity (or something like it) as common throughout. Hockett deliberately and carefully admits morphemes to idiom status.

The standard non-compositionality definition is that, ‘the essential feature of an idiom is that its full meaning… is not a compositional function of the meanings of the idiom’s elementary grammatical parts’ (Katz and Postal 1963).

Wood (1969) studied some of the problems raised by the meaning and form of idioms and idiom-like expressions and came out with conclusions on idioms as:

1. True idioms are wholly non-compositional, or opaque, in meaning.
2. Ambiguity is a common but not a necessary feature of idiomaticity.
3. Forms with a unique constituent need not be idioms, but those containing a cranberry-form are.
4. True idioms can be opaque in structure.
5. True idioms are wholly non-productive in form.
6. Single compound words can be idioms.

On the basis of these conclusions, Wood (1969) proposes the following definition:

An idiom is a complex expression that is wholly non-compositional in meaning and wholly non-productive in form.

1.5. IDENTIFYINGIDIOMS

A traditional definition of idiom is an expression whose meaning cannot be inferred from the meaning of its parts. Although at the first sight, straightforward, there is a curious
element of circularity in this definition. The definition, therefore, must be understood as stating that an idiom is an expression whose meaning cannot be accounted for as a compositional function of the meanings its parts have when they are not parts of idioms. The circularity is now plain: to apply the definition, idiomatic and non-idiomatic expressions must be distinguishable. Fortunately, it is possible to define an idiom precisely and non-circularly using the notion of a semantic constituent. Two things are required of an idiom: first, that it be lexically complex – i.e., it should consist of more than one lexical constituent, second, that it should be a single minimal semantic constituent.

Consider, for example: ‘This will cook Arthur’s goose’. The test of recurrent semantic contrast reveals that this, ‘will’ and ‘Arthur’ are regular semantic constituents; the rest, however, i.e., ‘cook someone’s goose’, constitutes a minimal semantic constituent, which as a whole contrasts recurrently with, say, help or destroy. ‘to cook someone’s goose’ is therefore an idiom.

As a contrast to the above, we shall regard as non-idiomatic (or semantically transparent) any expression that is divisible into semantic constituents, even if one or more of these should turn out on further analysis to be idioms. Most idioms are homophonous with grammatically well-formed transparent expressions. A few are not in this sense well formed, although some grammatical structure is normally discernible. Such cases, as ‘by and large’ and ‘far and away’, are often called asyntactic idioms.

Haas (1985) while defining an idiom briefly characterizes it as a lexical complex that is semantically simple. For Matthews (1972), the expressions that are syntactically complex, but semantically simple, are called idioms. Similarly, the expressions that are semantically peculiar are usually described as idioms (Cruse 1986).

Thus, Trask (1993) defines an idiom as an expression consisting of one or more words whose meaning cannot simply be predicted from the meanings of its constituent parts. For semantic reasons, an idiom requires its own lexical entry in the lexicon….
In fact, idioms add spice and bring evocative and exciting flavor in a language. Idioms have very specific applications that are not obvious from simply knowing the individual words (Applebee, Jane and Rush, Anton 1996). Certain consequences follow naturally from the fact that the apparent constituents of idioms do not have independent meanings (Cruse 1986). Consider some examples of idioms:

Te. kadupu maMta
Hi. I.M: IrRyA: ‘jealousy’
Hi. L.M: peta_meM jalana: ‘burning sensation in the stomach’

Here, the semantic sense of “kadupu maMta” would be “peta_meM jalana”. However, idiomatically, the expression represents ‘jealousy’. Such expressions are found commonly in languages wherein their usage changes according to the contexts in which they are used.

Apart from these, other such expressions are:

<table>
<thead>
<tr>
<th>Te.</th>
<th>Lexical sense in Hindi</th>
<th>Idiomatic sense in Hindi</th>
</tr>
</thead>
<tbody>
<tr>
<td>gAlilo xlpaM</td>
<td>havA_meM xiyA</td>
<td>saMxigXa sWiwi</td>
</tr>
<tr>
<td>nawwa nadaka</td>
<td>SaMbUka_kA calanA</td>
<td>bahuwa Xire calanA</td>
</tr>
<tr>
<td>mulYla bAta</td>
<td>kAztoM_se BarA raswA</td>
<td>KawaranAka sWiwi</td>
</tr>
<tr>
<td>cApakiMxa nIrur</td>
<td>catAyI_ke nIce pAnI</td>
<td>aprawyASiwa bAwa</td>
</tr>
</tbody>
</table>

Generally, all idioms are composed of elementary lexical units. It is interesting that although idioms consist of more than one word, they display to some extent the sort of internal cohesion that we expect of single words. For instance, they typically resist interruption and re-ordering of parts. Some of the restrictions of syntactic potential of idioms is clearly semantically motivated. For instance, the reason that ‘to pull someone’s left leg’ and ‘to kick the large bucket’ have no normal idiomatic interpretation as that
‘leg’ and ‘bucket’ carry no meaning in the idiom, so there is nothing for ‘left’ and ‘large’ to carry out their normal modifying functions (Cruse 1986).

The same is true of re-ordering. Many grammatical processes involving re-ordering of constituents are ruled out for semantic reasons, particularly those whose semantic function is to highlight a specific semantic constituent: thus, ‘What John pulled was’ his sister’s leg’ has no idiomatic reading, whereas, ‘What John did was pull his sister’s leg’, which leaves the idiom physically intact. In ‘Its my/John’s goose that was cooked, not yours’, and ‘Its my/Arthur’s leg he’s pulling, not yours’, both of which have a normal idiomatic interpretation, do not constitute evidence of semantic life in the elements of the idioms. What is being topicalized in these sentences is ‘John’, ‘Arthur’, etc., which are semantic constituents, and not part of the idioms; the possessive affix (which is part of the idiom) simply has to accompany the noun to which it is attached. ‘My’ and ‘your’ in these sentences must be analyzed as ‘I’/’you’ + ‘possessive’: only the possessive forms part of the idiom (Cruse 1986). But semantically innocuous re-orderings are also to some extent resisted:

A: John has a bee in his bonnet about it.
B: ? John has a bee about it in his bonnet.

At the same time, idioms show their status as phrases in various ways, too. For example, if an idiom is inflected, the inflectional affixes are carried by the grammatically appropriate elements within the idiom, whether or not they are semantic constituents; that is to say, the elements of an idiom retain at least some of their grammatical identity:

A: John has bees in his bonnet about many things.
B: * John has bee-in-his bonnets about many things.

Likewise, in certain regular grammatical re-formations the parts of an idiom may behave as they would in a transparent expression: thus, we have a ‘leg-pull’, formed on the same
pattern as ‘hand-shake’. For these reasons, it is observed that it would not be appropriate to assimilate idioms to the category of words.

Apart from this, words alone can function as anchors to clitics, and constituents of idioms are often seen to take clitics. This indicates that idioms are like phrases in their structure but semantically like words. The following table indicates the study of various linguistic structures in the context of their meaning, constituency, the changes they undergo when their meanings and constituents corresponding to each other, and their output in the canonical shape. The following table allows us to substantially distinguish various language structures basing on their meaning and constituency.

<table>
<thead>
<tr>
<th></th>
<th>Meaning (M)</th>
<th>Constituency(C)</th>
<th>M-C Correspondence (Compositionality)</th>
<th>Canonical Shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phrase</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Word</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Collocation</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>Idiom</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Compounds</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Table 1.1: Linguistic structures and their M-C Correspondence

An individual uses a phrase that generally includes words referring to everyday concrete objects, figuratively in a particular situation. His interlocutor, then the local speech community, and finally the body of native-speakers accept the metaphorical use into their own speech for its aptness, felicity, picturesqueness or even plain illogic. In fact, Smith (1925) considers irrelevance, illogical and absurdity to be important factors in the genesis
of many idioms. Where there is metaphorical use of the kind, the phrase in question may undergo the well-known processes of semantic extension, specialization, etc. that operate in any shifts of meaning. The resulting phrase may become an idiom, with no obvious link with its literal predecessor (which of course in most cases still exists independently). At any particular moment, therefore, synchronically there are (i) literal phrases, (ii) metaphors that are clearly connected with the literal, (iii) other more opaque but possibly interpretable phrases, and (iv) opaque idioms.

The conclusion to be arrived at from the discussion above is that a pure idiom must have constituent elements from which the overall meaning of the whole is not deducible. Such a viewpoint not only raises the issue of the connection between idiom and metaphors but also a number of others (Fernando and Flavell 1981).

Healey working within a tagmemic model, takes up the identification of idioms as a major concern and formulates three operational tests whereby idioms may be identified:

(1) The substitution or replacement-test whereby any morpheme replaceable by another is identified as a non-idiom. In ‘long live the King/the Queen/President Johnson’, ‘long live’ is an idiom which like ‘have cold feet’ cannot be lexically altered and retains its idiomaticity. The various tagmemic or functional slots in an idiom do not take slot-filler variables, whereas for instance the non-idiomatic slot following ‘live’ does.

(2) Once all non-idiomatic parts, if any, have been shorn off, a suspected idiom can be further subjected to a predictability test. “Darte” (the game) which qualifies as a possible idiom because it has a non-literal meaning emerging in contrast with the bimorphemic ‘dart + s (plural), turns out to belong to a lexical set consisting of ‘bowls’, ‘quoits’, ‘noughts and crosses’, ‘snakes and
ladders’, ‘billiards’, ‘dominoes’, etc. and consequently has the predictability of a member of a rule-governed set of forms.

(3) A third test of idiomaticity is to expose an expression that has been tested by means of (1) and (2) to as many transformational changes as its internal structure will permit. Since Healey believes, like Uriel Weinreich (1969) and Bruce Fraser (1970), that transformational constraints are an indication of idiomaticity, transformational deficiency in an expression which has passed tests (1) and (2), or seems doubtful by such criteria, is confirmation of that expression being idiomatic. Such transformational deficiencies reflect lot of idiomaticity rather than loss of grammaticality. Transformational constraints operating on idioms reflect, thus, a semantic rather than a syntactic phenomenon (Healey 1968).

Though Healey identifies idioms on essentially semantic grounds, he is quite right in pointing out that ‘…each idiom has a definite form, consisting of an integral number of morphemes and tagmemes’ (Healey 1968).

Makkai (1972: 58) argues that it is not only more economical but also more insightful from the point of view of identifying idioms “to use the term idiom only for units realized by at least two morphemes”. The potential ambiguity of idioms, their ‘disinformational potential’, arises from this distribution.

The four criteria advanced by Makkai (1972) for identifying idioms are:

(1) The presence of at least two free morphemes in a given expression,

(2) The ability of these morphemes to function with different meanings in more than one environment,

(3) The potential ambiguity of all idioms in decoding arise from the possibility of literal interpretation, besides the idiomatic sense,
The semantic irregularity of idioms arise from the fact that an idiom has a meaning that cannot be deduced from its component parts.

In other words, Idiom is a morpho-syntactic phenomenon, for it is at this level of the language that the key property of idiom, the asymmetry between semantics and syntax manifests itself most unequivocally. A stretch longer than the sentence may be rendered idioms by its non-literalness, but such mini-discourses do not become idioms but remain nonce-items bound to one specific situational context. Those items that eventually gain idiom status are generally short, easily memorable items like compounds, phrases and syntactically simple sentences such as ‘the coast is clear’, ‘a stitch in time saves nine’, ‘a rolling stone gathers no more’, etc. The general tendency is towards deletion and hence reduction. ‘To draw a red herring across the trail’, is usually shortened to ‘red herring’, ‘a bird in the hand is worth two in a bush’ to ‘a bird in the hand’, ‘a rolling stone gathers no more’ to ‘X is a rolling stone’, etc. The principle of least effort is in operation all the time (Fernando and Flavell 1981).

The operational tests most frequently used for establishing the boundaries of idiom are those of removing or replacing the morphemes one by one in the suspected idiom (Healey 1968). The replacement or substitution test is generally more effective in the establishment of idiom boundaries since the possibility of substitution either converts an idiom into a non-idiom (‘pay through the nose -> pay heavily’) or indicates the presence of some non-idiomatic variable: ‘X gave Y a kick in the pants’, ‘X paid through the nose for Z’, ‘X is a male chauvinist pig’, etc. Any attempt to replace any other morpheme in, for instance, ‘X paid through the nose for Z’ (‘* paid down the nose’, ‘* paid through the eyes’, etc.) results in loss of idiomaticity. Hence the minimum form of the idiom is ‘pay through the nose’ (Fernando and Flavell 1981).

In another instance, Healey (1968) agree that only forms which are completely non-productive, i.e., unique, should be called idioms.
1.6. PROPERTIES OF IDIOMS

As the process of ‘idiomaticization’ lies in diachronic evolution, idiomaticity cannot be adequately explained by generative rules. Neither a given sense nor a given syntactic structure by itself constitutes an idiom. Rather it is the regular association of one with the other that is the source of idiomaticity. Such an association is the product of contextual extension in the everyday situations of communicative use over a period of time. The majority of idioms, as agreed by idiomatologists (Smith 1925; Hockett 1958; Healey 1968; Makkai 1972), exhibit certain discernible stages in their development (Fernando and Flavell 1981).

Perhaps idiomaticity is too complex and pervasive to be captured within the narrow confines of a single definition. The central problem one comes up against in attempting to define idiom is identifying the property (or properties) that will adequately capture all the idioms in a language while excluding all the non-idioms. Makkai argues that it is more economical “to use the term ‘idiom’ only for units realized by at least two (free) morphemes” (Makkai 1972). We have, accordingly, two conflicting criteria, a conflict that is reflected in the variety of morphological forms that have been identified as idioms: bound forms, single free forms, compounds, phrases and sentences. As far as the typology of idiom go, such forms range from proverbs and metaphors to a variety of set phrases including rhetorical questions and social formulae.

However, the five properties of idiom that are most regularly invoked have been formulated by Fernando and Flavell (1981) as follows:

1. the meaning of an idiom is not the result of the compositional function of its constituents;
2. an idiom is a unit that either has a homonymous literal counterpart or at least individual constituents that are literal, though the expression as a whole would not be interpreted literally;
3. idioms are transformationally deficient in one way or another;
(4) idioms constitute set expressions in a given language;
(5) idioms are institutionalized;
(6) they are not productive.

Of the six properties listed, the fact that an idiom is an object of non-literal function has the highest common denominator of idiomaticity. Though the structural composition of an idiom is regarded as useful in determining the upper and lower limits of idiomaticity, other properties have been chosen in attempting to formulate a definition of idiom. In trying to separate the more idiomatic from the less idiomatic and from the non-idiomatic, those properties that establish the semantic unity of idioms are more important than those that establish points of grammatical contrast (e.g. transformational constraints). Such properties are of three sorts: semantic, syntactic and sociolinguistic (Fernando and Flavell 1981).

1.7. NATURE OF IDIOMS

The design features of Human Languages allow extreme flexibility in actual use, hence, language should not always be taken at face value. In fact, the phenomenon of using words and sentences in roundabout ways is extremely common. In addition to irony and downright lies, metaphors, idioms and proverbs are ways of saying things more or less indirectly. Usually, idioms cover slangs, proverbs, certain metaphors and similes (Applebee and Rush 1996). Apart from these some compounds also confuse by being similar to idioms.

Thus, idiomaticity is best defined by multiple criteria, each criterion representing a single property. If idiomaticity is so defined, certain types of idiom will be seen to possess more distinguishing properties than others. There exist, in other words, varying degrees of idiomaticity correlating with different types or categories of idiom. The adoption of multiple criteria would enable the investigator to filter out the non-idiomatic while retaining all those forms that show one or more of the properties of idiom. Hockett (1958), uses a single criterion for defining and identifying idioms: that the meaning of an
idiom is not the compositional function of its constituent parts. Makkai uses five: (1) morphological composition, (2) the susceptibility of an idiom to literal interpretation, (3) ambiguity, (4) semantic unpredictability, and (5) institutionalization. The result of using multiple criteria is that Makkai’s definition of idiom is more explicit and his identification of what forms are idioms are more selective than Hockett’s.

The point that emerges from this section is that idiomaticity is too all-pervasive to be correlated with a specific form of morpho-syntactic structure or the presence or absence of syntactic constraints such as given transformations.

As Randolph Quirk (1960) states: ‘The problem of idiomaticity is rather that most phenomena in language respond very well to treatment by the procedures that have evolved for handling ‘syntax’ on the one hand and ‘lexicon’ on the other’. Thus, idiomaticity is the outcome of the intersection of the phenomena of the form, the sense and the situational context. After examining several issues, the most satisfying and sensitive criterion to establish idiomaticity is undoubtedly the semantic one. Hence, there can be little doubt of the primacy of semantic criterion in establishing the idiomaticity of any expression (cf. Fernando and Flavell 1981).

An idiom has a literal homonymous counterpart that complements the fact that its syntax is non-correlative, i.e. a pure idiom constitutes a ‘double exposure’. The most important feature of idioms is that they are expressions that are ambiguous and therefore potentially misleading (Makkai 1972).

The specific term used by Makkai is ‘disinform’. Makkai contrasts the disinformation potential of idioms arising from the possibility of “logical yet sememically erroneous decoding” with the ‘misinformation’ which occurs as a result of “accidentally homophonous forms” having “equally meaningful decodings” in similar environments, for example ‘she bears children’ signifying both ‘carries’ and ‘gives birth to’.

Certain set expressions manifest what may be termed ‘a double exposure’ (Henry G.
Widdowson, personal communication), i.e. they manifest a non-literal and a literal meaning. Non-literal idiomatic meaning generally arises as a result of a figurative extension. Yet the original situation, which is the source of the extension, is often unperceived by the average speaker.

In a typical idiom such as, ‘blow one’s own trumpet’, or ‘be in a hot spot’, the idiomatic and literal meanings are capable of appearing simultaneously as a ‘double exposure’, a kind of pun, in one of the same context. Since pure idioms have literal counterparts, they show no special peculiarities of encoding in themselves. In other words, they are non-anomalous in terms of selectional and strict sub-categorial restrictions. To put it differently, the patterns of collocation and colligation such idioms manifest in their composition is normal and predictable. Being more covert, a pure idiom could also be more deceptive. Since the essence of idiomaticity is an asymmetry between syntax and meaning, the presence of a homonymous literal counterpart complements such asymmetry both structurally and contextually.

An idiom is a syntactic unit that manifests lexical integrity. Makkai (1972) identifies this property of an idiom as that of encoding. Pure idioms are simultaneously idioms of both encoding and decoding.

1.7.1. THE TRANSFORMATIONAL BEHAVIOR OF IDIOMS

Idioms can be modified by items not part of the idiom and sometimes even the order of the idiom reversed in order to achieve surprising and unusual effects. In such instances an element of the idiom may be changed, without its original import being changed at the same time, by substitution, inversion, or deletion.

Some deletion is possible in certain English idioms, but there are also large numbers, particularly those comprising verbs with either preposition or particle, where no deletion is possible: ‘see through someone’, ‘bring the house down’, ‘turn on’, ‘put up’, ‘step up’, etc. The presence of prepositions and particles is one of the salient criteria used by Cowie
and Mackin (1975) in identifying verbal idioms for inclusion in the ‘Oxford Dictionary of Current Idiomatic Usage’. The possibility of deletion in idioms is largely a matter of use and would vary widely from language to language.

Idioms composed of adjective + noun do not allow otherwise the characteristic predicative usages, nominalizations, the formation of comparatives or superlatives, or modification, whereas this would be possible in a literal use, as Weinreich (1969) has demonstrated in ‘a hot sun’, ‘a hot dog’; ‘the sun is hot’, ’*the dog is hot'; 'the heat of the sun', '*the heat of the dog', etc. Similarly, idioms including a verb in their make-up usually do not allow many transformations that are open to their literal counterparts.

Fraser’s (1970), hierarchy that looks at the transformational possibilities of idiom ranges from L0 (completely frozen) through L1 (adjunction), L2 (insertion), L3 (permutation), L4 (extraction), L5 (reconstitution) to L6 (unrestricted). Such a hierarchy is useful for classifying idioms from the point of syntacticity, but it would have greater value for idiomaticity per se if it were possible to establish a syntax-semantics correlation as a complementary basis for the hierarchy. In other words, the application of the transformation is linked to the achievement of a specific stylistic flexibility in idiom syntax than the formulations of the constraints on idiom manipulation.

The peculiarities of encoding in hyperbole and metaphor could result in certain textual incongruities. These could be semantic, as in 'wage indexation is a dead duck' (-Concrete, +Concrete) or collocational as in the co-occurrence of 'storm/tea cup' and 'born with/silver spoon'. Such semantic and collocational clashes occur within hyperbole and metaphor but they do not occur in what are called 'pure idioms' such as 'smell a rat', 'pull up one's socks', 'scratch somebody's back', etc.

An idiom is an institutionalized expression, i.e., ‘it is approved by the usage of the language’ (OED). Unless an idiom has currency among the members of a specific speech community or a sub-group of such a community for a reasonable period of time it cannot be regarded as institutionalized. By the institutionalization of idiom, we mean the regular
association in a speech community of a given signification with a given syntactic unit (a compound, a phrase or a sentence), such that the resulting expression is interpreted non-literally. In other words, part of the phenomenon of idiomaticity is the institutionalization of an asymmetry between sense and syntax in the case of compound, phrasal and sentential idioms.

### 1.7.2. SCALES OF IDIOMATICITY OR IDIOMATIC CLINE

Idioms are at the top of the scale. Hyperbole comes first with pure idioms in the second position. Approximately the middle of the scale is a mixture of metaphor ('dead duck', 'white lies'), simile ('as different as chalk from cheese'), pure idioms or idioms of decoding ('twist somebody's arm', 'in the dark'), and idioms of encoding ('off his hands'). The scale tapers off with items such as 'thank goodness' and 'sharing drinks' that verge on the literal and therefore are of marginal idiomatic status.

A different scale based on the structural properties of idioms rather than on pragmatic factors such as situational impossibility or absurdity may be considered. It is realized that the possibility of literal interpretation arising from an idiom having a homonymous counterpart is a factor dependent on the language-user's knowledge of the world. Literal or non-literal interpretation is a matter not only of the internal structure and correlation of the syntactic units (words, phrases and sentences) constituting the text of discourse, but also one of situational context. The possibility of literal or non-literal interpretation is a matter of beliefs and mental perceptions of the reader/hearer even more than of language structure and therefore variable, but such variables are an in-built factor in language use and interpretation. Fernando and Flavell (1981) proposed a ranking of idioms based on structural properties rather than on judgments such as 'absurd' but even the judgment 'more idiomatic' based on an item's possessing properties such as non-correlative syntax complement by homonymity must be seen in terms of a variable defining context. To use the terminology of the philosophy of language, idiomaticity involves us in the synthetic rather than in the analytic use of the language.
The scale of idiomaticity based on the structural properties of idiom is significantly different from that based on pragmatic considerations in one major respect: it gives priority to two-faced constructs, the two faces being the literal and the non-literal idioms over those with only one-face, either non-literal or literal. The scale ranges from constructs that have both literal and non-literal faces through those that are only non-literal to those which while showing a peculiarity of encoding are literal. Even this peculiarity of encoding is only semi-institutionalized, as in the case of predictable collocations. The items at the bottom of the scale are only marginally idiomatic. This second scale of idiomaticity also shows clearly that idiomaticity in a language is very much a matter of overlapping categories and intermediate zones.

In short, the following **characteristic features of idioms** are considered:

1. The meaning of an idiom is metaphorical rather than literal.
2. It is not a result of the componential function of the parts.
3. The grammatical form of an idiom is often invariable and fixed.
4. The process of substitution is not allowed.
5. The passive constructions cannot be formed.
6. Idioms vary a great deal on how metaphorical and invariable they are.
   In other words, idiomaticity (the quality of being idiomatic) is a matter of degree or scale.

### 1.7.3. CRITERIA FOR IDENTIFICATION OF IDIOMS

Makkai argues that it is not only more economical but also more insightful from the point of view of identifying idioms “to use the term idiom only for units realized by at least two morphemes” (Makkai 1972). Apart from this he also proposed five criteria:

1. The presence of at least two free morphemes in a given expression. As in the example of: ‘to hold one’s tongue’
2. The ability of these morphemes to function with different meanings in more than one environment.

   Example:
   i. ‘I held my tongue’
   ii. ‘We held our tongues’
   iii. ‘You held your tongue’
   iv. ‘She held her tongue’
   v. ‘He held his tongue’
   vi. ‘Hold your tongue’, etc.

3. The potential ambiguity of all idioms of decoding arising from the possibility of literal interpretation.

   Example:
   i. ‘To hold one’s tongue’ (literal interpretation of holding probably one’s tongue with the help of fingers, teeth etc.).
   ii. ‘To restrain oneself from talking’ (the semantic interpretation).

4. The semantic unpredictability of idioms arising from the fact that an idiom has a meaning that cannot be deduced from its component parts.

   Example:
   i. 'To have other fish to fry' (to have something to do that is more important or profitable)
      but * 'have other salmon to fry' or 'the other fish is to be fried'

5. Institutionalization. As the term suggests, it is the method in which the idiom is identified and freezed from its conflict between the literal meaning of the individual parts of the lexical unit and its semantic constituents.

In order to deduce a distinct definition for idioms, in this study, I have tried to pick up some such units that look like idioms and yet have a definite distinction as follows:
1.7.4. TYPES OF IDIOMS

The distinction between 'lexical idioms' and 'phrasal idioms' pervades all too often in the form of exclusive concentration on 'phrase idioms' (notably verb + object groups) and the exclusion or neglect of 'lexical idioms' (i.e. compound words, and noun + adjective, verb + particle and similar clusters).

Lexical idioms are listed as units in the lexicon. For phrase idioms, however, considerations of simplicity in syntactic and phonological description suggest that 'at least the members of the class of idioms whose occurrences also have compositional meanings must receive the ordinary syntactic structure assigned to occurrences of the stretches with compositional meanings' (Katz and Postal 1963).

Although not units in the syntactic lexicon, phrase idioms do have this status in the semantic dictionary. Indeed they make up a separate list, and are interpreted somewhat differently from 'lexical items'. It is suggested that syntactically deviant phrase idioms may be handled in the same way as semi-sentences (Wood 1986).

1.7.5. DISTINCTION BETWEEN COMPOUNDS AND IDIOMS

The prototypical compound is a word made up of at least two bases which can occur elsewhere as independent words, for instance, the compound ‘greenhouse’ contains the bases ‘green’ and ‘house’ that can occur as words in their own right (e.g. in the noun phrase, ‘the green house’, i.e., the house that is green). On the other hand, idioms are seen to form the end-point of a historical process by which word-combinations first establish themselves through constant re-use, then undergo figurative extension and finally petrify (Cowie, et al. 1983: xii).

Distinction between compounds and idioms:
Table 1.2. Distinction between compounds and idioms

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Idioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. They have a constituency at the word-level.</td>
<td>1. They have a syntactic constituency.</td>
</tr>
<tr>
<td>2. It is a word that consists of more than one lexical item, and which often triggers a sense other than the mere combination of its lexical units, for example: ‘candlelight’.</td>
<td>2. It is a unitary item, at least from the semantic point of view.</td>
</tr>
<tr>
<td>3. Among compounds can be found nouns, adjectives, verbs, adverbs, ad-positions, etc. for instance, ‘red-brown’ (adjective) and ‘air-condition’ (verb), and, very occasionally, among adverbs for instance, ‘everywhere’.</td>
<td>3. An idiom is a characteristic sequence of words that is particular and peculiar in nature and is often unique to that language.</td>
</tr>
<tr>
<td>4. They are non-unique instances.</td>
<td>4. They are unique instances.</td>
</tr>
<tr>
<td>5. Compounds are comparatively productive.</td>
<td>5. Idioms are completely un-productive.</td>
</tr>
<tr>
<td>6. Compounds are rule based.</td>
<td>6. Idioms are not derived by rule.</td>
</tr>
</tbody>
</table>

The meaning of a compound is generally not a mere sum of the meanings of its parts as ‘candlelight’ illustrates: ‘candlelight’ does mean ‘candle + light’, but something like ‘light’ from a ‘candle’ (Asher 1994). Thus, the literal meaning of only one or of neither element plays a part, are known as idiomatic compounds.

Consider some examples of such idiomatic compounds:

<table>
<thead>
<tr>
<th>Te.</th>
<th>Hi.</th>
</tr>
</thead>
<tbody>
<tr>
<td>grAmasiMhaM (grAmaM + siMhaM) 'village + lion'</td>
<td>kuwwA ‘dog’</td>
</tr>
<tr>
<td>kanuvippu (kannu + vippu) 'eye + open'</td>
<td>paraxA PASa ‘unveil’</td>
</tr>
<tr>
<td>Akupacca (Aku + pacca) 'leaf+ green'</td>
<td>harA ‘green’</td>
</tr>
<tr>
<td>pasupupacca (pasupu + pacca) 'turmeric + green'</td>
<td>pIIA ‘yellow’</td>
</tr>
</tbody>
</table>
Thus, compounds incorporate both semantic and pragmatic elements. If a compound is a curtailed sentence or phrase, it seems legitimate to derive it from its corresponding sentence or phrase (cf. Levi 1978). If it is a word, its underlying structure is fundamentally different, in particular lacking tense and modality (cf. Bauer 1978).

Katz and Postal (1963) define idioms in terms of non-compositionality of meaning, and then recognize that to be consistent, they must let this include the composition of words from morphemes. Bolinger (1975) argues that there is no clear dividing line between compounds and idioms.

We may conclude then that the compound words can in theory be idioms with morphological consistency.

1.7.6. DISTINCTION BETWEEN PROVERBS AND IDIOMS

Idioms may be seen as ready-made bits of language, and it is not always possible to draw a line between idioms, sayings and proverbs. With respect to the longer and more complex expressions in particular, it is hard to lay down rules or principles for usage, and this is yet another area of language where things simply have to be looked up and learnt.

Sayings and proverbs are perhaps not indispensable, but may be useful, and also fun to use, and they sometimes give some interesting insight into another culture. In translating idioms and proverbs one needs to be wary of the many differences that exist between different languages and cultures. Normally, proverbs are often used among peer groups. The use of proverbs may indicate superior native proficiency in the language. Often proverbs include idiomatic expressions to add authenticity and acceptance. Proverbs unlike other linguistic units like verbs, phrases, compounds and idioms have no specific status. They have only sociological, philosophical implications rather than linguistic. Thus, idioms vary from the proverbs in the sense that idioms are used colloquially. Proverbs are at times shortened and thus become idioms but it is not possible for idioms
to lengthen and be used as proverbs. Consider some examples of proverbs and also their idiomatic derivatives:

<table>
<thead>
<tr>
<th>Proverbs</th>
<th>Idioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>niMdu kuMda woNakaxu</td>
<td>niMdu kuMda</td>
</tr>
<tr>
<td>porugiMti pullakUra ruci annatlu</td>
<td>porugiMti pullakUra</td>
</tr>
<tr>
<td>kAIYlakiMxa nippulu posi nIYlu callinattu</td>
<td>kAIYlakiMxa nippulu</td>
</tr>
<tr>
<td>kukka woka vaMkara</td>
<td>kukka woka</td>
</tr>
<tr>
<td>gummadikAyala xoVMga aMte BujAlu wadumukunnatlu</td>
<td>gummadikAyala xoMga</td>
</tr>
<tr>
<td>koVMdanu wravvi eVlukanu pattinatlu</td>
<td>koMdau wravvu</td>
</tr>
<tr>
<td>aMxani xrAkRapalYlu pullana</td>
<td>aMxani</td>
</tr>
<tr>
<td>uppu, pulusu winna viSvAsaM</td>
<td>uppuwinu</td>
</tr>
</tbody>
</table>

The distinction between proverbs and idioms could be shown as follows:

<table>
<thead>
<tr>
<th>Proverbs</th>
<th>Idioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. They are clausal.</td>
<td>1. They are phrasal.</td>
</tr>
<tr>
<td>2. They are the constituents of discourse unit.</td>
<td>2. They are the constituents of syntactic units (sentence).</td>
</tr>
<tr>
<td>3. They may change from transparent to opaque.</td>
<td>3. They are always non-transparent.</td>
</tr>
</tbody>
</table>

Table 1.3. Distinction between proverbs and idioms

While some proverbs range from semi-opaque to opaque, as:

i. ‘don’t count your chickens before they’re hatched’;

ii. ‘he who pays the piper calls the tune’;

iii. ‘he who laughs last, laughs longest’, etc.,

others are quite literal, as:
1.7.7. DISTINCTION BETWEEN COLLOCATIONS AND IDIOMS

Collocations are collections of words that "fit together"; i.e. they are predictable patterns and phrases or sequences of words that we typically use together. They include what have traditionally been considered vocabulary items, as well as structural patterns and combinations of words that simply "go together".

By definition, collocation is ‘a composite of lexical items with a specialized, but strictly unpredictable meaning’. They are of roots, not of words, which are essentially means of reference (Mitchell 1971).

On observation, collocations (such as ‘ask a question’, ‘high winds’, ‘on foot’) are similar to idioms in that they involve relatively fixed sequences of words, but differ in that they are not recognized culturally or stylistically as expressions in themselves. Some linguists prefer to distinguish collocations as idioms on syntactic and semantic grounds. According to Cruse (1986), Benson et al. (1986) and others, collocations are syntactic units that can be broken down intuitively into smaller recognizable and independent semantic units (ask + a question, ask + the price). Other linguists refuse to distinguish between idioms and collocations, on the grounds that they often see one form as a subordinate category of another (e.g. Moon 1992, Fernando 1996, Gross 1996). Van der Wouden (1997), for example, states: ‘I will use the term collocation as the most general term to refer to all types of fixed combinations of lexical items; in this view, idioms are a special subclass of collocations, to wit, those collocations with a non-compositional, or opaque semantics.” (Van der Wouden, 1997).
Van der Wouden does point out that this entails problems. He cites the examples of commonly considered collocations such as ‘a murder of crows’ and thus idiomatic. Similarly, the formulation ‘ask for money’ is considered to be a collocation, although it is not completely compositional. The expression cannot be broken down further than: ‘ask for + money’, so that ‘for’ appears to be stuck, morpheme-like, to the verb.

The following table depicts the distinction of collocations from idioms as:

<table>
<thead>
<tr>
<th></th>
<th>Collocations</th>
<th>Idioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Pattern:</td>
<td>Ordered sequences.</td>
</tr>
<tr>
<td>2.</td>
<td>Structure:</td>
<td>No hierarchy, Flat.</td>
</tr>
<tr>
<td>7.</td>
<td>Distribution:</td>
<td>Frequent</td>
</tr>
</tbody>
</table>

Table 1.4. Distinction between collocations and idioms

By seeing idioms as essentially ‘marked’ expressions and collocations as ‘unmarked’ or normal means of expressing a concept, we are trying to make a distinction that is not categorical or binary and which lends itself to the notion of a continuum. Very common collocations such as Wierzbicka’s examples of prepositional phrases (‘in April’, ‘on Thursday’, ‘at ten o’ clock’), are clearly unique formulations in that the prepositions are obligatory for each formulation, but are also unmarked, standard ways of expressing those concepts. Bound collocations are a little more unusual (‘blond hair’) and yet these represent the preferred way of saying things in general discourse. To take Moon’s (1992) examples, ‘out of the blue’, ‘to call the shots’, ‘foot the bill’: all of these of course semantically opaque, but they are also marked forms of more prosaic formulations,
namely: ‘unexpected’, ‘to take command’, ‘to pay the bill’. These phrases are idioms, because they bring some rhetorical force to the basic expression (usually by the use of explicit metaphors: the first two expressions increase the intensity of the expression, while ‘to foot the bill’ also implies ‘a reluctance to pay’). At times, it may also be the case that there are a cluster of related core statements with no really central phrase (such as ‘finally’, ‘in summary’, ‘at last’) which coexist with more idiomatic expressions (‘at the end of the day’, ‘when all is said and done’, ‘all’s well that ends well’). It is also perhaps worth noting that collocations appear to be neutral in terms of style, whereas idioms can be seen as at times inappropriate in terms of formality.

Frequently paired words or a set of words usually get together to form collocations. Superficially, idioms also appear to be frequently paired words or a phrase with a set pattern in the wording. The difference between a collocation and an idiom, however, is the meaning. Often, the meaning of a collocation can be interpreted by combining the meanings of the separate words in the phrase; the meaning of the idiom is more than (and often quite different from) the meaning of the separate words within the idiom. Unlike idioms, individual words in a collocation can contribute to the overall semantics of the compound. So idioms like "take a break", structures like "If I had the chance, I would . . . "and word combinations like "get on a bus /get in a car" are all considered collocations. Sometimes the same set of words can function as an "ordinary" phrase, an idiom, or a collocation.

Compounds include word pairs that occur consecutively in language and typically are immutable in function. Noun + noun pairs are one such example, which not only occur consecutively but also function as constituents. Cowie (1981), notes that compounds form a bridge between collocations and idioms, since, like collocations, they are quite invariable, but they are not necessarily semantically opaque. Since collocations are recursive (ibid.), collocation phrases, including more than just two words, can occur. For example, a collocation such as by chance in turn collocates with verbs such as find, discover, notice etc. Flexible word pairs include collocations between subject and verb,
or verb and object; any number of intervening words may occur between the words of the collocation.

The examples of ‘to foot the bill’, ‘maMdi padu’, we are obviously approaching another transitional area bordering on idiom. These examples are semantically transparent. These are also un-idiom like in the fact that they are fairly freely modifiable (Cruse 1986).

1.7.8. DISTINCTION BETWEEN DEAD METAPHORS AND IDIOMS

From a linguist's point of view, idioms are complex lexical items, that is, word-like entities that have some of the properties of phrases. They are more than a word but less than a sentence. Some specialists often call them "dead metaphors". A 'dead metaphor' is one in which the sense of a transferred image is not present. Example: money, so called because it was first minted at the temple of Juno Moneta. (Shipley 1970).

An "idiom", on the other hand, is simply a construction that one cannot analyze by taking it apart. Like ‘get rid of’, for instance, or ‘lots of’, or ‘right off the bat’, or ‘take off’ (in either the sense of ‘an airplane’ or ‘a striptease’), or thousands of other familiar phrases in English that don't come apart into easily digestible pieces. These are chunks of the language that come pre-assembled, with their own meanings and grammar, and they're learned as such by native speakers. Many of them are metaphors, but not all. Some of them are slang, but most aren't. A "metaphor" is a species of meaning that is extraordinarily common in language, and is neither idiomatic nor slang, for the most part. A metaphor treats some aspect of meaning as if it were something else. When one says, ‘I spent an hour on this’, you're using the ‘Time is money’ metaphor theme, since ‘spend’ refers to money and you're using it for time. When you say ‘Stock prices went down today’, you're using the ‘Less is down’ metaphor theme, since prices are expressed in abstract numbers, while ‘up/down’ refers to gravitational effects on physical objects. Slang is full of metaphor, and so even the ordinary language has considerable amount of metaphor. Often, Metaphor is a common basis for idioms, but not all idioms come from
it, and most metaphoric usage are usually interpreted by native speakers as being ordinary, non-idiomatic prose.

While idioms and dead metaphors must be distinguished, it should also be recognized that they have certain characteristics in common. It is probable that the majority of idioms began their lives as metaphors; and synchronically, transitional cases, which are idioms for some and metaphors for others, are not uncommon (Long and Summers, 1979). Dead metaphors have in common with idioms that their constituent elements do not, in the straightforward sense, yield recurrent semantic contrast: consider for instance, the contrast “amAvAsya/pUrnima” in “amAvAsya caMxrudu”. They are not therefore, semantically transparent.

1.8. OBSERVATIONS

A literal rendering of an idiom is very rarely capable of serving as an approximate translation. It is most likely to be either non-interpretable or quite unrelated in meaning to the original expression. Even apparently simple translation equivalents of idioms are treacherously deceptive. Consider the Telugu idioms: “masipUsi mAredukAya ceyu” and “kAlu kAlina pilli”. A translation of the first is ‘to swindle’ and the second is equivalent ‘to roam in discomfort and aimlessly’. But neither of these when translated literally give the slightest clue to the idiomatic meaning of the original Telugu expression.

These difficulties experienced with apparently simple lexemes are all the more troublesome with regard to idioms. An idiom may be translated into the target language by means of the following with increasing order of preference and acceptance:

1. a paraphrase;
2. an idiom of similar meaning but dissimilar structure and/or lexical constituents;
3. an idiom of similar structure and/or lexical constituents but dissimilar sense;
4. a match structure, lexical constituents and meaning.

It is by no means uncommon for an idiom in one language to be at least roughly equivalent to a lexically unrelated idiom in another language.

Example:

Te.     Hi.
nippu peVttu       Aga lagAnA ‘to set fire’
kadupubbu        peTa PulanA ‘to swell (stomach)’
wala wirugu        sira cakarAnA ‘to swirl (head)’
goVMwu koyu        galA kAtanA ‘to cut (neck)’

In the above examples, the idioms mentioned in Telugu when translated into Hindi produce idioms that require no separate entry in the lexicon. Whether lexically unrelated idioms can ever be considered exact translation equivalents, however, is debatable. Literal translation fares rather better with dead metaphors; the results are usually a little bit odd, but are nonetheless interpretable in the manner of live metaphors, such as:

Te.     Hi.
gAnugeVxxu        kolhU_kA bEla ‘donkey’
patta pagalu       xina xahAde ‘broad day-light’
matti suxxa        gobara ganeSa ‘stupid’
rAwi guMdeV        pawWara_kA kalejA ‘cold hearted’
SilAkRaraM        pawWara_kI lakIra ‘golden rules’

Interestingly, a high proportion of dead metaphors have similar (although not often identical) dead metaphor equivalents. These close equivalents among dead metaphors can present the translator with a dilemma (one or many!). If the translation is word-for-word, the achievement has greater fidelity in one respect but to the detriment of fidelity in another respect, wherein, the source language idiom is a live metaphor and the target
language idiom is not; if, however, a greater value on the latter type of fidelity is put, then a sacrifice of the former is a must.

Not surprisingly, dead metaphors as a rule present fewer problems to foreign learners of a language than idioms do. Their interpretability, however, must not be exaggerated; their meanings are not necessarily wholly predictable on first acquaintance. Indeed, some can only be appreciated as metaphors with hindsight, as it were. It is only when the figurative meaning is pointed out that the path from the literal to metaphorical meaning becomes traceable.

Similarly, since compounds, proverbs, and collocations are often language-specific and cannot be translated compositionally in most cases, researchers have expressed interest in statistical methods that can be used to extract bilingual pairs of each one of these for parallel and non-parallel corpora. Note that one cannot assume that a concept expressed by way of a collocation in one language will use a collocation in another language.

How one translates an idiom into another language depends on the idiom stock in the receptor-language. The rhetorical effect of a given idiom in the source language as opposed to the receptor-language and the text that is being translated are crucial in rendering idioms appropriately. On the whole the rhetorical effect of semantic idioms has a better chance of being retained in a translation than that of structural idioms. Translation is an exacting art. Idiom more than any other feature of language demands that the translator be not only accurate but highly sensitive to the rhetorical nuances of the language.

What is most noteworthy in the brief contrastive survey of idiom is that languages drawn from different language families, while showing dissimilarities, also show many similarities. The similarities are reflected in structural processes, in common idiom types (homonymous idioms, metaphor, hyperbole, set phrases, etc) and in parallel cognate idioms, while the dissimilarities show up in different rhetorical functions and different connotations being attached to given lexemes and their referents. All in all, it seems that
the similarities are more striking than the differences. Equally striking is the fact that the idiomatic use of language is such a widespread phenomenon. Languages which are only idiomatic inter-lingually (i.e., by virtue of having a distinctive structural cut in contrast with other languages are rare. Thus, idiom is a near-universal of language.

1.9. METHODOLOGY

The intention of the work presented here is to set the foundation for a further intensive study in identifying an idiom, to enlist the defining criteria of idioms and distinguish them from the other language forms such as compounds, proverbs, collocations and metaphors. We would focus idioms in the context of larger than the sentence in order to prove their idiomaticity and try to distinguish them from the other language structures viz. compounds, proverbs, collocations and dead-metaphors in particular to be used in the machine translation system as a component of translator's tool.

For this, we begin to compare the views presented by some prominent Indian grammarians then analytically discuss some of the important definitions of the late 20th century linguists. For the identification of idioms, we have sieved and compiled the various criteria that would become the basis for further work.

We needed to consider both western and Indian authors definitions regarding idioms and their observations in identifying the idioms. It was necessary to know the nature of idioms, their features, their classification.

We also have sought for the sources of idioms, their features, and the classification of Telugu idioms (Vijayalakshmi, 1998).

This is followed by a collection of data of idioms from various Telugu-Hindi and Telugu-English dictionaries. For this purpose we took the help of Telugu-Hindi anusaaaraka Machine Translation Dictionary developed at CALTS, University of Hyderabad (cf. Uma Maheshwara Rao, 1999), Vemuri Radhakrishnamurthi’s Telugu-Hindi Dictionary (2002),

The collection of data was extensive but we selected only those idioms that are ambiguous and which posed problems while translating them. The study was dealt from various angles viz. morphological, syntactic, semantic angles and also on the level of their frequency. They are again studied on the pattern of transfer that posed problems in getting the exact meaning transferred from the source language, in this study Telugu to the target language, Hindi. From this wide source of data, I have manually selected around 190 idioms that form part of my third chapter 'Appendix’.

As part of my fourth chapter, the idioms listed in the data were subjected to critical scrutiny to identify the issues in identifying or recognizing idioms and their idiosyncratic sense when they are ambiguous. Out of this data, I selected approximately 85 idioms for analyzing their idiomaticity in terms of their semantic features such as + / -human, + / -animate, + / -concrete, + / -abstract, etc. Again, these idioms are studied using a sentence as context by observing the strings of words that occur either as prefixing or suffixing the idiomatic verb. Here, each idiomatic verb is checked for it's intransitive nature and then each idiomatic verb is studied against its agent [Ag], the subject and theme [Th], the object for the basic information. After an extensive study, I came with the classification of idioms that in turn helped in categorizing them under different headings.

As a final step in identifying a given sequence of words as an idiom, we looked for the need of information, i.e., a context larger than the sentence that I took up in my fourth chapter. This required me to study a Telugu corpus containing 3 million word strings. From this corpus we extracted word strings in the form of ‘n-grams’, which are then segregated as bi-grams, tri-grams, tetra-grams. We extracted 2.26 million word strings of ‘bi-grams’, 2.17 million word strings of ‘tri-grams’ and 1.8 million word strings of ‘tetra-grams’. These provided me to identify the minimum probability of length for an idiom to occur in a sentence.