Chapter 8

Pravṛtti-nimitta and Śabdasūtra

8.1 Pravṛtti-nimitta

In Indian Grammatical Tradition, we see abstractions at phonological, morphological, and morphophonemic level as well as at semantic level. The abstraction of the Sanskrit words rāmeṇa and ramayā as rāma + ṭā and ramā + ṭā is an example of suffix level abstraction. Postulation of dhātus (verbal bases) in various classes gaṇas of the dhātu-pāṭhā is an example of abstraction at the level of morphology. We find similar abstraction at the level of semantics in the concept of pravṛtti-nimitta.

All the Indian schools admit a relation between a word and its meaning. This relation is termed as śakti. However, these schools differ in the nature of this relation. According to the mīmāṃsākas, the relation is natural\(^1\). The Grammarians also accept this view\(^2\). They believe ‘just as the organs of perception have an innate capability to perceive, similarly the words have an innate capability to express the meaning\(^3\). The

\(^1\)mīmāṃsāsūtra 1.1.5 : autpattikas tu śabdasyārthena sambandhāḥ.
\(^2\)siddhe śabdārthasambanddhe – Patañjali in mahābhāṣya.
\(^3\)indriyānām svaviśayesu anādiḥ yogyatā yathā | anādiḥ arthāy śabdānām sambandho yogyatā tathā ||(vākyaapadiyam III 3.29)
naiyāyikas however are more realistic. They do not accept this as a natural relation. According to them, this relation is not ‘natural’ or ‘siddha’ but established by somebody. Naiyāyika’s classify the relation into two types: primary and secondary. The primary relation further is of two types – a) saṁketa or īśvarecchā⁴ and b) paribhāṣā. For example, the relation between a word ‘cow’ and an object cow or the relation between a technical term such as a ‘circle’ and a real object circle are primary relations. Of these, the first one is termed as a saṁketa, since nobody knows how this relation came into existence. In other words they are arbitrary. On the other hand, the technical terms are coined by somebody and hence the relation of a technical term to its object is termed as a paribhāṣā. To account for the metaphorical usages of a word, naiyāyikas admit a secondary relation, also known as laksanā.

It is the delimiter (avacchedaka) of this relation (śakti) - śakyatavacchedaka, which causes a proper understanding of the meaning of a word. Śakyatāvacchedaka is also known as ‘pravr̥tti-nimitta’ - the reason (nimitta) that warrants the use (pravr̥tti) of a word to specify the desired meaning⁵.

8.2 Example

The concept of Pravr̥tti-nimitta helps one in understanding the underlying ‘unity’ among the seemingly different meanings of a word, and thus accounts for the polysemy. Let us take an example of the word hari in Sanskrit. Amarakośa lists 16 different meanings of this word (See fig 8.1).

Then how do we account for these seemingly different meanings? The word hari is

---

⁴Earlier naiyāyikas considered this relation to be īśvarecchā, but later naiyāyikas termed it as a saṁketa

⁵pravr̥ttaḥ śabdānāṁ artha-bodhana-śakteḥ nimittam prayojakam (Nyāya kośa, pp 580)
derived from the verbal root *hr* which means *harane* (to steal). Thus according to this derivation one who steals is *hari*. And now everything falls in place. Viṣṇu is called hari because he takes away (steals) the sins of his devotees, monkey is called hari because it steals the eatables, a lion is called hari because he takes away one’s life, and so on. Proper use of this pravṛtti-nimitta also gives a scope for creative use of such words.

8.3 The ‘core meaning’ in Modern Linguistics

In modern linguistics, though we find abstraction at the level of form (in terms of phonemes and morphemes, etc), abstractions at the level of semantics are very rare. Words are typically treated as polysemous and thus various meanings are just listed in the dictionary. This approach fails to give a holistic view of the meanings of a word. There are few exceptions, however [78].
“A monosemic approach, on the other hand, posits only one abstract sense for a word, and the various contextual meanings or readings that the word receives are predictable from features of the context in which the word occurs [82].”

For example, Saxton gives various senses of *just* as below.

- I live *just* around the corner. (in the sense of immediacy)
- I *just* got home from school. (in the sense of recently)
- Tina looks *just* like her mother. (in the sense of exactness)
- Your beet salad is *just* delicious. (in the sense of very)
- May be it is *just* hormonal. (in the sense of only)
- I can *just* squeeze through that fence. (in the sense of barely)

Saxton, by postulating the “core” meaning of *just* as *without deviating*, extracts the appropriate meanings from the features in the context. This concept of monosemic approach, thus, comes very close to the concept of Pravṛṭti-nimitta.

Of late there have been attempts by the dictionary makers also to show how various meanings of a word are related, thereby, giving a holistic view of the various meanings a word expresses (see appendix D). Such an explanatory note helps a secondary language learner to acquire the ‘seemingly diverse meanings of a word’ with less effort.
8.4 Relevance of pravṛtti-nimitta in developing a Language Accessor

In anusāraka, the first layer provides the śabdāśūtra of each word. This śabdāśūtra is developed on the basis of the pravṛtti-nimitta of the source language word. It is not an easy task to decide the ‘core’ meaning of a polysemous word. Though many times the etymology helps, factors such as diachronic changes, borrowing of words from other languages - sometimes leading to homonymy, make it difficult to discover the ‘core’ meaning. The task becomes more difficult when the ‘core’ meaning is to be expressed through other language. This then calls for a mechanism to handle the differences between two languages with regards to the packaging and labeling of concepts at lexical level, as discussed in chapter 3. These differences lead to five different combinatorial possibilities. We show in what follows the solution anusaaraka provides in each of these cases, except the one-to-one mapping, of course.

8.4.1 Lexical Gap

The lexical gap may be present either at a concept-word level or at the function-word level. If it is at the level of a concept-word, we borrow the word as it is, till an equivalent in the target language is defined / coined and is in usage. For example the technical terms such as ‘digital electronics’, are borrowed into Hindi as it is, and displayed in the Devanāgarī script. The persons who read texts with such words are either familiar with the terms, or the texts they read define and explain these technical terms. And thus no more explanation of these terms is necessary.

Here is an example of an introduction of a technical term.

These satellites have devices which can receive signals from an earth
station and transmits them back in various directions. Such a device is called a TRANSPONDER.

In the anusaaraka output of such sentences, the word TRANSPONDER, for which there is a lexical gap in Hindi, will be borrowed into Hindi and transliterated into Devanāgarī script in the output. But if such a word represents some concept which needs cultural background, then anusaaraka interface pops-up a footnote providing necessary description.

If such a word is a function-word, then either we borrow the word, or use special markers and provide a help on the function it serves. For example, there is no functional equivalent to the English determiners in Hindi. So we borrow them and also provide a note on their usage. But in case of Telugu adverbial ‘gā’, instead of borrowing the suffix, we put a mark ‘*’ as its Hindi equivalent, and provide an online help (see appendix C).

8.4.2 Many-One mapping

One may say many-one mappings do not pose any problem as far as the translation is concerned. It is not so. The many-one mapping may lead to either mis-interpretation or an ambiguity. For example, as is already noted in chapter 3, the following English sentence leads to ambiguity when translated into Hindi.

Eng: He gave a flower to her.  (1)
Gloss: usane diyā eka phūla usako.  (1a)
Hin: usane usako eka phūla diyā.  (1b)

From the Hindi translation it is not clear who gave a flower to whom, whereas in English it is clear that it is the male person who gave a flower to the female person. The source of ambiguity in Hindi is the third person pronoun because Hindi does not
mark third person pronouns for gender. The verb agreement in Hindi gives the gender information. Thus the two languages differ in ‘where’ the information is coded. Since the parsers are not reliable and anusaaraka claims 100% information preservation, anusaaraka proposes mappings for the pronouns from English into Hindi as shown in table 8.4.2.

Thus the anusaaraka output for the above English sentence would be:

\[ \text{anu: usane} \{ \text{pu.} \} \text{ diyā eka phūla usako} \{ \text{strī.} \} \].

If we describe the meaning of a word by a two dimensional region on a plane, then the union of 4 non-overlapping regions corresponding to the words *he*, *she*, *it* and *that\(^7\) of English correspond to a region *vaha* of Hindi. Using the nyāya terminology, the meaning of ‘He’ may be described as ‘pumāṅgatvāvacchinna vaha’(third person pronoun, delimited by masculine-ness), which we denote by *vaha*\{pu.\}.

### 8.4.3 One-Many mapping

When a word in English maps to more than one word in Hindi, we either try to express the underlying concept in words or represent all the disjoint meanings by ‘\(^\wedge\)’,

\(^7\)the dash sign after vaha in the meaning of that is used to indicate that it is an adjective in this context.
as follows:
Uncle: māmā cācā phūphā mausā
cousin: bhāī bahana \{asahodara\}

In the first example, we have listed all possible meanings joined by ‘~’, to indicate that uncle may mean any of these. In the second case, we could have listed out all possible relation names such as cacerā (paternal), mamerā (maternal), phūpherā (of husband of father’s sister), etc. denoting its meaning. But since we have a word asahodara which means not born of the same womb, asahodara bhāī athavā bahana will be an exact equivalent of cousin. But since asahodara is an adjective, and is not used in normal conversation in Hindi, we provide the equivalent of cousin as bhāī bahana with asahodara within \{\} providing an extra information. Here the meaning of English word is an union of more than one non-overlapping regions in Hindi. The information within \{\} serves as an extra information which delimits the meaning. In nyāya terminology, then, the meaning of the English word cousin may be described as asahodaratvavacchinnabhāīyābahana.

8.4.4 Overlapped regions: A real challenge to develop śabdasūtra

Real challenging are those cases where one can not express a region in English as a union or intersections of various regions in Hindi, but needs much more complex sets of operations.

Let us look at the English word ‘light’.

The dictionary entries for the word ‘light’ from the online English-Hindi dictionary\(^8\) are:

light Adj. 1. halakā-

Example: This suitcase is light and good.

\(^8\)available at http://ltrc.iiit.net/onlineServices/Dictionaries/DictFrame.html.
light N. 1. prakāśa
Example: I could see light in the room.

light V.
– 1. jalānā
Example: Mohan lighted the match-stick.
– 2. sulagānā
Example: Mohan lighted the cigarette.
– 3. prakāśita karanā
Example: The torch lighted the way for him.
– 4. prasanna honā
Example: His face lighted up when he heard the news.

The meaning of ‘light’ as an adjective is totally different from the meanings which correspond to the nominal and verbal usages of it. Thus this is a clear case of homonymy. It is obvious from the examples that the meanings of light as a noun and as a verb are related. But still, Hindi reader will definitely be puzzled by the range of usages of light as above and will fail to see the underlying ‘thread’ that connects these meanings. Here is an attempt to develop a thread running through different meanings of the word light.

It is very important to note that we are neither claiming to give a logical justification of why one may use the word ‘light’ in all the above senses, which for a Hindi speaker look as totally unrelated at the first instance, nor are we developing any theory for how the usage of a word and thereby its meaning gets shrinked or widened diachronically.
We are looking at the English usages and trying to establish a thread, which might be an artificial one, which may not reflect the correct historical developments, but which helps a Hindi reader to get a handle to relate seemingly different meanings and remind him at an appropriate time so that he can understand the English text correctly.

In the first example, *to light* may mean an activity that results in producing a *light*. According to vaiyākaranas a verbal root (dhātu) represents an activity (vyāpāra) and a resultant of the action (phala). Typically when the sense of a verbal root gets extended, the extension covers the similarity of an action or the similarity of the resultant. For example, lighting of the cigarette does not result in the production of light, but the underlying activity (of using the match stick to produce the flame and using it to light the cigarette) is same as that of lighting of the lamp. Lighting of a torch may not have the same underlying activity, but the result of the action is same, that is, it produces light. Similarly the presence of light illuminates the path and hence one can understand the usage *to light the way*. Finally the *merriment or spark* on the face may be explained as the metaphorical use of light.

So the **underlying thread** may be explained as

\[
\text{prakāśa} \Rightarrow \text{prakāśita karanā} \Rightarrow \text{prajvalita karanā} \Rightarrow \text{prakāśamāna karanā} \Rightarrow \text{jalānā} \ (2)
\]

Having arrived at the **thread** capturing all the senses, now we try to capture its essence in the form of a **concise expression** as

\[
\text{prakāśa} \ [^* \ karanā]/\text{halakā} \ (3)
\]

The ‘*’ indicates that prakāśa may take optionally some derived suffix. Note the placement of halakā as the second alternative and with ‘/’ to mark the homonymy.
This concise expression\(^9\) in (3) as well as the thread\(^10\) in (2) is called a śabdasūtra.

Thus a śabdasūtra stands for both a concise formula to represent the meaning and also a thread showing the connections between various senses of the meaning, which are apparently unrelated from the target language speaker’s point of view. Around 1000 śabdasūtras are developed for high frequency English polysemous words.

Sometimes, the ambiguity is only at the word level, and not at the base (prātipadikā) level. For example the word *leaves* has two analyses:

leaves : leaf, n, pl, and
leaves : leave, v, 3, sg

In such cases, we provide a ‘sūtra’ at the word level, such as

leaves = pattā{ba.}/choda

The ‘/’ indicates that the meanings are unrelated.

### 8.5 Guidelines for developing śabdasūtra

The step by step procedure for developing the śabdasūtra is given below.

- Collect different senses of the word for which śabdasūtra is to be developed. Various monolingual and bilingual dictionaries, various electronic resources such as WordNet, etc. may be used for this purpose.

- Collect at least one example sentence for each of the described senses.

- Translate the sentences into the target language.

\(^9\)alpāksharam asandigdham sāravat vishvatomukham |
astobham anavadyam ca sūtram sūtrandā viduḥ ||

\(^10\)is called a sūtra in Sanskrit
• Try to minimise the number of target language equivalents to the maximum possible extent.

• Develop a thread linking all these minimised target language equivalents.

• Describe the connections in the thread.

• Develop a concise expression, which serves as an handle to refer to all these senses.

Various notations used in developing the concise formula are provided in the appendix B. The appendix E and F contain śabdasūtra for the words ‘as’ and ‘case’ respectively.

8.5.1 Śabdasūtra and fidelity

Now we come back to the original question: In anusaaraka how do we ensure that a Hindi reader has complete ‘access’ to the English text?

The śabdasūtra in the form of a concise formula is used in the first layer of the anusaaraka output. Every word is split into a root and a suffix (prakṛti and pratyaya). We provide the meaning of each of these as a śabdasūtra. The underlying assumption is that the meaning of a word can be composed from its constituents. In case the meaning is non compositional, we provide a śabdasūtra for the complete word. So the first time reader will have to go through the threads understanding the explanation. In the user interface therefore,

• We provide the śabdasūtra for each word.

• We also provide additional help necessary to understand the sūtra. This help consists of
8.6 Śābdabodha in anusaaraka

The process of śābdabodha describes the steps involved in ‘understanding’ the meaning of a sentence after listening to it. Anusaaraka produces the output in several layers. We illustrate below with examples from English-Hindi anusaaraka, how the layers produce the śābdabodha of the English sentence.

• Example 1:

Eng: Rats kill cats.
anu-Hin: cūḥā{ba.} māra{0} billī{ba.}

An anusaaraka reader who has undergone the training on English grammar from Hindi viewpoint, would remember about the special vr̥tti, viz. V , i.e., a transitive verb requires two arguments, whose positions are fixed, and are to its immediate left and immediate right. The argument on the left is the subject and the one on the right is the object, and that the subject is abhihīta. With this knowledge, then he establishes the relations between different words as, cūḥā is the abhihīta, the verb is in active voice, and thus now appealing to the knowledge of Hindi grammar, the abhihīta should be kartā and hence should take nominative case, and the other word, viz. billī should be the karma, and hence should take accusative case. And thus a anusaaraka reader understands the English sentence through Hindi as

Hin: cūhe mārate haim billiye m ko.
Chapter 8: Pravṛtti-nimitta and Śabdasūtra

• Example: 2

Eng: Shankar was returning home on his bicycle after a football match.

The first layer output of this sentence, where each word is split into a stem and an suffix, and their meanings are provided is shown below.

anu-Hin: Shankar thā vāpasa_lauṭa2 {ing} ghara on{→ para} vaha{pu.}_kā sāikila after{→ ke_bāda[pīche]} a phuṭabālā maicåjoḍa.

The ‘2’ in the meaning of the verb return indicates that the verb can be both intransitive as well as transitive. Accordingly, in Hindi a reader has to choose the correct alternative between lauṭa and lauṭā. The → indicates that the preposition should not be read ‘in situ’, but should be placed at an appropriate position to the right. Thus a user has a heavy load in interpreting the meaning of this sentence:

– Deciding the correct meanings of the words return, after, match, and the verbal suffix ing, and

– Deciding the preposition phrase boundary.

At this stage the reader may look at the second layer output of the same sentence, to reduce the burden of interpretation (with a possible cost of misinterpretation).

anu-hin-layer2: Shankar — vāpasa_lauṭa2{0_rahā thā} ghara on{→ para} vaha{pu.}_kā sāikila after{→ ke_bāda[pīche]} a — phuṭabālā maica.

At this stage, the words ‘are returning’ and ‘football match’ are grouped together. Hence the burden of disambiguation of the word match and the suffix ing is reduced at this stage. The expectancy now comes into play and helps in deciding the meaning of the verb return. The verb return as an intransitive
verb (in the sense of lauṭa) has an expectancy of an object, but as a ditransitive verb (in the sense of lauṭā) has expectancy of a direct and an indirect object. The absence of an indirect object then leads to the choice of lauṭa.

Yogyatā (competency) helps in the proper choice of meaning of the preposition after viz. ke bāda. Ke pāche is used when the noun specifies a location in the space whereas, when the noun indicates an event or time, then the meaning is ke bāda. The reader on the basis of football match, which is an event, therefore chooses ke bāda for after. Yogyatā also helps further to decide the preposition boundaries of the preposition on.

8.7 Conclusion

In this chapter we have seen the innovative use of the concept of pravṛtti-nimitta to represent the meaning of a word in the target language. The concept of śabdasūtra is evolved as a means to express the pravṛtti-nimitta in a compact form. We have further shown with examples, how the śabdasūtra together with the notion of ākāṅkṣā and yogyatā help in getting the meaning of the sentence. Assuming that the reader of the target language shares the world knowledge, domain knowledge etc. with the source language reader, and taking into consideration the incompatibility between the source language and the target language in coding the information at various levels, the notion of śabdasūtra plays an important role in the formulation of image of source language words in the target language.