Chapter 7

Dorr’s Divergence and Anusaaraka

7.1 Dorr’s Divergence

In Chapter 3 we have discussed the divergences between English and Hindi from the information coding point of view. The much discussed translation divergence is due to Dorr [32]. Dave et al [30] discussed these divergences with respect to English and Hindi, and Goyal et al [37] discussed them with respect to English, Hindi and Sanskrit. In this chapter we will look at these divergences to see how anusaaraka handles these divergences.

Dorr classifies the divergences into two major classes viz. syntactic and lexical semantic. These are further classified as:

- Syntactic divergence
  - Constituent Order divergence,
  - Adjunction Divergence,
  - Preposition stranding Divergence,
  - Movement Divergence,
Chapter 7: Dorr’s Divergence and Anusaaraka

- Null Subject Divergence,
- Dative Divergence and
- Pleonastic Divergence.

- Lexical Semantic Divergence:

The lexical semantic divergence [32] discussed in Dorr and Dave et al [30] in particular with reference to English and Hindi are

- Conflational Divergence,
- Structural Divergence,
- Categorial Divergence,
- Head swapping Divergence and
- Lexical Divergence.

7.2 Anusaaraka solution

Anusaaraka being a language accessor, we look at these divergences from comprehensibility point of view. The important layer of anusaaraka is the first layer i.e. the layer of śabdasūtra. We address the issue of divergence with respect to this layer. In case, this layer can not handle the divergence leading to uncomprehensible output, we show how the next layer(s) handle it.

- Syntactic divergence

  - Constituent Order divergence:

This divergence concerns with the word order in English and Hindi. As we saw earlier in chapter 6, we train the anusaaraka reader in the special vṛtti -V- English exhibits. So in case of parser failures, the reader can still resort to the śabdasūtra layer and understand the meaning.
Chapter 7: Dorr’s Divergence and Anusaaraka

- Adjunction Divergence:
  This is associated with the positioning of the adjective phrase. Hindi is flexible with two types of adjective phrase constructions, viz.
  
  Hin: vaha ladakā jo kala yahāṁ āyā thā (1)
  gloss: the boy who yesterday here came

  Hin: kala yahāṁ āyā huā ladakā (2)
  gloss: yesterday here came boy

  whereas English has only one construction viz.
  
  Eng: The boy who came here yesterday (3)
  which corresponds to (1). Thus English does not have an equivalent construction for (2).

  So when we translate from English into Hindi, there is no problem. Problem arises when we translate from Hindi into English. Hindi allows an adjectival participial construction, and English does not have a parallel construction. This case is exactly parallel to the case of missing adjectival participials in Hindi corresponding to the ones in Dravidian languages. Anusaaraka [70], discusses a solution to this problem by suggesting a special ‘jo*’ construction in Hindi.

- Preposition Stranding:
  We saw in chapter 6 (examples 43 and 49) that there is more affinity between a preposition and a verb than between a noun and its preposition, leading to preposition stranding. This is thus a part of training on special sannidhi in English between verb and its argument markers.

- Null Subject Divergence:
  Hindi is a null subject language whereas English obligatorily requires a subject. So translation from Hindi to English poses a problem. But through anusaaraka complete information is available. For example, look at the following anusaaraka output from Hindi into English.
Though the subject is missing, for an anusaaraka reader, the information regarding the subject is still available indirectly through the verb features, and as such this does not pose any problem, at the level of śabdasūtra.

- Pleonastic Divergence:
  This divergence occurs with syntactic constituents having no semantic contents such as dummy *it*, *there*, etc. The anusaaraka reader is trained for this divergence giving logical reasons behind the presence of this divergence, or rather presence of dummy *it* and expletive *there*.

- Lexical Semantic divergence

- Conflational Divergence
  This divergence arises when lexicon incorporates periphrastic mechanism to fill in gaps in lexicalisation describing an action. For example *stab* means *to strike with a knife*. Hindi does not have a word conveying the same meaning, and thus uses a phrase *chūre se mūranā*, which is a literal translation of *to strike with a knife*. At śabdasūtra level, it is perfectly all right having a phrasal equivalent for a word. In fact the śabdasūtra may be much more complex than just a phrase.

- Structural Divergence
  Consider a sentence
  Eng: He entered the house.  (5)
  Hin: usane gharaṃ me praveśā kiyā.  (5a)
  Here *the house* is a NP, whereas the phrase in Hindi *gharaṃ meṃ* is a prepositional phrase. This is taken care of in anusaaraka by vibhakti transformation rules for different verb frames. For example a transformation rule
in this case will be
enter → *praveśa karanā<obj:merī>*

- Categorial Divergence:

This arises because of change in the lexical category during translation. The example discussed in Dave [30] is

Eng: They are competing.

(6)

In anusaaraka, *are competing* will be treated as a verb group, with *compete* as the verbal base and *are*-*ing* as a suffix. Thus the categorial divergence problem disappears!

- Head Swapping:

Under this, two cases of swapping viz. demotional and promotional swapping are discussed. The examples discussed in Dave [30] are:

Eng: It suffices. (demotional head swapping), and

Eng: The play is on. (promotional head swapping).

(7) (8)

In both of these cases, anusaaraka does not see any problem. In *It suffices*, *it* is dummy. Thus, with the meaning of *suffice* as *kāphī honā*, there is no problem of head swapping. Similarly, in the sentence *The play is on*, the śabdasūtra layer may not give a comprehensible output, since this is a kind of idiomatic expression with preposition providing the aspect of continuity, treating *is on* as a meaningful unit, the word group level provides the desired output. Thus there is no question of Head swapping.

- Lexical Divergence:

The example discussed under this category, in Dave [30] is

Eng: John broke into the house.

(9)

This also is not problematic. At the level of word grouping, the particle *into* will be grouped with the verb *break* to yield *break_into* as a meaningful unit, and providing its Hindi equivalent.
Thus we see that all these divergences can be taken care of at the śabdasūtra level or at the level of word grouping.

7.3 Problematic Cases

There are certain typical syntactic construction in English, which cause a divergence between English and Hindi. They are:

- Resultative constructions such as
  
  She washed the plates clean. \hspace{1cm} \text{(10)}
  
  The river froze solid. \hspace{1cm} \text{(11)}
  
  The gardener watered the flowers flat. \hspace{1cm} \text{(12)}
  
  They painted the house green. \hspace{1cm} \text{(13)}

- Verbs of motion specifying the manner of motion as in
  
  She danced into the room. \hspace{1cm} \text{(14)}
  
  She smiled her thanks. \hspace{1cm} \text{(15)}
  
  She beamed her adoration. \hspace{1cm} \text{(16)}

- The absolutive constructions such as:
  
  Julie accepted the award, tears streaming down her face. \hspace{1cm} \text{(17)}
  
  The cat confronted the burglar, its back arched and teeth bared. \hspace{1cm} \text{(18)}

In all such cases, plan is to provide an online intelligent help. The interface will be intelligent enough to identify the difficult constructions and provide necessary help.