Chapter 6

English from Hindi viewpoint: A Paninian Perspective

6.1 Introduction

With the world wide web spreading all over the world, information is now available at the click of a mouse. Most of this information is in English. In India hardly 10\% of the population can understand English. Thus the language barrier leads to a digital divide. Hence, if India has to take real advantage of the new technology, it is necessary to make this information available to the Indians in their native languages. Experiments have shown that English-Hindi anusaaraka, as a tool to access English text has been useful [13]. Since the top layers of the anusaaraka output, produce an output that follows the grammar of the source language, it is necessary for a serious user to ‘learn’ the ‘contrastive grammar’ between the source and the target language. A courseware, explaining the differences between English and Hindi will therefore be needed to access anusaaraka.

\(^1\)source http://en.wikipedia.org/wiki/List_of_countries_by_English_speaking_population
In Chapter 5 we have seen the information centric point of view of Pāṇini towards analysis of a language. In this chapter we use Pāṇinian way of analysis to discover the structural differences between English and Hindi. This study will not only be useful for an anusāraka reader, but also to the Machine Translation community working on English-Hindi translation systems, since it will throw open the problematic cases in English to Hindi Machine Translation.

6.1.1 Traditional view

The structural differences between English and Hindi are mostly attributed to the differences in their word orders. Language typologists\[8\] classify English as an SVO language and Hindi as an SOV language. However comparing English and Hindi on the basis of word order is like comparing apples with oranges! The reason is: English uses position to code crucial information of the relation between the words in a sentence. So when one says English is an SVO language, one is asserting a fact about the encoding of grammatical relations, viz. subject and object, with respect to verb in English. The position immediately preceding a verb marks the subject and the one immediately following marks an object. On the other hand, in case of Hindi, a relatively free word order language and claimed as an SOV language, one is just stating a statistical fact about the order of words in a typical Hindi sentence!

To make the point clear, the following two English sentences have exactly opposite meanings.

Rats kill cats. \( (1) \)

Cats kill rats. \( (2) \)

whereas, the following two Hindi sentences with similar change in the order of words as above, have the same meaning (ignoring the topicalisation, of course).
Chapter 6: English from Hindi viewpoint: A Pāṇinian Perspective

In the following sections, we investigate the reasons behind the structural differences between these two languages from information coding point of view. In the second section, on the basis of the basic structure of declarative sentences in English and Hindi, we conclude that English does not have a morphological formative for an accusative marking. The missing accusative marker is compensated by freezing the subject position in English. As a consequence, this fixed subject position gives rise to some structural differences between the two languages. The third section discusses these structural differences. In the fourth section, it has been pointed out that, English does not have a morphological formative marking the yes-no question. Therefore English resorts to the word order again. The structural differences arising because of this phenomenon are discussed in the fifth section.

6.2 Missing accusative marker

Look at the following English sentence and its Hindi gloss

Eng: Rats kill dogs. (5)
Hin gloss: cūhe māra\textsuperscript{2} kutte. (5a)

\textsuperscript{2}0 stands for no overt suffix. The verbal form with 0 suffix in English has more than one interpretations – non 3\textsuperscript{rd} person singular, or to-less infinitive, etc. The one which is relevant in this particular case is present tense non 3\textsuperscript{rd} person singular.
One may interpret the above Hindi gloss as

cūhe mārateḥaṁ kutte. \hfill (6)

Though this is not a grammatical Hindi sentence, still, if a Hindi reader is asked to interpret this sentence, he will interpret this as

Hin: kutte mārateḥaṁ cūhori ko. \hfill (7)
gloss: dogs kill rats{acc.} \hfill (7a)
Eng: Dogs kill rats. \hfill (7b)

which is exactly the reverse of what is being said in English!

Why does this happen? First let us try to understand the reason for why a Hindi reader analyses it in this way, and later we will see what mechanism in English triggers the desired meaning.

6.2.1 Arguments by Hindi reader for this interpretation

The Hindi sentence (5a) is ungrammatical, because Hindi requires an accusative marker (ko) to mark the karma role. However, Hindi also has a tendency to drop the karma vibhakti, wherever there is a possibility of recovering the information from other sources, such as world knowledge. For example, consider the sentence (8) below.

Hin: rāma phala khāṭā hai. \hfill (8)
gloss: Ram fruit eats. \hfill (8a)
Eng: Ram eats fruits. \hfill (8b)

In sentence (8), phala does not have an accusative marker. In spite of this, a Hindi
reader appealing to the world knowledge (in this case the yogyatā or competency), interprets this sentence as rāma is the kartā of the action of eating and phala is the karma of the action.

At the same time, Hindi obligatorily requires an accusative marker, if anything against yogyatā is to be communicated, as is obvious from the following Hindi sentence.

Hin: latā śarāba nahīṁ pīṭī, śarāba latā ko<sup>3</sup> pīṭī hai. (9)
Gloss: Lata wine not drink, wine Lata-{acc.} drinks. (9a)

Following the same argument, since sentence(5a) does not carry any explicit accusative marker to mark the karma, Hindi reader appeals to his world knowledge, since the dogs have yogyatā to kill the rats and not the other way, interprets the sentence as (6).

Unlike Hindi, English does not have an explicit morphological formative for accusative marking. Rather it codes the information to indicate grammatical relations of subject and object by their positions.

Now one may ask the question, which position is crucial, the subject position or the object position or both? In other words, what is invariant, the S-V order or the O-V order or the S-V-O order?

6.2.2 Initial Hypothesis (S-V-O order)

Both the subject as well as the object positions carry crucial information, hence the S-V-O relation is invariant in English.

However, we come across such sentences as

<sup>3</sup>the marker ko with latā, is to mark latā as a karma. This ko can not be dropped.
Who likes sweets? **Sweets**, I like.  

Mrs. Venables turned a little pale. 
Lord Peter presented no difficulties, but **Bunter** she found rather alarming. 

(source: D. Sayers, The Nine Tailors)

In sentences (10a) and (10b), the object is in the topic position, i.e., topicalised to mark the contrast.

Therefore, in this case, the information that **sweets** is an object of the verb **like**, is not coded in position or the order. In other words, the orders S-V-O as well as V-O are not invariant.

### 6.2.3 Revised Hypothesis (S-V order)

This leads us to reframe the observation as, 

**It is the subject-verb order which is invariant. Objects may move around.**

But life is not simple as is expected. There are examples showing movement of subject also!

Here are the examples:

Uneasy *lies the head* that wears a crown.  

Never *was the sea* so calm!  

Here *comes the bus*!  

On the bed, *hung a mosquito net*. 

In the above examples, the **head, the sea, the bus and a mosquito net** are the subjects of the verbs *lies, was, comes*, and *hung*, and do not precede the verbs.
However, in all these examples, the verb is monovalent. That is, they have an expectancy of only one argument, which agrees with the verb in number and person. Therefore, its position in a sentence is not crucial. However, in case of transitive verbs, there are two arguments, and hence it is necessary to mark at least one of them. From the above examples, what we observe is:

In case of transitive verbs, subject is always to the left of the verb, or in other words, S-V order is invariant!

There is evidence, which goes against this hypothesis too. Look at the following sentences

Something had to give. And give it did.  (12a)
Last October our good friend in South Africa, wanted to come to England this year and come he did, with his wife Annie.  (12b)
We all said she was bound to leave him, and leave him she did.  (12c)
She could only hope that Harriet was mistaken in his feelings...

Wish it she must, for his sake... (J. Austen, Emma)  (12d)
Ride in a taxi with Pamela and Bredon he could not, even if it meant losing her forever... (D. Sayers, Murder must Advertise)  (12e)

In all these examples, the subject is after the verb phrase! But at the same time, we also note that subject is followed by an auxiliary! Finally it precipitates (to the observation) that it is the subject-auxiliary proximity (sannidhi) that is invariant in English. The normal proximity between auxiliary verbs and the main verb gets violated in English, and a new proximity is established between a subject and an auxiliary verb.

This leads to a concept of ‘Subject Position’ - a position which is to the immediate left of the auxiliary verb, or the main verb (in case auxiliary verb is absent). And
thus, we revise our observation as:

6.2.4 Final Observation

In case of transitive verbs, the missing accusative marker in English has been compensated by a position called ‘Subject Position’.

We see below the consequences of the missing accusative marker.

6.2.5 Consequences of missing accusative marker

Difference in word order

Major consequence of coding the information in terms of position is reflected in the difference in word order.

- Hindi has post positions, while English has pre-positions

  e.g. at the door -> daravāje\textit{para}

  Leaving apart the subject position, the remaining part of English sentence structure reflect the mirror image of the corresponding Hindi sentence structure.

  e.g. look at the English sentence and its Hindi translation in table 6.1.

<table>
<thead>
<tr>
<th>Eng:</th>
<th>There$^1$</th>
<th>is$^2$</th>
<th>somebody$^3$</th>
<th>knocking$^4$</th>
<th>at$^5$</th>
<th>the$^6$</th>
<th>door$^7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gloss:</td>
<td>0</td>
<td>hai</td>
<td>koi</td>
<td>khaṭakhaṭatā haū</td>
<td>para</td>
<td>–</td>
<td>daravāje</td>
</tr>
<tr>
<td>Hin:</td>
<td>daravāje$^7$</td>
<td>para$^5$</td>
<td>khaṭakhaṭatā haū$^4$</td>
<td>koi$^3$</td>
<td>hai$^9$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1: English-Hindi: Mirror reflection

- English uses phrasal verbs formed of verbs followed by particles giving different shades of meaning. These particles are in post verb position, whereas the upasargas in Sanskrit (and in Hindi also) have the same role, but are used as prefixes. For example,

  Eng: look at, look for, look after, etc.
  Hin: āhāra, vihāra, prahāra, etc.
Chapter 6: English from Hindi viewpoint: A Pāṇinian Perspective

- The order between main verb and the auxiliary is reversed. In Hindi the verb groups are formed by the main root followed by auxiliary verbs, as in \( jā \) \( rahā \) \( hai \). However, in English the order is: auxiliary verbs followed by the main verb as in \( is \) \( going \).

- It should be noted that the reversal is not everywhere. For example, in case of word formation, at morphological level, the order is still root or stem followed by a suffix, as we see in the words \( going \), \( goes \), \( chairs \), etc. Similarly in case of noun groups, the relative order of adjective and noun is same as in Hindi. For example: compare the red book with \( lāla \) \( pustaka \).

**Exceptional Case Marking (Subject-Object raising)**

Another phenomenon in English is raising of subject to object position or also termed as exceptional case marking (ECM).

Consider the sentence:

Eng: I want [him to go there].

gloss: maiṁ cāhatā hūṁ usako jānā vahāṁ.

From Hindi point of view, there are two problems in this sentence.

- In case of icchārthaka (indicating desire etc.) verbs in Sanskrit (and also in most Indian languages), if the subordinate verb has \( tumun \)(infinitive) suffix, then it shares the kartā with the icchārthaka dhātus (\( samānakartṛkeṣu tumun \), Pāṇini: 3.3.158). For example,

San: aham bhoktum icchāmī.

gloss: I eat to desire{1p sg}.

Eng: I want to eat.
Hence, to express a reading such as

\[ I \text{ want [him to go there].} \quad (15) \]

wherein the \textit{kartā} of want is different from the \textit{kartā} of \textit{go}, Hindi can not use \textit{nā} (tumun) construction to convey this reading. Hindi either uses a finite clause such as

\[ \text{main ċāhatā ġūm ki vaha jāye.} \quad (16) \]

or, uses non-finite clause such as

\[ \text{main ċūakā jāna ċāhatā ġūm.} \quad (17) \]

- Second problem with this construction is, this structure is inherently ambiguous. Consider the sentence,

\[ I \text{ want this pen to write.} \quad (18) \]

The most likely meaning is

\[ I \text{ want [this pen] [to write].} \quad (19) \]

and not

\[ I \text{ want [this pen to write].} \quad (20) \]

as in (9). It is the world knowledge which triggers the desired meaning.

The question is what phenomenon in English does make such constructions inherently ambiguous? Secondly why the \textit{he} which is the subject of \textit{go} has an accusative marker?

Answer to the first question is not very difficult. It is the subject of the infinitive verb which also happens to be the object of the main verb that makes the
sentence inherently ambiguous.

This construction in generative grammar has been termed as Exceptional Case Marking (ECM) or raising to object phenomenon. *He*, which is at the subject position of *go*, gets case marked by the verb *want*. This is because, *go* not being in finite form, can not assign case to its arguments whereas, *he* being also in the object position of *want*, *want* can assign a case to *he*, making it *him*. Here it should be noted that the word *him* should not be analysed as *he* + accusative marker. *Him* here is just an oblique case of *he*, a morphological requirement. Thus, a verb is assigning case to something which is not its own argument. That is why such verbs are referred to as Exceptionally Case Marking verbs. This phenomenon is also sometimes referred to as subject-object raising, since the subject of the subordinate clause has been raised to the object position of the main verb. This explanation makes one feel that the sentence formation is *asamartha*, and in spite of it, a native speaker finds the sentence to be good. So there should be some *gamakatva* (the ability to convey the desired meaning) because of which an English reader is able to get the correct meaning. This *gamakatva* is in the subject position of the verb *go*. 

Figure 6.1: ECM phenomenon in English
Subject sharing (gapping)

The quest ‘where’ is the information coded in a English lead us to the concept of Subject Position. ‘How’ is information coded is equally important. Whether the information is coded explicitly or implicitly has direct bearing on the parsers. If the information is coded explicitly, then it is available directly for parsing. However, if it is coded implicitly, following language convention, then one has to extract this information indirectly. Further, if other language does not follow the same convention, it may lead to catastrophe as well.

Here is an example where English and Hindi exhibit different behaviour in the sharing of kārakas. This may lead to improper or misunderstanding of English sentence by Hindi native speakers.

For example, consider the sentence

Eng: Mohan dropped the melon and burst. \hspace{1cm} (21)

The Hindi gloss of this is

Gloss: mohana ne girāyā tarabūja aura phuṭā. \hspace{1cm} (21a)

The kartā of phutanā is missing, and hence appealing to the yogyatā, a Hindi reader interprets this as

Hin: mohana ne tarabūja girāyā aura tarabūja phuṭā\(^4\). \hspace{1cm} (22)

whereas the English sentence (21) means

Mohan dropped the melon and Mohan burst. \hspace{1cm} (23)

\(^4\)Hindi, unlike English, does not have a subject sharing rule. Hindi allows such usages as rāma ne subaha kapade dhoye aura dopaharataka sukha bhī gaye.

In this sentence, the kartā in the second sentence (kapade) is same as the karma in the previous sentence.
Though this sentence may sound senseless to an English reader, still s/he can’t get any meaning from this sentence other than the above one. To get the other meaning viz. that the melon burst, English has to use another construction viz.

Mohan dropped the melon and it burst. \(24\)

### 6.3 Missing yes-no interrogative marker

Look at the following two sentences in English:

Eng: Ram is going to school. \(25a\)

and

Eng: Is Ram going to school? \(26a\)

The first one is declarative and the second one is a yes-no question. If we look at the words in the sentences, they are the same, except for the word order. So it is natural that the information of ‘interrogativeness’ or ‘declarativeness’ of the sentence is in the word order. There is no explicit morpheme to mark the interrogativeness.

From the Hindi translations of these sentences

Hin: rāma skūla jā rahā hai. \(25b\)

and

Hin: kyā rāma skūla jā rahā hai? \(26b\)

it is clear that Hindi has an explicit word ‘kyā’ to mark the ‘yes-no’ question. The counterpart of this morpheme is missing in English. As a consequence, English codes this information in the form of ordering of words.
6.3.1 Observation

The missing marker corresponding to yes-no question is compensated by the ‘subject-auxiliary verb inversion’ in English.

As a consequence the normal proximity between auxiliary verbs and the main verb is weakened, and a new proximity is established between the subject and the auxiliary verb.

6.3.2 Consequences of Missing yes-no interrogative marker

- Subject Position can’t be empty:

For, if it were empty, it would not be clear whether the given sentence is interrogative or declarative.

- Insertion of auxiliary do in interrogatives: If a verb form does not involve an auxiliary verb, then a dummy do is inserted, as shown below.

He goes to school. \hspace{1cm} (27)

Does he go to school? \hspace{1cm} (28)

Here goes is split as does+go by introducing an auxiliary do, and the auxiliary does then is inverted with the subject to give an interrogative sentence (28).

- extra overheads:
Since English codes information in Subject Position as well as in subject-auxiliary order, Subject position can’t be empty. This forces English to bear an extra overload of dummy *it* and existential *there* to fill the Subject Positions.

– Dummy/expletive *it*

Consider the following English sentences

It is raining. (29)

It is very hot outside. (30)

and their Hindi translations

bārisha ho rahī hai. (31)
bāhara bahuta garamī hai. (32)

Hindi translations do not have any counterpart of the dummy *it*. This *it* is termed as expletive or dummy *it*. As the name implies this *it* does not carry any information, and is just a place holder or a stand-by.

– Expletive *There*

Consider the following sentences

There are flowers in the garden. (33)

There could have occurred several riots. (34)

There are thought likely to be awarded several prizes. (English Syntax, Radford, pp 246) (35)

In these sentences, the subject position is occupied by the word *there* and the word with which verb shows agreement (*ukta*) is moved away from the subject position. This *there* is not an adverbial there, since one can say
There are flowers there.  

(36)

The second *there* in the above sentence is an adverbial *there*.

Hindi translations of above sentences are

- bagice_mem phula haini.  
  (37a)
- kaai darse ho_sakate_the.  
  (37b)
- kaai paritoshoḵin kā vitaraṇa hone kī saḿbhāvanā vyakta kī jā rahī hai.  
  (37c)

In Hindi translations we do not see any counterpart of English *there*.

*There* in all these cases is also called an existential *there*, since it expresses the existence or appearance.

Why does English require this existential *there*?

The words or phrases that are to be focused are normally moved into a focus position at the front of a clause in order to highlight it. When a verb is to be focused, it is not possible to bring it to the front, since it either renders the subject position empty or it may read as an interrogative sentence. Hence in such cases, the subject position is filled with an expletive *there*.

It is interesting to note that there are certain transitive verbs which use expletive *there*. For example,
There entered a hall an ugly old man. (Levin, 1993:p90) (38)

This *there* thus serves as a focus element to express the ‘factuality’ or ‘happeningness’ of the event.

### 6.3.3 Subject-Subject raising

Since the dummy *it* does not carry any lexical information, it is an extra overhead. So there is a tendency to drop it whenever such an opportunity exists. This is natural and consistent with the principle of economy. For example, consider the following English sentences.

It seems that the boys have eaten fruits. (39)  
The boys seem to have eaten fruits. (40)

These two sentences are equivalent. The phenomenon where the subject of the subordinate clause has been moved to the subject position of the matrix clause, is known as subject-subject raising in western grammar.

This phenomenon is purely structural/syntactical and has nothing to do with the semantics. Once the subject of the subordinate clause is moved to the subject position of the main verb, the main verb shows agreement with this noun. Thus we see that, in sentence(40), the *boys* is in the subject position of *seem* and also is *ukta*, as it shows agreement with the verb *seem*. However, semantically, *the boys* is not a *kāraka* for the verb *seem*. It is the *kartā* of *eat*!

This inconsistency in the sentence structure thus should render the sentence incompetent (*asamartha*) to convey the desired meaning. Or in other words, the sentence has certain syntactic operations that do not convey any semantic connections. But
then, how is that an English reader does not find such constructions odd? or how is that the sentence is acceptable to the native speaker? How does a language allow proper communication in spite of apparent inconsistency at structural level to convey the desired meaning?

It may be said that it is the gamakatva (arthabodhakatva, the ability to convey the desired meaning) that takes care of proper communication. Patañjali has discussed this aspect under the commentary of samartha padavidhiḥ(Pañini: 2.1.1). He takes an example of Sanskrit compound devadattasya gurukulam (school of D’s teacher). Devadatta is related to guru. So semantically, it should be analysed as in figure1 of 6.2. However, its syntactic decomposition is as described in figure 2 of 6.2.

This is possible only because, the word guru is sākāmksha(having expectancy). Hence
even if it joins with other word to form a compound, this śākāṅkṣhatā (the property of having expectancy) makes it possible to convey the desired meaning. This śākāṅkṣhatā is the gamakatva which makes it possible to interpret the given compound properly.

On similar lines, in case of (40) the verb eat requires a subject. But its subject position is empty. At the same time, seem does not need any subject, and its subject position is occupied by the boys. The gamakatva is in the fact that seem does not require any subject. It expects only a clause following it.

So English native speakers seems to push the occupant of the subject position of the verb seem to the right towards the first subordinate verb whose subject position is empty. Thus for a native English speaker, the subject position is more important than the agreement between the verb and the subject.

6.3.4 Tough movement

There is a class of adjectives (tough, difficult, easy, hard, etc) which also exhibit a phenomenon of raising. This phenomenon is often referred to as tough movement. Look at the following pair of sentences.

It is hard to see John.  (41)
John is hard to see.  (42)

As is clear from the above example, the object (John) of the subordinate verb is moved to the subject position of the main verb (is) which was occupied by dummy it. This is again a case of asāmartya (incompetency), since there is an agreement between the occupant of subject position (John) with the verb (is), whereas the occu-
pant of subject position is not an argument of the verb!

Not only the objects, but even the complements of prepositions can also move to the subject position. However, when the complement of prepositions move to the subject position, the prepositions are left behind. This then gives rise to ‘violation of normal sannidhi/expected proximity’. For example

This violin is tough to play these sonatas on. (43)

where, the normal sannidhi between preposition (on) and the noun (violin) is violated.

Further with the believe type of verbs, English allows constructions such as

John is tough to believe that University would fire. (44)

Students are tough to believe that University would fire. (45)

Here also, the agreement is just a structural requirement and does not carry any semantic information. The gamakatva is in the ākāṅkṣā of the adjective tough and the verb in the subordinate clause.

6.3.5 Wh questions

The phenomenon of subject-auxiliary inversion is also observed in wh- questions, where the wh-element is brought forward to the topic position in order to focus. Hence, with an exception of subject of wh- questions (where the wh is already in focus), the wh questions also show a ‘subject auxiliary inversion’. Here are some examples:

Whom did Ram kill? (46)

Where did Ram go? (47)

with an exception of wh-questions on subject, as in
Who killed Ravana?

Further in case of wh questions on NPs which are objects of a preposition, the prepositions normally are not moved along with the wh elements, as in

Who did Ram talk to?

English does allow pied piping wherein the sannidhi between prepositions and it’s object is intact as in

To whom did Ram talk?

However, this order weakens the focus on wh element. The preferred order in English, is as in (49) with stranded preposition and not as in (50).

6.3.6 Inversion in tagged questions

English does not have any separate morpheme for tagged questions, hence again, it resorts to the subject-auxiliary inversion, as in

He has gone to Mumbai. Hasn’t he?

He won’t win. Will he?

6.3.7 Inversion in other constructions

Phenomenon of subject auxiliary inversion is seen in other constructs also.

Ram has gone to the market. So has Shyam.

Only then, did he understand the joke.

No other colleague, would I trust.

Suffice it to say that ...
(53) is an example of gapping phenomenon. It is also interesting to note how the word order helps in triggering the correct sense of the word *so* between its two senses, viz. *therefore*, and *also*. Compare (53) with

\[
\text{Ram has gone to the market, so Shyam was waiting for him. (57)}
\]

(54) and (55) are examples of focus on the factuality, and (56) is an idiomatic expression.

### 6.4 Conclusion

The purpose of the foregoing exercise is to look at the structural differences between English and Hindi from an information theoretic point of view. The major reason behind the structural differences between English and Hindi is the absence of accusative marker and yes-no marker in English. To compensate for this absence, English resorts to the word order. This further gives rise to more structural differences between the two languages. The interaction between different phenomena have been explained in figure 6.3.

We conclude that a Hindi reader while reading an English text has to ‘tune’ himself to the following:

- Acquire a new ‘vṛtti’ – the ‘quazi compound’ \_V\_.
- Do away with the normal ‘sannidhi’ (proximity) between a verb and its auxiliary and also between a noun and its post-position (which are integral part of Indian languages), and acquire new ‘sannidhi’s between:
  1. a subject and auxiliary and
  2. a verb and its preposition,
Figure 6.3: Contrast between English and Hindi
and finally,

- The occupant of subject position need not have any kāraka role with the corresponding verb.

We discuss, with examples, how this grammar helps an anusaaraka reader to understand the text, in chapter 8.