An Insight into Ordering of Linguistic Units for Translation
(The Case of English-Hindi)

Sanjay Kumar Jha

Director and Professor (Amity School of Liberal Arts)

Amity University Gurgaon, Haryana, India

Abstract

As a well-ordered or well-organized text is easier to read and comprehend, it is essential to for a human translator or a machine translation system to understand the right ordering of four major units of language while translating one language into another. The objective of this short communication is to restrict its concern to the ordering of English text only especially in terms of body, paragraph, sentence, clause, phrase, and word.

Discussion:

The phenomenon of ordering can be viewed with respect to SL, TL, and SL-TL. At this level, the ordering is concerned with SL. Since a well-ordered or well-organized text is easier to read and comprehend, it is essential to discuss the ordering of six major units of language as follows:

<div align="center">

Body
Paragraph
Sentence
Clause
Phrase
Word

</div>

As far as the ordering of six above-mentioned units is concerned, only human is able to order all the six units with greater precision. As for machine, only the ordering of clause, phrase, and word seems feasible. The ordering of body, paragraph, and sentence (in terms of cohesion and coherence) seems rather difficult, as machine will require high degree of precision in terms of

algorithmic linguistic conditions. Getting such a task done by machine may seem difficult but not impossible. The details of such issues fall outside the scope of the present research. The present research, in this respect, explores merely the possibility of such task to be implemented in the field of machine translation under the purview of advancements in text linguistics and cognitive science.

Ordering as a part of SL is concerned with understanding the organization of text in terms of cohesion and coherence amid larger units of translation, i.e. body (set of paragraphs), a paragraph, and a sentence.

Body

The body may minimally consist of a set of paragraphs in the form of an article or an essay or any piece of message; and maximally in the form of a thesis, dissertation, treatise, or a book. Since body is the maximal stretch of a translating text, it meets all the required conditions of textuality of (Beugrande,)especially in terms of cohesion and coherence.

Before venturing upon the task of body ordering, it is important to understand the types of text, which can be broadly viewed from two perspectives, viz., function, and domain.   A translator ought to be sensitive to the types of text (e.g. expressive, informative, and vocative), as ordering (e.g. climactic, spatial, chronological, psychological or topical and so on) may differ as per the types of text. It is the functional features of the text, which decide the ordering of the text. For example, one should discard the ordering of body if it is of expressive type. From the perspective of translatability, the ordering of body is of least significance. The ordering of body is a matter of priority in terms of chapterization or giving prominence to an event or a piece of message. Since such issues falls under the art of writing, so they are being skipped here. What matters above all by the ordering of body is, a translator should have a high degree of understanding of the overall text and organize it in certain chronological order to establish logical coherence amid events stated in the whole body of the SL text (English).

Paragraph

Stretches of language longer than a sentence is called paragraph. Text as the realization of body is the maximum stretch in the form a set of coherent paragraphs, whereas text as the realization of single paragraph consists of a set of coherent and cohesive sentences. Truly speaking, it is the level of paragraph, which most of the translators attempt to comprehend and subsequently translate.

But it is an irony that grammatical research has, for ages, been focusing more on clauses and sentences than on discourse. The reason is, as Roger Bell (1995) says, "the regularity in text structure is considerably of less obvious nature". Therefore, the contrastive grammatical competence used for the purpose of translation yet falls under sentence or word level grammar rather than text grammar. But in recent years after the advent of **text- linguistics,** the attention shifted towards text in the form of paragraph(s). It has also been found that like sentences, texts also have systems, which bind them together, and give them shape and unity between their boundaries. Such a binding is termed *discourse structure*. The translators now vested their interest in issues related to text-cohesion, text-coherence, and text-style, etc.

Unlike body, paragraph or discourse is more prone to be ordered. Ordering of paragraph has two connotations here. First is from extra linguistic perspective, which can be viewed under the purview of writing system, and the second is from intralinguistic perspective, which is concerned with **cohesion** and **coherence.** As a part of the first connotation, a translator is required to see the ordering of paragraphs (rather than paragraph) in terms of logical effect or prominence based on the number of paragraphs. In case of restricted number of paragraphs, the sequence has to be traditionally unmarked as (1) Introductory (2) Descriptive (3) Concluding. But what matters above all is prominence as per the type of text. For example, in narrative discourse prominence is given to major events. But in procedural discourse, the goal is of greater prominence than the steps leading to the goal and so on.

Ordering of paragraph as a part of second connotation is concerned with **cohesion** in terms of transitional elements. By transitional elements, a translator comes to know the connections of meaning between preceding and following parts of the paragraph. Some of the main transitional elements within the paragraph are conjunctions, certain adverbials, pronouns, deictics, etc. An example of cohesion in terms of reference is being cited here from Newby (1988:60).

1.   Yesterday, Mary and her husband went to buy a picture.
2.   In the little gallery above a restaurant in the center of town, they browsed among the watercolors by local artists, the small oil sketches of poppies and primroses, the bold landscapes and the chaste interiors.
3.   He suggested they buy a finely-detailed study of rocks on the crest of a dark hillside.
4.   Her preference, however, was for the flowing lines of a lithograph showing a local church.
5.   They spent an hour there in pleasant disagreement.
6.   He wanted to postpone a decision until another day.
7.   She didn't.
8.   Eventually, they both decided on a small hand-coloured etching of trees against a winter sunset.
9.   She proposed that they go to buy another picture at the local Art Society exhibition on Saturday.
10.   He didn't want to do so as they had spent too much money already.

"They" in sentences (2), (3), (5), (8), (9), and (10) refers to Mary and her husband; *"He"* in sentences (3), (6) and (10) refers to "husband". "*her*" in sentences (1) and (4) refers to Mary; so does "*she*" in sentences (7) and (9). "*There*" in sentence (5) has as its referent "*the little gallery above a restaurant in the centre of town*" in (2) sentence. "*didn't*" in sentence (7) refers to "*wanted to postpone*". "*So*" in sentence (10) refers to "*go to buy another picture at the local Art Society Exhibition on Saturday*" in sentence (9).

Such a system of reference makes a web of connection from one part of discourse to another. Cohesion can be useful in terms of solving pronominal reference to avoid anaphoric ambiguities.

Sentence

The purpose of sentence ordering is to stress on the fact that sentences must occur in logical order, which gives meaning flow in a pargraph. If the SL sentences are placed in haphazard sequence, it may create havoc by making the TL text not only incomprehensible but also the SL text untranslatable as follows:

One line of approach would be to assume that our standard family will not change its demands from year to year. Whoever wants it can have it. Stunned by the blow, the cat collapsed. If the boxers are well-matched, an exciting contest can be safely predicted.

Individually each of these sentences makes perfect sense. But, in combination, they are not making any sense because of random ordering. Therefore, a translator ought to look into the flair of co-text and context of the SL text in order to establish a logical coherence in a text. A human translator can think consciously of the logical order of the co-texts but an MT system is far behind in understanding logical order of the events, as it will require high degree of world knowledge to be programmed.

Besides evaluating the ordering of body, paragraph, and sentence, the translator (especially an MT system) should evaluate the order of clause, phrase, and word in order to ensure the wellformedness of the SL input text. The following are the brief explanation of each.

Clause

The issue of clause ordering arises when a sentence has more than one clause especially in the case of complex and compound- complex sentences. For this, an MT system needs to find clause boundary, and order the clauses in terms of focus and emphasis before translating it. The lack of such evaluating competence may not only cause poor translation but also no translation, as shown in the translation output of three MT systems below. Resorting to lexical translation by S1 and S2 indicates that the systems lack required linguistic knowledge base as to separating clausal boundaries and order them in logical order.

English: That you hate her is known to everybody.

Natural Hindi: *sabko patā hē ki tum usse naphrat karte ho*

S1	< yaha traansaleta nahii ho paayaa, dobaaraa phrema kariye - shaabdika artha - vaha tuma ghriiaa usa hai jnaata taka pratyeka aadamii > .
S2	! vahii aapa ghRnd-aa karataa hei vaha pratyeka vyakti ko jaanii gayii hei
S3	Unable to translate

Phrase

For the ordering of phrase, an MT system should assess two facts- identifying phrase boundaries (within the clause) and establishing logical sequence of the phrases irrespective of the phenomena of inversion, focus, or emphasis. For example the sentence below:

All alone in her kitchen my aunt was making a little tea quite slowly.

can be split with different possible choices, but it's our grammatical sense which not only appeals to organize words under five types of phrase (i.e., Adjective Phrase, Preposition Phrase, Noun Phrase, Verb Phrase, and Adverb Phrase) but also order them in a logical sequence as follows:

(All alone)	(in her kitchen)	(my aunt)	(was making)	(a little tea)	(quite slowly).
Adj Phrase	Prp Phrase	NP	VP	NP	Adv Phrase

The idea of identifying phrasal boundaries is highly useful in terms of parsing by both human and machine.

Word

The order of words needs to be minutely assessed not only within the phrase in terms of hypotactic and paratactic relation but also within its logical occurrence with respect to other constituents of the sentence in terms of case relations. The failure of understanding these issues may produce unintended meaning in TL text. An example from *The Oxford Dictionary of English Grammar*, (1994:192) will make the point clearer.

A light green shirt

Light is hypotactic to green rather than shirt. The shirt itself is not light, whereas in

An expensive green shirt

Both the adjectives (i.e. expensive and green) are in paratactic relation.

As for ordering in terms of case relationship, the phrase can be overtly coordinated as "the shirt was expensive and green".
The translator should ensure whether the string of words conform to the unmarked word order of English or not.

References:

Beaugrande, R. (1981). Introduction to Text Linguistics. Longman
Bell, R. T. (1995). Translation and Translating. Longman, New York.
Chalker, S. & Weimner, E. (1994). *The Oxford Dictionary of English Grammar*, OUP, New York.
Newby M. 1988. The Structure of English. Cambridge, Great Britain.