ICON - 2008

6th International Conference on Natural Language Processing

Details Of The Selected Paper

Title	**Comparative Analysis of phrasal cohesion of English-Hindi with English-French**
Topic	**Machine Translation**
Abstract	**In this paper, we study the property of phrasal cohesion for the English-Hindi language pair. The property of phrasal cohesion states that the linguistic phrases in source language translate as a unit in the target language. The study of phrasal cohesion is important in the context of various applications such as Machine Translation. We compare the degree of phrasal cohesion of English-Hindi and English-French language pairs. We have empirically verified that the degree of phrasal cohesion of English-French language pair is more than the phrasal cohesion of the English-Hindi language pair.**
Authors	**Vineet Yadav Language Technologies Research Centre, International Institute of Information Technology, Hyderabad Sriram Venkatapathy Language Technologies Research Centre, International Institute of Information Technology, Hyderabad Karthik Gali Language Technologies Research Centre, International Institute of Information Technology, Hyderabad**
Contact	**vineet.yadav.iiit@gmail.com**
Download	

Close

Comparative Analysis of phrasal cohesion of English-Hindi with English-French

Abstract

In this paper, we study the property of phrasal cohesion for the English-Hindi language pair. The property of phrasal cohesion states that the linguistic phrases in source language translate as a unit in the target language. The study of phrasal cohesion is important in the context of various applications such as Machine Translation.

We compare the degree of phrasal cohesion of English-Hindi and English-French language pairs. We have empirically verified that the degree of phrasal cohesion of English-French language pair is more than the phrasal cohesion of the English-Hindi language pair.

1 Introduction

In this paper, we *study* the property of phrasal cohesion for the English-Hindi language pair and *compare* it with the degree of phrasal cohesion between English and French. English and French are closely related languages and it is well-known that they have a low degree of language divergence when compared to English and Hindi which are known to be highly divergent languages. One of the components of language divergence is the property of Phrasal Cohesion. It is an important parameter to study, with use in practical applications like Machine Translation. As expected, we verified empirically that phrasal cohesion of English-French was higher than phrasal cohesion between English and Hindi.

Typically, in an Machine Translation (MT) system (both Rule-based and Statistical), phrasal cohesion is assumed. The property of phrasal cohesion states that a linguistic phrase in the source sentence translates as a unit in the target language. This means, for a source language phrase, the property of phrasal cohesion is violated if the translation of an external phrase is embedded within the translation of the phrase in consideration. It is convenient for the MT systems to assume phrasal cohesion because it limits the number of transformations of the source syntactic tree to be explored in order to obtain the target language word order. The number of transformations are limited as the existence of phrasal cohesion ensures that the source and target syntactic structures are isomorphic to each other ie., the dependency links between the source words and the links dependency between the translations of source words remain the same. The transformations permitted in such cases are just the reordering (or ordering if the syntactic tree in source language does not have order information) of the nodes of the source syntactic tree with respect to their parents. Hence, it becomes extremely important to study the prevalence of such an assumption (of phrasal cohesion) in a particular language pair (especially between language pairs English-Hindi where there is a large degree of word-reordering). We present our observation regarding phrasal cohesion between English-Hindi in detail in this paper.

The phrasal cohesion between a language pair is computed by measuring the number of violations of the property of cohesion. The violations are known as crossings. A crossing is said to have occurred between a phrase pair if the translations of the phrases have overlapping spans in the target language sentence. For example, consider a source sentence '$a\ b\ c\ d\ e$' which has been translated to the target sentence '$t_b\ t_a\ t_d\ t_c\ t_e$', where the translations of valid source phrases '$a\ b\ c$' and '$d\ e$' are '$t_b\ t_a\ t_c$' and '$t_d\ t_e$' respectively. Here, the source phrases '$a\ b\ c$' and '$d\ e$' cross each other. As we can observe, maintaining the cohesiveness of each of the phrases '$a\ b\ c$' and '$d\ e$' will not allow the appropriate translation '$t_b\ t_a\ t_d\ t_c\ t_e$' to be generated.

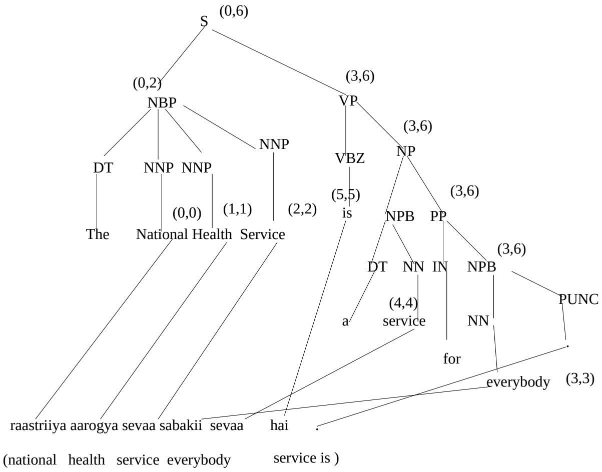

Figure 1: Crossings in tree

For this study, we need to extract syntactically valid phrases in the source language sentence. We use a source language parser for this purpose. Each node (representing the sub-tree rooted at a node) in the parsed output denotes a valid phrase in the source sentence. For example, a couple of valid source phrases in Figure 1 are "The National Health Service" and "for everybody.". We experiment with both a phrase-structure parser and a dependency parser to compute the phrasal cohesion. To check the existence of a crossing between two phrases, the span of their translations is computed and it is checked if their spans overlap. The spans of the phrases are computed by using a word-aligned corpus. The word aligned corpus is used to project the words of the source phrases to the words of the target sentence. From Figure 1, the spans of the source phrases (0,2) are (3,6) respectively. We have organized the paper into the following sections, (1) Introduction (2) Related work, (3) Formal definitions, (4) Computing Phrasal Cohesion, (5) Experimental Set-up, (6) Experiments and Results, (7) Discussion and (8) Conclusion.

2 Related Work

The popular approaches for doing Machine Translation employ the technique of phrase based translation (Koehn et al., 2003; Och and Ney, 2003). In this approach, the translation of possible phrases in the source language (not required to be linguistically accurate) are learnt from a word aligned parallel corpus. Word-aligned parallel corpus is obtained from a sentence aligned corpus using the IBM models (Brown et al., 1993; Och, 2000). Such an approach for doing Machine Translation works best for language pairs which require only local word reordering during translation, but doesn't work very well for language pairs requiring long-distance reordering.

To handle long-distance reordering, syntax based language models have been recently proposed (Fox, 2005; Yamada and Knight, 2002; Chiang, 2007; Quirk et al., 2005). Some of these models use syntactic parsers of either one of the languages or both the languages. Other models do not use any syntactic parser but learn hierarchical rules (that resemble a synchronous grammar) between the language pair which are then used for translation. In models, where the reordering component is distinct within their translation (or decoding) algorithm (Yamada and Knight, 2002), it is assumed that the syntactic structures of the source and target languages are isomorphic to each other. The effectiveness of such models is directly effected by the degree of phrasal cohesion between the language pair, and hence, it is important to study the property of phrasal cohesion.

(Fox, 2002) measures the degree of phrasal cohesion between English and French . One of the observations from her paper was that ordering words by phrasal movement is a reasonable strategy ie., it is reasonable to assume that one can largely carry out an isomorphic transformation of the source syntactic tree in order to obtain the target language word order. Another observation was that the phrasal cohesion of the dependency structures was higher than that of the phrase structure trees. This is important in the context of Indian Languages where dependency structures explain the language structure better than the phrase structure trees.

3 Formal Definitions

Word-aligned data maps the source language words with words in the target language. Given the English sentence $\mathbf{e}$ = e_1 e_2 .... e_n and the Hindi sentence $\mathbf{h}$ = h_1 h_2 .... h_m . The alignment $\mathbf{a}$ contains the mapping of each English word. a_j denotes the position of the Hindi word mapped to the j^{th} source word ie., e_j is aligned to h_{a_j}.

Let 'e_i ... e_j' be an valid English phrase. The span of this English phrase in the source in the target language is represented as $(p_{i,j}, q_{i,j})$. $p_{i,j}$ is the minimum alignment position in the target and $q_{i,j}$ is the maximum alignment position in the target.

$$p_{i,j} = min(a_i...a_j)$$

$$q_{i,j} = max(a_i...a_j)$$

Two phrases e_i.....e_j and e_k.....e_l are said to have crossed each other if their spans ($p_{i,j}$,$q_{i,j}$) and ($p_{k,l}$,$q_{k,l}$) overlap.

4 Computing Phrasal Cohesion

In this paper, we observe two types of phrasal crossings.

1. Head Crossings

   Head Crossings of a node is the overlap of span of the source phrase rooted at the node with the span of its head [1].

   In Figure 1, there is a head crossing for the phrase rooted at the node VP which has “is” as the head. The head crossing take place between phrase “is” and NP. Here span of “is” is (5,5) and span of NP is(3,6).

2. Modifier Crossings

   Modifier Crossings is computed on pairs of nodes which are siblings of each other. It is the overlap of the spans of the two phrases rooted at the two nodes.

There might be cases where the span of one source phrase is equal to the span of another source phrase. For example, consider the following spans, ($p_{i,j}$,$q_{i,j}$) and ($p_{k,l}$,$q_{k,l}$). The phrases e_i.....e_j and e_k.....e_l are phrasal translations if $p_{i,j}$=$p_{k,l}$ and $q_{i,j}$=$q_{k,l}$.

Phrasal translations may not necessarily qualify as instances of crossings. We compute the crossings by both considering phrasal translations as crossings as well as treating them as special cases and not marking them as crossings. We define a filter which is ON when the phrasal translations are not marked as crossings and is OFF when the phrasal translations are considered as crossings.

5 Experimental Set-up

We have used an English-Hindi parallel corpus of 900 word aligned sentence pairs. The English side of parallel corpus has 10702 words with an average of 11.89 words per sentence. The Hindi side of parallel corpus has 11448 words with an average of 12.72 words per sentence. The total number of alignment links between source and target language words are 11865. Our experiments require two types of source sentence analysis : (1) dependency and (2) phrase structure. To obtain phrase structure analysis , we used Collins' parser (Collins, 1997). Collins' parser also contains head information with non-terminals which is needed in our experiments. We used Brill's tagger to obtain part of speech tagged output. To obtain the dependency analysis, we used two different parsers, (1) Bi-directional dependency parser (Shen and Joshi, 2007) and (2) Ryan McDonald's parser (McDonald et al., 2005). We label the alignments links with labels. Strength of 0 is used to indicate primary alignment link and strength of 1 is used to indicate a secondary alignment link (see Figure 2). “They” in English is translated as “unhonne” two times. So, 0 strength is given to the alignment link having first “unhonne” , and 1 strength is given to second link having “unhonne”.

The English-Hindi parallel corpus was taken from TIDES MT project and later refined at IIIT-Hyderabad, India. But, there was some noise in the aligned data which needed to be eliminated. The aligned data was refined in three stages.

- English-Hindi word-aligned corpus was given to the language experts for manual correction.

[1] Note: the span of the head is computed by considering only the head node as a phrase

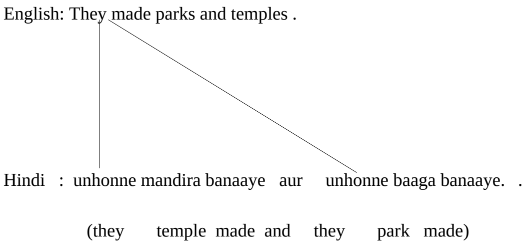
English: They made parks and temples .
Hindi : unhonne mandira banaaye aur unhonne baaga banaaye. .
(they temple made and they park made)

Figure 2: An example of repeated translation

- We identified some of the common errors made by the annotators during the creation of word-aligned data. A tool was developed that marks suspicious cases and helps human annotators to make appropriate corrections. The main criteria used for marking the suspicious cases were the part-of-speech tags associated with the words in both languages. For example, if a source word which is tagged as noun is aligned to a postposition in target sentence, the tool reports an error.

- We treat source language verb groups as phrasal units which engage in many-to-many alignment with the corresponding verb groups in the target alignment (see Figure 3). Any missing links are pointed out by alignment correction tool, which are subsequently corrected by the annotators.

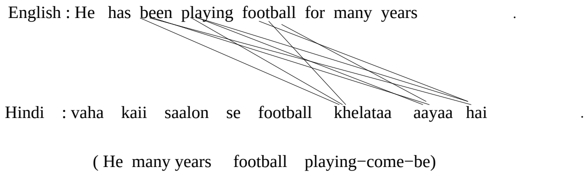
English : He has been playing football for many years .
Hindi : vaha kaii saalon se football khelataa aayaa hai .
(He many years football playing−come−be)

Figure 3: An example of verb group alignment

We use a simple rule-based local word grouper to identify the verb groups at both the sides of the corpus.

6 Experiments and Results

The degree of crossing is computed using two measures.

- Crossings per sentence
- Percentage crossings

Section 6.1 presents *Crossings per sentence* which talks about the coverage of crossings. Section 6.2 presents *percentage crossings* which talks about likelihood of crossings. We have compared our phrase-structure crossing results with Heidi J. Fox's results and have discussed them in Section 6.3.

6.1 Crossings per sentence

Crossing per sentence (σ) is defined as,

$$\sigma = \frac{c}{n} \tag{1}$$

where c is the total number of crossings and n is the total number of sentences.

Crossings	Filter OFF	Filter ON
Head crossings	3.221	2.798
Modifier crossings	0.594	0.456
Phrasal translations	–	0.561

Table 1: Crossings per Sentence in phrase structure trees

In Table 1 we discuss crossings per sentence using phrase structure trees and in Table 2, we discuss the crossing per sentence using dependency trees. We observe that both head crossings and modifier crossings are less when dependency analysis are considered in source. When filter is ON, there is some decrement in crossings per sentence. A test for modifier crossing for a node in phrase structure tree can be performed only if its parent has at least three children. A node with three children would have one head node and two modifiers nodes. A test of head crossing can be performed if the parent has two or more children. Hence, there are more head crossing tests when compared with modifier crossing tests. This leads to a smaller value of average modifier crossings when compared with the value of average head crossings. The phrasal translation filter causes only a minor decrease in the value of "crossings per sentence". As seen in Table 1, the average phrasal translations are 0.561 i.e., there is roughly one translation in every two sentences. When filter is ON, some of the crossings get eliminated and thus, there is a decrease in average head crossings by 0.423 from 3.221 to 2.798. Similarly, there is decrease in modifier crossings per sentence by 0.138 from 0.594 to 0.456.

Table 2 shows the results of average crossings per sentence in which we have used Bi-directional dependency parser.

Crossings	Filter OFF	Filter ON
Head crossings	1.494	0.583
Modifier crossings	0.188	0.176
Phrasal translations	–	1.005

Table 2: Crossings per Sentence in dependency trees

The average crossings of dependency trees is nearly half as compared to the phrase structure trees when filter is OFF, the phrasal translations value is also nearly twice as compared to the phrase structure trees. Filter does not have an effect for the dependency structures in case of modifier crossings but it has some effect on the values of head crossings.

Table 3 shows the results of average crossings in which we have used Ryan McDonald's parser (McDonald et al., 2005)

Crossings	Filter OFF	Filter ON
Head crossings	1.582	0.976
Modifier crossings	0.266	0.257
Phrasal translations	–	0.655

Table 3: Crossings per Sentence in dependency trees

6.2 Percentage crossings

Percentage crossing (γ) is defined as,

$$\gamma = \frac{c}{d} * 100 \tag{2}$$

where c is the total number of crossings and d is the total number of crossings tests performed.

In the previous results we have shown crossings per sentence. which depends heavily on average sentence length. So, percentage crossing is an useful measure to measure the degree of phrasal cohesion which captures the likelihood of a head-modifier pair or modifier-modifier pair to engage in a crossings. As seen in Table 4, when filter is turned ON, head crossings decreases by 3.24% from 37.59% to 34.35% in phrase structure trees. Similarly, there is a decrement in modifier crossings. It decrease by 4.42% from 23.58% to 19.16%, when filter is turned ON. The value of phrasal translations is 5.06%.

Crossings	Filter OFF	Filter ON
Head crossings	37.59%	34.35%
Modifier crossings	23.58%	19.16%
Phrasal translations	–	5.06%

Table 4: Percentage crossings in phrase structure trees

Table 5 shows the percentage crossings results in which we have used Bi-directional dependency parser (Shen and Joshi, 2007)

Crossings	Filter OFF	Filter ON
Head crossings	18.77%	8.27%
Modifier crossings	8.29%	8.09%
Phrasal translations	–	9.82%

Table 5: Percentage Crossings in dependency trees

Table 6 shows the results of percentage crossings in which we have used Ryan McDonald's parser (McDonald et al., 2005)

Crossings	Filter OFF	Filter ON
Head crossings	19.77%	13.10%
Modifier crossings	11.71%	11.63%
Phrasal translations	–	6.44%

Table 6: Percentage Crossings in dependency trees

Similar to the results on average crossings per sentence, the dependency analysis express a lower likelihood of crossings (both head and modifier crossings). We have used two dependency parsers and compared the results. Bi-directional dependency parser has better percentage phrasal translation as compared to Ryan McDonald's Parser.

6.3 Comparison with English-French Results

Heidi J. Fox (Fox, 2002) computes the same measures (average crossings and percentage crossings)

Crossings	Filter OFF	Filter ON
Head crossings	4.790	2.772
Modifier crossings	0.880	0.516
Phrasal translations	–	2.382

Table 7: Crossings per sentence in English-French pair

Crossings	Filter OFF	Filter ON
Head crossings	32.16%	18.61%
Modifier crossings	14.45%	8.47%
Phrasal translations	–	11.35%

Table 8: Percentage crossings in English-French pair

for both phrase structure and dependency parsers. We have compared our phrase structure results with Heidi J. Fox's phrase structure results. Heidi J. Fox's results are shown in Table 7 and Table 8. She had considered three types of alignment links, (1) Sure alignments, (2) Possible alignments and (3) Union of both alignments. As we consider only sure alignments, we compare our results with the Heidi J. Fox's results on the union of sure and possible alignments. The unaligned words are ignored since they do not cover any span on the target side.

Here are the major observations while comparing the results for English-Hindi with the results for English-French.

- Head Crossings : The average head crossing per sentence for English-French (4.790) is more when compared to the same in English-Hindi (3.221) when the phrasal translation filter is OFF. The main reason is that the average length of sentences in the English-Hindi word-aligned corpus (11.9) is much less than the average length of sentences in the English-French corpus. However, the percentage crossings of English-Hindi (37.59%) is higher than the percentage crossings of English-French (32.16%) which is expected as the English and French are closely related languages.

- Modifier Crossing: The observations are same as the observations for the head crossings.

- Phrasal Translations: The percentage of alignment links in English-French which are part of phrasal translations are 11.35% and the percentage for English-Hindi are 5.06%. Hence, when the phrasal translation filter is turned on, the drop in degree of head crossings for English-French is expected to be higher. When filter is ON, the percentage head crossings for English-French is 18.61% (down from 32.16% when filter was OFF) and the percentage head crossings for English-Hindi is 34.35% (minor drop from 37.59% when filter was OFF).

The important observation is that the phrasal cohesion of English-Hindi is much less when compared to the phrasal cohesion of English-French when the phrasal translations are not considered as crossings.

7 Discussion

In this section, we discuss major causes of crossings between English and Hindi. In Table 9, we present the head and modifier pairs which display high values of average crossings per sentence. For the head, part-of-speech tag of head is considered and for modifying sub-tree, head of the root is considered for error analysis.

Head	Modifier	Head Position	Ave. Cross.
MD	VB	left	0.168
IN	NN	left	0.15
NN	IN	left	0.126
TO	VB	left	0.111
VBZ	NN	left	0.067
IN	NNS	left	0.062
VB	IN	left	0.056
VBN	IN	left	0.056

Table 9: Causes of average head crossings in phrase structure trees

In Table 9, we have taken the order into account (represented by head position). (Modal Verb, Verb) pair participate in the maximum number of crossings per sentence (0.168 average head crossings). This is followed by the pairs (Preposition, Noun) and (Noun, Preposition). The average crossings per sentence for the (Preposition, Noun) pair is 0.18 where the preposition is the head. For the cases, where the noun is the head, the average crossings per sentence is 0.126. The fourth highest average crossing per sentence is displayed by to-infintive pair.

Head	Ave. Cross.	Percent Cross.
VBZ	0.22	48.76%
IN	0.197	43.30%
VBP	0.194	43.96%
VB	0.175	48.76%
MD	0.116	59.28%
NN	0.0977	13.39%
NNS	0.066	75%

Table 10: Analysis of head for head crossings in phrase structure trees

In Table 10, we analyze the heads which exhibit a high degree of crossings. The cumulative percentage is computed by considering various modifiers. The heads are represented using their part-of-speech tags.

As seen in Table 10 Verb (VBZ) has maximum head crossings. The average head crossings of auxiliary verbs is 0.22 and percentage crossings is 48.76%. The second highest average crossing is shown by preposition (0.197 average head crossings and 43.30% percentage head crossings). In the Table 11, we analyze only those cases which have high number of occurrences with high percentage crossings.

Head	Modifier	Head Position	Count	Percent Cross.
MD	VB	left	152	96.20%
TO	VB	left	100	88.49%
VBZ	VBN	left	45	84.90%
VBP	VBN	left	38	82.60%
VBD	VBN	left	26	81.25%
VB	VBN	left	19	86.36%

Table 11: Causes of percentage head crossings in phrase structure trees

Table 12 shows the top modifier pairs according to the average modifier crossings in phrase structure trees.

Table 13, shows the top modifier pairs of phrase structure trees, according to the percentage modifier crossings.

1st Modifier	2nd Modifier	Ave. cross
RB	VB	0.042
NN	IN	0.036
IN	IN	0.023
RB	VBN	0.117
IN	NN	0.014
RB	NN	0.012
NNS	IN	0.012
TO	IN	0.010

Table 12: Causes of average modifier crossings in phrase structure trees

1st Modifier	2nd Modifier	Count	Percent. Cross.
RB	VB	38	80.85%
RB	VBN	15	75%
VBZ	NN	5	100%
VBP	RP	3	100%
IN	JJ	3	100%

Table 13: Causes of percentage modifier crossings in phrase structure trees

The percentage crossings of verb-noun (VBZ-NN) pair, verb-particle (VBP-RP) pair and preposition-adjective pair is 100%, but their total number of occurrences are very less.

Table 14 shows the average head crossing of top modifier pairs.

Head	Modifier	Head Position	Ave. Cross.
IN	NN	left	0.0411
IN	NN	right	0.0255
VBZ	IN	left	0.0166
VB	RB	left	0.0166
VBN	IN	left	0.0155
IN	NNS	left	0.0155

Table 14: Causes of average head crossings in dependency trees

(Preposition, Noun) pair and (Noun, Preposition) pair are mainly responsible for the major average head crossings in dependency trees. The third highest average head crossing in dependency trees is shown by (Verb, Preposition) pair. It has 0.0166 average head crossings.

In Table 15, we analyze the heads which exhibit a high degree of crossings in the dependency trees. The cumulative percentage is computed by considering various modifier siblings.

Head	Ave. Cross.	Percent. Cross.
IN	0.0977	13.15%
VB	0.0966	10.38%
VBZ	0.0677	7.58%
NN	0.0644	5.03%
VBP	0.0522	9.69%
VBN	0.0511	9.34%
NNS	0.0388	5.79%

Table 15: Analysis of head crossings in dependency trees

Table 16 shows to the top modifier pairs of dependency trees, according to the average modifier crossings.

1st Modifier	2nd Modifier	Ave. Cross.
NN	NN	0.0077
VRB	RB	0.0066
NN	JJ	0.0055
MD	PRP	0.0055
RB	NNS	0.0044
RB	NN	0.0044

Table 16: Causes of average modifier crossings in dependency trees

In dependency trees, percentage modifier crossings are very less (usually below than 50%). So we have not analyzed percentage crossings for the dependency trees.

Figure 4 shows head crossing due to Modal verb -Verb pair (MD-VB), in which there is a head crossing between the modal verb “could” and VP containing “do” as head. The span of VP is (2,7) and span of “could” is (4,6). Since “could” is head of SQ and it has crossing with it’s sibling (VP chunk) which has a span (2,7). So this is a head crossing.

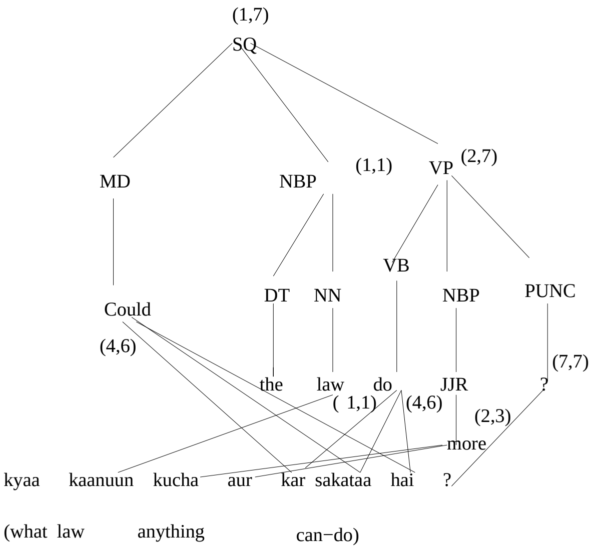

Figure 4: Head crossing of modal verb, verb pair

In Figure 5 there is phrasal translation between “to” (TO) and VP containing “play” as head, as span of ”to” is (6,6) and span of VP is (6,6). We treat it as phrasal translation when filter is ON, and head crossing when filter is OFF, as ”to” is the head. There is another head crossing occurring between “has” (VBZ) and NPB, which has “role” as the head. The span of “has” is (7,7) and the span of NBP is (3,8). In this case “has” is the head.

8 Conclusion

In this paper, we have studied the property of phrasal cohesion for English-Hindi language pair. The phrasal cohesion is computed both using the phrase-structure trees and the dependency trees. For obtaining the phrase-structure trees, we have used Collins’ parser while for the experiments on dependency trees, we have used Shen’s parser and McDonald’s parser.

We compare the degree of phrasal cohesion of English-Hindi and English-French language pairs. We have empirically verified the following,

1. The degree of phrasal cohesion of English-French language pair is more than the phrasal cohesion of the English-Hindi language pair.

2. The degree of phrasal cohesion for dependency structures is greater than the degree of phrasal

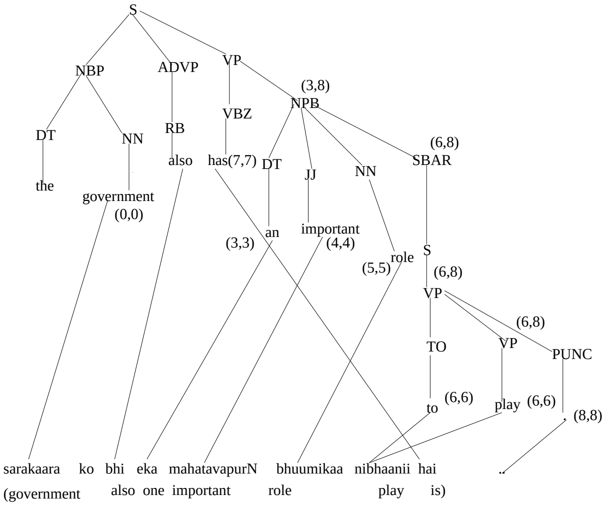

Figure 5: Phrasal Translation

cohesion for phrase-structures. This observation is important, specially in the context of Indian Languages, as the structure of Indian Languages is better expressed using the dependency formalisms.

References

P. Brown, SAD Pietra, VDJ Pietra, and RL Mercer. 1993. The Mathematics of Machine Translation. *Computational Linguistics*, 19(2):263–312.

D. Chiang. 2007. Hierarchical Phrase-Based Translation. *Computational Linguistics*, 33(2):201–228.

M. Collins. 1997. Three generative, lexicalised models for statistical parsing. *Proceedings of the 35th conference on Association for Computational Linguistics*, pages 16–23.

H.J. Fox. 2002. Phrasal cohesion and statistical machine translation. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 304–3111.

Heidi Fox. 2005. Dependency-based statistical machine translation. In *Proceedings of the ACL Student Research Workshop*, pages 91–96, Ann Arbor, Michigan, June. Association for Computational Linguistics.

P. Koehn, F.J. Och, and D. Marcu. 2003. Statistical phrase-based translation. *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1*, pages 48–54.

R. McDonald, F. Pereira, K. Ribarov, and J. Hajic. 2005. Non-projective dependency parsing using spanning tree algorithms. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 523–530.

F.J. Och and H. Ney. 2003. A Systematic Comparison of Various Statistical Alignment Models. *Computational Linguistics*, 29(1):19–51.

F.J. Och. 2000. Giza++: Training of statistical translation models.

C. Quirk, A. Menezes, and C. Cherry. 2005. Dependency treelet translation: syntactically informed phrasal SMT. *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 271–279.

L. Shen and AK Joshi. 2007. Bidirectional LTAG Dependency Parsing. Technical report, Technical Report 07-02, IRCS, UPenn.

Kenji Yamada and Kevin Knight. 2002. A decoder for syntax-based statistical mt. In *Proceedings of 40th Annual Meeting of the Association for Computational Linguistics*, pages 303–310, Philadelphia, Pennsylvania, USA, July. Association for Computational Linguistics.