Proceedings of the 16th EAMT Conference, 28-30 May 2012, Trento, Italy

LTG vs. ITG Coverage of Cross-Lingual Verb Frame Alternations

Karteek Addanki, Chi-kiu Lo, Markus Saers and **Dekai Wu**
Human Language Technology Center
Dept. of Computer Science and Engineering
Hong Kong University of Science and Technology
{vskaddanki|jackielo|masaers|dekai}@cs.ust.hk

Abstract

We show in an empirical study that not only did all cross-lingual alternations of verb frames across Chinese–English translations fall within the reordering capacity of Inversion Transduction Grammars, but more surprisingly, about 97% of the alternations were expressible by the far more restrictive Linear Transduction Grammars. Also, about 71% of the cross-lingual verb frame alternations turn out to be monotonic even for diverse language pairs such as Chinese–English. We also observe that a source verb frame alternation pattern translates into a small subset of the possible target verb frame alternation patterns, based on the construction of the source sentence and the frame set definitions. As a part of our evaluation, we also present a novel linear time algorithm to determine whether a particular syntactic alignment falls within the expressiveness of Linear Transduction Grammars. To our knowledge, this is the first study that attempts to analyze the cross-lingual alternation behavior of semantic frames and the extent of their coverage under syntax-based machine translation formalisms.

1 Introduction

In this paper we present a first empirical study on the cross-lingual verb frame alternations by aligning semantic role fillers in parallel sentences. We evaluate how many of these alignments fall within the expressiveness of two well known syntax based machine translation formalisms: Inversion Transduction Grammars (Wu, 1997) and Linear Transduction Grammars (Saers, 2011). As a part of our evaluation, we discuss the reordering of semantic roles within a frame and across frames within a sentence. We also present a novel algorithm to determine whether there exists a canonical parse for an alignment under Linear Transduction Grammars.

© 2012 European Association for Machine Translation.

While recent years have seen continued improvements in the accuracy of SMT using tree-structured and syntactic models (Wu, 1997; Wu and Chiang, 2009; Wu, 2010; Wu and Fung, 2009b,a), only a few attempts (Wu and Fung, 2009b) have been made towards using semantic roles to guide SMT. Recent studies (Wu and Fung, 2009a) show that most of the glaring errors made by statistical machine translation systems are a result of confused semantic roles which result in serious misunderstanding of the essential meaning. Semantic roles have also been successfully used in evaluating translation utility (Giménez and Màrquez, 2007, 2008; Callison-Burch *et al.*, 2007, 2008; Lo and Wu, 2011a,b). However, no effort has been made to identify the reordering of semantic role fillers across languages. Such an analysis is interesting for two reasons: (1) to determine how much reordering we really need in order to preserve meaning while translating, and (2) to determine which existing syntactic SMT models have an inherent bias towards such a reordering. The first reason helps us determine an upper bound on the expressiveness and hence the computational complexity of the syntactic models. The second enables us to choose syntactic SMT models that can be adapted to incorporate semantic knowledge. Such a system should theoretically be able to capture semantically valid syntactic generalizations, thereby improving translation accuracy.

To fulfill the above requirements, we evaluate two well known syntax-based machine translation formalisms: Inversion Transduction Grammars or ITGs (Wu, 1997) and Linear Transduction Grammars or LTGs (Saers, 2011). As discussed in Wu (1997), ITGs allow nearly all possible reorderings (22 out of 24) given up to four semantic role labels within a semantic frame. Further, various forms of empirical confirmation for the effectiveness of ITG expressivity constraints (Zens and Ney, 2003; Zhang and Gildea, 2005, 2004) motivate us to choose it as a likely candidate. Though ITGs are far more constraining than other higher order syntax directed transduction grammars and IBM models, it would be interesting to see how far an even more constrained model is able to handle reorderings of semantic role fillers. For this purpose, we choose LTGs which are bilingual generalizations of linear grammars and highly constrained compared to the ITGs (Saers *et al.*, 2011).

In order to identify reorderings that produce semantically good translations, one should approach the task of aligning semantic role fillers carefully. Such an alignment should accurately match at least the corresponding basic event structure "Who did what to whom, when, where and why" in both source and target languages in order to preserve meaning (Pradhan *et al.*, 2004). Further, a complete analysis of the syntactic alignments generated as a result of aligning semantic role fillers entails examining the reordering of roles both within a frame and across all the frames in one sentence. The possibility that an exact alignment might not exist at all even for semantically valid translations should also be considered.

In this paper, we present an empirical study of the semantic reorderings as a result of aligning cross-lingual semantic role fillers. We also determine to what extent the alignment constraints of ITGs and LTGs permit such reorderings. We use semantically annotated Chinese–English parallel resources and manually align the semantic role fillers. In order to identify the alignments permitted by LTGs we propose a novel linear time algorithm. Our results indicate that ITGs permit all the syntactic reordering occurring from aligning semantic role fillers. Interestingly, we also show that about 97% of the alignments are handled by LTGs. We also observe that all the verb frame alternations of semantic frames fall within the reordering capability of both LTGs and ITGs.

The rest of the paper is organized as follows. In the next section, we state and prove an algorithm to determine whether an alignment corresponding to a given permutation can be parsed by a bracketing linear transduction grammar. In section 3, we describe our experimental setup. Results and the conclusion follow in sections 4 and 5.

2 LTG parsability algorithm

In this section, we present a linear time algorithm to determine whether or not there exists a canonical parse for an alignment under LTGs. Although, LTGs are restricted forms of ITGs, the algorithm for determining whether a permutation corresponding to an alignment can be parsed by a LTG is not a special case of the linear time skeleton algorithm for binarization of synchronous grammars (Huang *et al.*, 2009). The linear time skeleton algorithm builds canonical binarization trees by reducing greedily but such an approach would not work for a LTG. For example, the permutation $[3, 2, 0, 1]$ which can be parsed by an LTG reduces to 2-3, 0-1 on the stack which cannot be further reduced.

We propose an algorithm that makes use of a technique similar to top-down parsing of bisentences using linear transduction grammars. The algorithm is as shown in the procedure parsable. In order to prove the correctness of the algorithm, we use the definition of *permuted sequence* from Huang *et al.* (2009) but we redefine *proper split* in the context of BLTGs. The proof is as follows:

Definition 1. *A* ***permuted sequence*** *is a permutation of consecutive integers. If a permuted sequence of sequence* $\mathbf{a}$ *can be split into the concatenation of a permuted sequence* $\mathbf{b}$ *and a single element of permutation* α *such that* $\mathbf{a} = (\mathbf{b}; \alpha)$ *or* $\mathbf{a} = (\alpha; \mathbf{b})$*, then the corresponding split is called the* ***proper split*** *of* $\mathbf{a}$.

The definition of a proper split implicitly imposes the constraints of a linear transduction grammar. Restricting one of the elements in a split to be a single element in the permutation is equivalent to allowing at most one nonterminal in the right hand side of productions. The definition of a permuted sequence enforces the projection constraints of transduction grammars by allowing no gaps in the reorderings within a constituent.

Lemma 1. *A split* $\mathbf{a} = (\mathbf{b}; \alpha)$ *or* $\mathbf{a} = (\alpha; \mathbf{b})$ *is proper if and only if* $\alpha = \max(\mathbf{a})$ *or* $\alpha = \min(\mathbf{a})$.

Proof. We prove both the forward and reverse implications as follows:

1. If $(\mathbf{b}; \alpha)$ is a proper split of $\mathbf{a}$, then $\mathbf{b}$ is a permuted sequence. From Definition 1, all the elements in $\mathbf{b}$ should be consecutive.Hence α is either greater than all the elements in $\mathbf{b}$ or less than all the elements in $\mathbf{b}$ which implies $\alpha = \max(\mathbf{a})$ or $\alpha = \min(\mathbf{a})$ respectively. Similar conclusions can be made for the case when $\mathbf{a} = (\alpha; \mathbf{b})$.

2. If $\alpha = \max(\mathbf{a})$ and there exists a split of $\mathbf{a}$ such that $\mathbf{a} = (\mathbf{b}; \alpha)$, then $\mathbf{b}$ is a permuted sequence from $[min, \ldots, max - 1]$. This makes $(\mathbf{b}; \alpha)$ a proper split. Similarly, when $\alpha = \min(\mathbf{a})$, $\mathbf{b}$ is a permuted sequence from $[min + 1, \ldots, max]$ and $(\mathbf{b}; \alpha)$ is a proper split. The case when $\mathbf{a} = (\alpha; \mathbf{b})$ is similar.

□

Lemma 2. *If* $\mathbf{a}$ *is a permuted sequence covering* $[min, \ldots, max]$*, and there exists a proper split of* $\mathbf{a}$ *such that* $\mathbf{a} = (\mathbf{b}; \alpha)$ *or* $\mathbf{a} = (\alpha; \mathbf{b})$*, then* $\mathbf{b}$ *is a permuted sequence covering* $[min, \ldots, max - 1]$ *or* $[min + 1, \ldots, max]$.

Proof. From Lemma 1, $\alpha = \max(\mathbf{a})$ or $\alpha = \min(\mathbf{a})$. Therefore, $\mathbf{b}$ covers the range $[min, \ldots, max - 1]$ or $[min + 1, \ldots, max]$ according to whether α is $\max(\mathbf{a})$ or $\min(\mathbf{a})$ respectively. □

Procedure parsable($\mathbf{a}$,min,max)

```
input  : A permuted sequence a of range [min, ..., max]
output: true or false depending on the whether or not a is BLTG parsable
begin
    if max − min + 1 = 1 then                                   // base case
        return true
    else
        if first(a) = min then                                  // a = [α : b]
            shift(a)               // remove the first element of a
            return parsable(a,min + 1,max)
        else if first(a) = max then                             // a = ⟨α : b⟩
            shift(a)
            return parsable(a,min,max − 1)
        else if last(a) = min then                              // a = ⟨b : α⟩
            pop(a)                 // remove the last element of a
            return parsable(a,min + 1,max)
        else if last(a) = max then                              // a = [b : α]
            pop(a)
            return parsable(a,min,max − 1)
        else
            // no proper split exists
            return false
```

Definition 2. *A permuted sequence* $\mathbf{a}$ *is said to be parsable if:*

1. $\mathbf{a}$ *is a permuted sequence of size 1 i.e.,* $\mathbf{a} = (\alpha)$
2. *there exists a proper split of* $\mathbf{a}$ *containing a permuted sequence* $\mathbf{b}$*, which is also parsable.*

This is a recursive definition and associates a hierarchical tree structure with each permutable sequence. The tree structure is equivalent to the biparse tree which parses the alignment represented by the permutable sequence. For the sake of completeness, we define the parse tree below:

Definition 3. *A parse tree* $t(\mathbf{a})$ *of a parsable sequence* $\mathbf{a}$ *is either:*

1. α *if* $\mathbf{a} = (\alpha)$*, or*
2. $[\alpha\ t(\mathbf{b})]$ *if* $\mathbf{a} = (\alpha; \mathbf{b})$ *and* $\alpha = \min(\mathbf{a})$*, or*
3. $\langle t(\mathbf{b})\ \alpha\rangle$ *if* $\mathbf{a} = (\alpha; \mathbf{b})$ *and* $\alpha = \max(\mathbf{a})$*, or*
4. $[t(\mathbf{b})\ \alpha]$ *if* $\mathbf{a} = (\mathbf{b}; \alpha)$ *and* $\alpha = \max(\mathbf{a})$*, or*
5. $\langle\alpha\ t(\mathbf{b})\rangle$ *if* $\mathbf{a} = (\mathbf{b}; \alpha)$ *and* $\alpha = \min(\mathbf{a})$ *where* $t(\mathbf{b})$ *is the parse tree of* $\mathbf{b}$.

We use the same notation as Wu (1997) for representing the straight and inverted configurations. We also note that there might exist more than one parse tree for a parsable sequence but we are interested only in whether or not there exists at least one parse tree.

Theorem 1. *Procedure parsable runs in time linear to the length of the input and succeeds (i.e., returns* ***true****) if and only if the input permuted sequence* $\mathbf{a}$ *is parsable.*

Proof.

1. If the procedure returns ***true***, then $\mathbf{a}$ is binarizable as we can recover a parse tree from the algorithm.

2. If $\mathbf{a}$ is parsable, then the procedure must return ***true***.

   We prove this by a complete induction on n, the length of $\mathbf{a}$.

   Base case: $n = 1$, trivial. Assume that the condition holds for all $n' < n$.

   From Definition 2, if a permuted sequence is parsable then there exists a proper split. We check for all possible values of α in a proper split (see Lemma 1 and Definition 1). By induction hypothesis, the procedure succeeds as the procedure is called on a permuted sequence of length $n - 1$ after the first split.

   As the procedure is recursively called a maximum of n times where n is the length of $\mathbf{a}$ and each procedure call takes $O(1)$ time, the algorithm is linear with respect to the length of the input. The total complexity is $O(n)$.

□

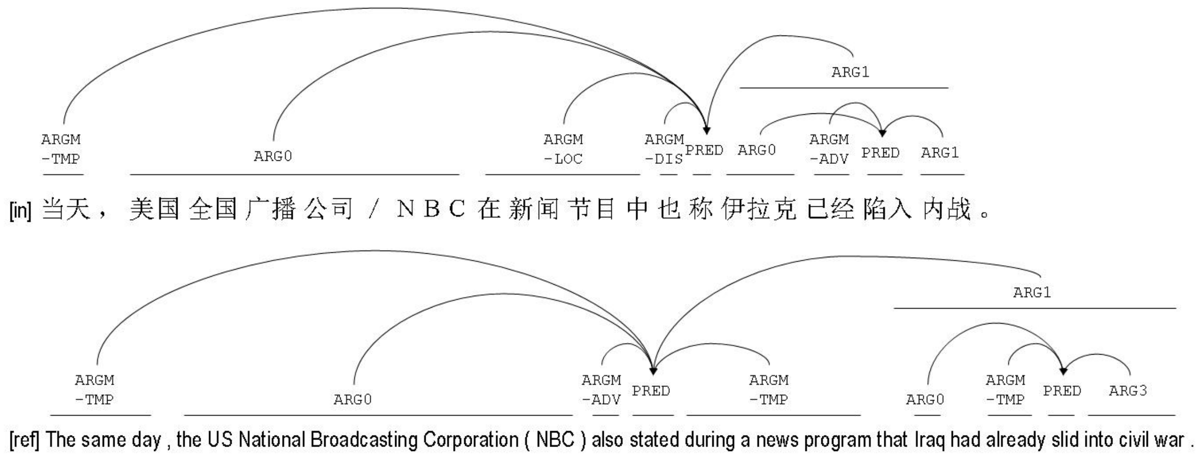

Figure 1: An example of nested semantic role fillers

3 Experimental setup

3.1 Semantic role alignment

As a first step in our experiment, we would like to identify semantic role fillers in the target language sentence that match the basic event structures of "Who did what to whom, when, where and why" in both source and target sentences. We use a randomly sampled subset of 100 sentence pairs from the Chinese–English parallel corpus derived from Phase 2.5 of the DARPA GALE program. The Chinese and English sentences are annotated with gold-standard semantic roles in Propbank Style and belong to the news wire genre. We use a bilingual speaker to manually align the semantic roles. We do not attempt to automatically align semantic role fillers with identical semantic role labels in a frame as the gold standard annotation was done monolingually leading to a possibility of mismatch between the source and target role fillers.

The aligner was instructed to align role fillers that are precise translations of each other. First the predicates corresponding to different frames in source and target sentences are aligned. For each aligned predicate, the role fillers for the frame modifiers are aligned. We assume that if there is not an exact match between predicates in source and target sentences, none of the role fillers for the other modifiers can be aligned. Such a scenario would occur only when the source sentence is paraphrased using a totally different construction in the target language and we ignore such sentence pairs. We only found one such example in our sample and it is shown below:

Source: 报道说，今年以来英国有关征收"绿色税"的争论和猜测不断。

Gloss: Report said , this year throughout Britain related levying "green tax" 's controversy and speculation did not stop .

Target: According to the report , "green taxes " have come under constant controversy and speculation in Britain this year .

One can notice that both source and target sentences convey the same semantic information but have no predicates that can be aligned. While the source sentence has two predicates 征收 (levy) and 不断 (did not stop), none of them match with the predicate in the target sentence which is "come". We assume that such constructions are rare enough in parallel corpora and treat them as an exception rather than a rule.

We also do not align partially matching semantic role fillers i.e., semantic role fillers that do not contain the same level of information. We have observed that such a mismatch primarily occurs due to the independent annotation of the source and target corpora.

3.2 Extracting semantic reorderings

We extract the semantic reordering information from the manually aligned semantic role fillers both within a single frame and across all the frames in a sentence. While extracting the reorderings we ignore the tokens that: (1) correspond to an unaligned semantic role filler and, (2) are not annotated with any semantic role. Tokens that are not annotated with any semantic role perform the task of providing the syntactic structure to the meaning contained in the semantic role fillers in a sentence. As our goal is primarily to identify the kind of reorderings necessary to preserve meaning we do not deal with these tokens.

Semantic role fillers may contain nested semantic frames. So a bispan corresponding to a semantic role filler alignment might contain bispans corresponding to other semantic roles. This leads to a hierarchical or a compositional syntactic alignment between source and target sentences. However, preserving the compositionality of the syn-

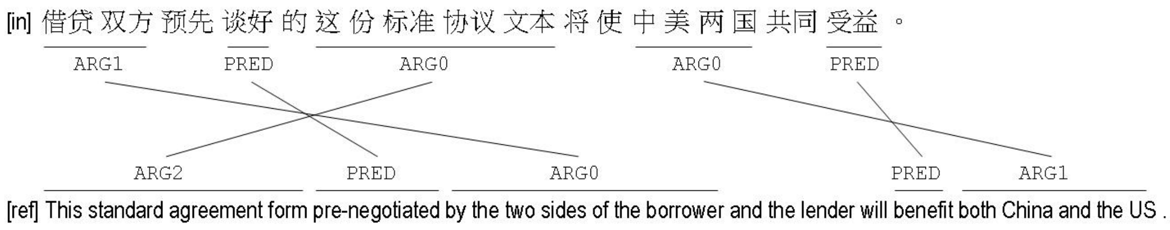

Figure 2: An example of a semantic alignment not parsable by LTG

tactic alignments adds little value in understanding the syntactic reordering necessary to preserve meaning. For example, if a semantic role filler contains a nested semantic frame (or frames) then all the semantic information contained in the encompassing role filler is captured by the role fillers in the nested frame (or frames). An example is shown in figure 1.

So we extract an alignment permutation by identifying the mapping between disjoint semantic role fillers among all the frames in a sentence. We did not encounter any non-disjoint alignments between the role fillers, and we could successfully extract a permutation from all the sentence pairs. In the next stage of the experiment we evaluate whether a given permutation falls within the reordering capacity of ITGs and LTGs.

3.3 Evaluating alignments

We evaluate the alignments of the semantic role fillers both within a frame and across all the frames in a sentence. The reordering within a frame indicates the relation between the cross-lingual verb frame alternation patterns. We also extract the alignments of disjoint semantic role fillers across all the frames in the sentence pairs as discussed in the previous subsection. In both the cases, we determine whether or not there exists a canonical parse for the alignment using an ITG or an LTG. For this purpose we use the shift reduce algorithm proposed in (Huang *et al.*, 2009) for ITGs and the algorithm proposed in Section 2 for the LTGs.

4 Results

We observed that all the cross-lingual alternations of verb frames fall within the reordering capability of both LTGs and ITGs. We did not find any semantic frames which had more than four arguments (including the predicate) in our sample. Both LTGs and ITGs are capable of generating all possible alternations for semantic frames up to three arguments. In the case where there were four arguments we did not encounter any examples where the role fillers formed an *inside-out* ($[2, 0, 3, 1]$ or $[1, 3, 0, 2]$) which both LTGs and ITGs cannot generate, nor did we find *constituent swapping* ($[2, 3, 0, 1]$) or *serial inversion* ($[1, 0, 3, 2]$) which LTGs cannot generate. Note that this result corresponds to the alignment of semantic role fillers within one frame. The alignment of semantic role fillers across all the semantic frames in a sentence is discussed in the next subsection.

4.1 LTGs have high semantic alignment coverage

We observed that the alignment of the semantic role fillers across all the frames in a sentence fall within the reordering capacity of ITGs for all the sentence pairs in our sample. We did not find any translations of semantic frames wherein the role fillers formed an *inside-out* alignment. This observation is consistent with the universal language hypothesis of ITGs (Wu, 1997).

Surprisingly, for about 97% of the sentences, the generated alignments could be expressed by the far more restrictive Linear Transduction Grammars. There were only three sentence pairs that had reordering of verb frame alternations which could not be parsed by an LTG. All the alignments that could not be parsed by an LTG contained the *serial inversion* permutation pattern which can be parsed by an ITG but not by an LTG. Figure 2 shows an example.

In Figure 2, the order of the arguments in two adjacent semantic frames is inverted. Although, the alternation of each semantic frame can be independently parsed by an LTG, the reordering caused by these alternations at the sentence level cannot be parsed. All the alignments in our sample, that could not be parsed an LTG exhibited similar pattern.

Further, we noticed that for about 71% of the sentences, the alignments were monotonic. It is interesting to note that despite reordering at a surface level, the parts that carried the semantic information remained in the same order in both source and target sentences. A possible reason for such a high percentage of monotonic alignments could be the similarity in the word orders of Chinese and English as both languages follow a *subject-verb-object* construction. Further empirical testing is needed to determine whether or not this observation holds true for language pairs with a difference in word order.

Though these sentence pairs are an example of the limitation of LTGs to express reordering that occurs in natural languages, it is interesting to note that the *serial inversion* alignment pattern was infrequent in our sample and there is no occurrence of the *constituent swapping* alignment pattern.

4.2 Topicalization versus fluency

We noticed that reordering of semantic role fillers depends on factors deeper than the syntactic structure of the source sentence. It depends on whether the translation aims to capture the intentional or the extentional semantics of the source sentence. This results in some interesting trade-offs that could be made between topicalization and fluency. Consider the sentence pair in Figure 2. The source sentence contains three semantic frames with predicates 谈好 (negotiated), 使 (cause) and 受益 (benefited) whereas the translation contains only two semantic frames corresponding to the predicates negotiate and benefit.

Source: 借贷双方预先谈好的这份标准协议文本将使中美两国共同受益。

Gloss: Borrower and lender both sides pre-negotiated *DE*/的 this standard agreement form will cause China US two countries together receive benefit .

Translation: This standard agreement form pre-negotiated by the two sides of the borrower and the lender will benefit both China and the US .

Alternative translation: This standard agreement form pre-negotiated by the two sides of the borrower and the lender will result in both China and the US getting benefited .

In the second frame the VP in the Chinese sentence undergoes a dative alternation upon translation as the double-object construction (NP-NP) for the verb *benefit*, is less fluent and possibly awkward in English. One could argue that the proposed translation is not semantically equivalent to the source sentence because "a difference in the syntactic form always spells a difference in meaning" (Goldberg, 1995). We provide an alternative translation which preserves the topicalization on the possession facet and the possessor rather than on the transfer. While the permutation generated by aligning the source and target sentences cannot be parsed by an LTG ($[2, 1, 0, 4, 3]$), the alternative translation generates a permutation that can be parsed ($[2, 1, 0, 3, 4, 5]$).

While both translations manage to convey the meaning in the source sentence correctly, one focuses on fluency and the other on preserving the topicalization. It is not our purpose to compare the translation quality in both cases but to provide an example of the subtle transformation of semantics that occurs while translating and how they affect reordering.

4.3 Verb frame alternation patterns

We studied how alternation patterns change when verb frames are translated from one language to another. If the alternation of the translated verb frame can be estimated based on the source language alternation, it might provide information as to how the target language sentence should be constructed, rather than solely relying on the surface reordering rules. Although Schulte im Walde (2000) showed that verbs can be clustered into semantic categories based on their alternation behavior, little work has been done towards understanding cross-lingual verb frame alternation patterns. As a first step towards understanding the cross-lingual alternation behavior, we collected some statistics and performed a rudimentary qualitative analysis on the alternation patterns of the target language given a source language alternation pattern.

We observed that for about 77% of the semantic frames, the alternation pattern is preserved when translated and for about 4%, the target alternation pattern was a permutation of the source pattern. Surprisingly, for about 19% of the frames, the target frame alternation pattern had a different label which was not present in the source frame. For example, the $[arg0 : action]$ pattern in Chinese gets converted into $[arg1 : action]$ in English. Table 1 shows the counts for the target alternation patterns for some of the frequently occurring source alternations.

A given source alternation pattern is aligned only to a small subset of the possible target frame alternations. For Chinese–English, the alternation pattern remains the same in most cases which could be attributed to the similarity in word order. From Table 1, one can observe that the $[arg0 : action : arg1]$, the most frequent source alternation pattern, remains unaltered 88 out of 97 times.

In cases where the target alternation pattern was a permutation of the source pattern, we observed a difference in the voice of the source and target sentences. In most cases, the Chinese sentence in active voice was translated into an English sentence in passive voice. In a few cases, translation demanded a reordering of the source alternation pattern as English had no equivalent construction. For example, in Chinese, when a verb qualifies a noun the verb comes after the noun, while in English it comes before. Hence the phrase 信心增强 (confidence increased) translates into strengthened confidence in English.

For sentence pairs where the source and target alternation patterns differed in labels, we noticed that there were some inconsistencies in the annotation. The sentence pairs were manually annotated with the frame sets defined for Chinese and English

Zh/En alt. patterns	[arg0:action:arg1]	[arg0:action	[action:arg1]	[arg1:action]	Sum
[arg0:action:arg1]	88	0	0	0	88
[arg0:action]	0	11	0	0	11
[action:arg1]	0	3	39	1	43
[arg1:action]	0	12	6	3	21
[action:arg2]	0	1	5	0	6
[arg0:action:arg2]	3	0	0	0	3
[action:arg4]	0	1	1	0	2
[arg1:action:arg2]	3	0	0	0	3
[arg1:action:arg4]	3	0	0	0	3
Sum	97	28	51	4	

Table 1: Frequency of source and target alternation pattern occurrence

in the Propbank (Palmer *et al.*, 2005). We argue that it is due to the limitation of frame set definitions as they were defined to be consistent within one language but not across languages. For example, in the frame set definition of 死于 (died of), the *arg0* is the *entity who dies*, while in the frame set definition of its translation die, *the deceased* is *arg1* and there is no *arg0* defined. Similar observations could be made for most of the sentence pairs which differed in source and target alternation labels.

As our initial analysis of cross-lingual verb frame alternation patterns suggests that patterns in one language align with only a restricted subset of patterns in the other language, we believe that it might be possible to learn the target frame alternation patterns given a source frame alternation pattern. However, it is worth noting that it is important to deal with the inconsistencies in the frame set definitions across languages before one attempts such a task. Larger scale experiments are needed in order to reliably identify the relation between source and target alternation patterns.

5 Conclusion

In this paper, we reported a first empirical study of cross-lingual verb frame alternations and made the following observations: (1) the alignments of the semantic role fillers fall within the reordering capacity of ITGs for all the sentences, (2) even highly constrained models such as LTGs are capable of parsing most of these alignments and (3) there appears to be a correlation between the alternation patterns of the source and target verb frames. We also presented a novel algorithm to determine whether or not a permutation falls within the reordering constraints of LTGs.

The first two observations indicate that alignments of parts that carry the semantic information in sentences (i.e., predicates and semantic role fillers) do not warrant a highly expressive model. Further, since the evaluated models have an inherent bias towards generating these alignments, the constraints they enforce would be useful if one were to automatically align and/or induce semantic role fillers from parallel sentences. It would be interesting to evaluate the performance of the alignments generated by using the semantic role fillers as anchors.

Our observation about the verb frame alternation patterns suggests that it might be possible to predict the target frame alternation pattern given a source frame alternation pattern which would be useful for aligning the verb frames. Although a qualitative evaluation indicated that source sentence construction and frame set definitions can affect target alternation pattern, further evaluation is needed in order to reliably identify features that affect alternation patterns.

As for future work, we think it is interesting to explore methods to incorporate semantic frames in generating robust alignments. It would also be interesting to see whether cross-lingual alternation patterns provide information about verb classes in the bilingual case similar to Schulte im Walde (2000).

Acknowledgments

This material is based upon work supported in part by the Defense Advanced Research Projects Agency (DARPA) under BOLT contract no. HR0011-12-C-0016, and GALE contract nos. HR0011-06-C-0022 and HR0011-06-C-0023; by the European Union under the FP7 grant agreement no. 287658; and by the Hong Kong Research Grants Council (RGC) research grants GRF621008, GRF612806, DAG03/04.EG09, RGC6256/00E, and RGC6083/99E. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the RGC, EU, or DARPA.

References

Chris Callison-Burch, Cameron Fordyce, Philipp Koehn, Christof Monz, and Josh Schroeder. (Meta-) evaluation of Machine Translation. In

Proceedings of the 2nd Workshop on Statistical Machine Translation, pages 136–158, 2007.

Chris Callison-Burch, Cameron Fordyce, Philipp Koehn, Christof Monz, and Josh Schroeder. Further Meta-evaluation of Machine Translation. In *Proceedings of the 3rd Workshop on Statistical Machine Translation*, pages 70–106, 2008.

Jesús Giménez and Lluís Màrquez. Linguistic Features for Automatic Evaluation of Heterogenous MT Systems. In *Proceedings of the 2nd Workshop on Statistical Machine Translation*, pages 256–264, Prague, Czech Republic, June 2007. Association for Computational Linguistics.

Jesús Giménez and Lluís Màrquez. A Smorgasbord of Features for Automatic MT Evaluation. In *Proceedings of the 3rd Workshop on Statistical Machine Translation*, pages 195–198, Columbus, OH, June 2008. Association for Computational Linguistics.

Adele E. Goldberg. *Constructions: A construction grammar approach to argument structure*. University of Chicago Press, 1995.

Liang Huang, Hao Zhang, Daniel Gildea, and Kevin Knight. Binarization of synchronous context-free grammars. *Computational Linguistics*, 35(4):559–595, 2009.

Chi-kiu Lo and Dekai Wu. MEANT: An Inexpensive, High-Accuracy, Semi-Automatic Metric for Evaluating Translation Utility based on Semantic Roles. In *Proceedings of the Joint conference of the 49th Annual Meeting of the Association for Computational Linguistics : Human Language Technologies (ACL-HLT-11)*, 2011.

Chi-kiu Lo and Dekai Wu. SMT vs. AI redux: How semantic frames evaluate MT more accurately. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*, 2011.

Martha Palmer, Daniel Gildea, and Paul Kingsbury. The proposition bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1):71–106, 2005.

Sameer Pradhan, Wayne Ward, Kadri Hacioglu, James H. Martin, and Dan Jurafsky. Shallow Semantic Parsing Using Support Vector Machines. In *Proceedings of the 2004 Conference on Human Language Technology and the North American Chapter of the Association for Computational Linguistics (HLT-NAACL-04)*, 2004.

Markus Saers, Dekai Wu, and Chris Quirk. On the expressivity of linear transductions. In *Proceedings of the 13th Machine Translation Summit (MT Summit XIII)*, 2011.

Markus Saers. *Translation as linear transduction: Models and algorithms for efficient learning in statistical machine translation*. Acta Universitatis Upsaliensis„ 2011.

Sabine Schulte im Walde. Clustering verbs semantically according to their alternation behaviour. In *Proceedings of the 18th conference on Computational linguistics-Volume 2*, pages 747–753. Association for Computational Linguistics, 2000.

Dekai Wu and David Chiang, editors. *Proceedings of the Third Workshop on Syntax and Structure in Statistical Translation (SSST-3) at NAACL HLT 2009*. Association for Computational Linguistics, Boulder, Colorado, June 2009.

Dekai Wu and Pascale Fung. Can Semantic Role Labeling Improve SMT? In *Proceedings of the 13th Annual Conference of the European Association for Machine Translation (EAMT-2009)*, pages 218–225, Barcelona, Spain, May 2009.

Dekai Wu and Pascale Fung. Semantic Roles for SMT: A Hybrid Two-Pass Model. In *Proceedings of the 2009 Conference of the North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT-09)*, pages 13–16, 2009.

Dekai Wu. Stochastic Inversion Transduction Grammars and Bilingual Parsing of Parallel Corpora. *Computational Linguistics*, 23(3):377–403, 1997.

Dekai Wu, editor. *Proceedings of the 4th Workshop on Syntax and Structure in Statistical Translation*. Coling 2010 Organizing Committee, Beijing, China, August 2010.

Richard Zens and Hermann Ney. A comparative study on reordering constraints in statistical machine translation. In *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics-Volume 1*, pages 144–151. Association for Computational Linguistics, 2003.

Hao Zhang and Daniel Gildea. Syntax-based alignment: supervised or unsupervised? In *Proceedings of the 20th international conference on Computational Linguistics*, page 418. Association for Computational Linguistics, 2004.

Hao Zhang and Daniel Gildea. Stochastic lexicalized inversion transduction grammar for alignment. In *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*, pages 475–482. Association for Computational Linguistics, 2005.