How Synchronous are Adjuncts in Translation Data?

Sophie Arnoult
ILLC
University of Amsterdam
`s.i.arnoult@uva.nl`

Khalil Sima'an
ILLC
University of Amsterdam
`k.simaan@uva.nl`

Abstract

The argument-adjunct distinction is central to most syntactic and semantic theories. As optional elements that refine (the meaning of) a phrase, adjuncts are important for recursive, compositional accounts of syntax, semantics and translation. In formal accounts of machine translation, adjuncts are often treated as modifiers applying *synchronously* in source and target derivations. But how well can the assumption of *synchronous adjunction* explain translation equivalence in actual parallel data? In this paper we present the first empirical study of translation equivalence of adjuncts on a variety of French-English parallel corpora, while varying word alignments so we can gauge the effect of errors in them. We show that for proper measurement of the types of translation equivalence of adjuncts, we must work with non-contiguous, many-to-many relations, thereby amending the traditional Direct Correspondence Assumption. Our empirical results show that 70% of manually identified adjuncts have adjunct translation equivalents in training data, against roughly 50% for automatically identified adjuncts.

1 Introduction

Most syntactic and semantic theories agree on the argument-adjunct distinction, although they vary on the specifics of this distinction. Common to these theories is that adjunction is a central device for language recursion, as adjunction modifies initial but complete sentences by adding optional phrases; adjunction also contributes to semantic compositionality, albeit in various ways, as syntactic adjuncts may take different semantic roles. Shieber and Schabes (1990) transfer the role of adjuncts from monolingual syntax (Joshi et al., 1975) to the realm of translation equivalence using a Synchronous Tree Adjoining Grammars (STAG), and propose to view adjunction as a synchronous operation for *recursive*, *compositional* translation. STAG therefore relies substantially on what Hwa (2002) calls the Direct Correspondence Assumption, the notion that semantic or syntactic relations correspond across a bitext. We know from various works–notably by Hwa et al. (2002) for dependency relations, Arnoult and Sima'an (2012) for adjuncts, and Padó and Lapata (2009) and Wu and Fung (2009) for semantic roles–that the Direct Correspondence Assumption does not always hold.

A question that has not received much attention is the degree to which the assumption of synchronous adjunction is supported in human translation data. This is crucial for the successful application of linguistically-motivated STAG, but attempts at answering this question empirically are hampered by a variety of difficulties. Linguistic structures may diverge between languages (Dorr, 1994), translations may be more or less literal, and annotation resources may be inaccurate, when they are available at all. Besides, automatic word alignments are known to be noisy and manual alignments are rather scarce. The work of Arnoult and Sima'an (2012) reports lower and upper bounds of one-to-one adjunct correspondence, using rather limited resources to identify French adjuncts making their results not directly applicable for measuring the stability of the synchronous adjunction assumption.

In this paper we aim at redefining the translation equivalence of adjuncts in ways that allow us to report far more accurate bounds on their cross-linguistic correspondence. In particular, we are interested in measuring adjunct correspondence robustly, in training data.

Consider for example the sentence pair of Fig-

ure 1. Most adjuncts in each sentence translate as adjuncts in the other sentence, but one of these translation equivalences appears to be many-to-many, because of parsing mismatches across the bitext; both parses and adjunct labellers on both sides of the bitext must be on par for adjunct translation equivalences to be established. Besides, one generally establishes translation equivalence using word alignments, which may be noisy. Another factor is that of the degree of translation equivalence in the data in general; while parallel bitexts express the same meaning, meaning may diverge locally.
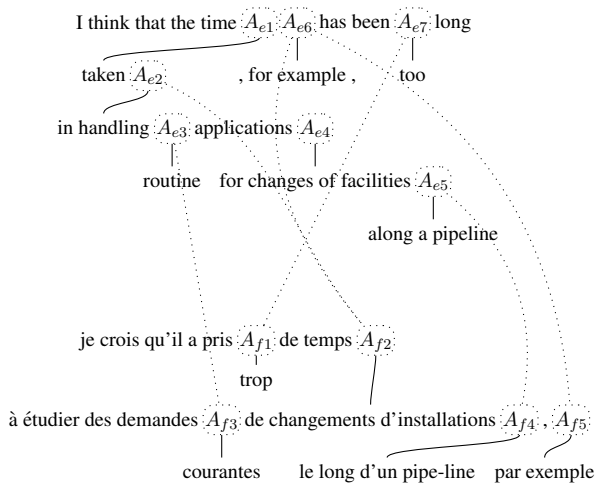

Figure 1: Example sentence pair

This paper contributes the first study to measure the degree of adjunction synchronicity: we derive many-to-many pairings between adjuncts across a bitext, thus supporting a generic view of translation equivalence, where meaning can be expressed by distinct entities and redistributed freely in translation; practically, this also allows us to capture equivalence in spite of mismatched parses. We abstract away from word alignments to a certain degree, as we directly pair adjuncts across a bitext, but we still use word alignments– namely the overlap of adjunct projections with target adjuncts–to decide on these pairings. We further distinguish between adjunct pairings that are bijective through the word alignment, and other pairings, where the translation equivalence does not exactly agree with the word alignment; we qualify these pairings as weakly equivalent.

Under this new view of adjunct translation equivalence, we perform measures in French-English data. We show that adjunction is preserved in 40% to 50% of the cases with automatically labelled adjuncts, with differences between data sets, word aligners and sentence length; about 25% more adjuncts form weakly translation-equivalent pairings. With gold adjunct annotations, the proportion of translation-equivalent adjuncts increases to 70%.

These results show that adjunct labelling accuracy on both sides of the data is crucial for adjunct alignment, while suggesting that applications that exploit adjunction can gain from decreasing their dependence on word alignments and idealized experimental conditions , and identifying favorable contexts for adjunct preservation.

2 Alignment-based role pairing

How can one find translation-equivalent adjuncts using word alignments, without being too constrained by the latter? Obviously, adjunct pairs that are consistent with the word alignments are translation equivalent, but we also want to be able to identify translation-equivalent adjuncts that are not exactly aligned to each other, and also to accept many-to-many pairings; not only to get linguistically justified discontinuous pairs, as with the French double negation particle, but also for robustness with regard to dissimilar attachments in the French and English parses.

2.1 Translation equivalence under the alignment-consistency constraint

Consider for instance Figure 2, which represents a word alignment for part of the sentence pair of Figure 1. We would like to match $\bar{f}_2$ to $\bar{e}_2$ and $\bar{e}_6$, $\bar{f}_3$ to $\bar{e}_3$, $\bar{f}_4$ to $\bar{e}_5$, and $\bar{f}_5$ to $\bar{e}_6$. If one only pairs adjuncts that are consistent with the word alignment, one obtains only half of these adjunct pairs: $\langle \bar{f}_3, \bar{e}_3 \rangle$ and $\langle \bar{f}_4, \bar{e}_5 \rangle$; one cannot pair up $\bar{f}_5$ and $\bar{e}_6$ because the latter is also aligned outside of the former; and one can also not find the equivalence between $\bar{f}_2$ on one hand and $\bar{e}_2$ and $\bar{e}_6$ on the other hand if one assumes one-to-one correspondence between adjuncts.

2.2 Translation equivalence through projection

We align adjuncts across the bitext by projecting them through the word alignment and finding, for each adjunct, the shortest adjunct or sequence of adjuncts that overlaps the most with that adjunct's
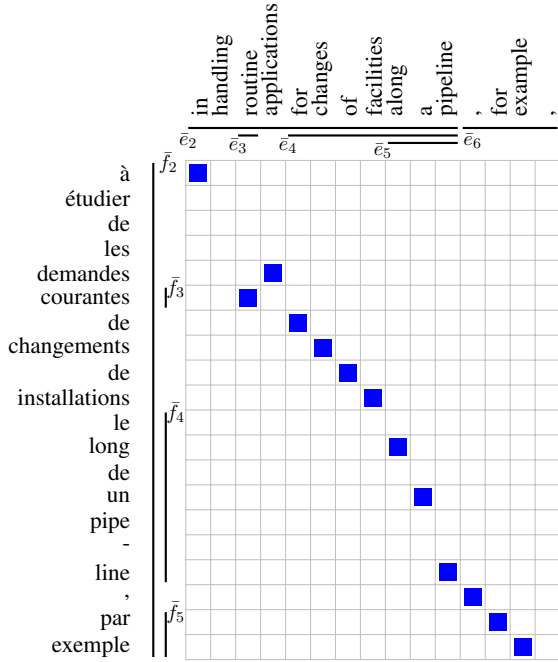
Figure 2: Example with word alignment

projection. To prevent source adjuncts from being aligned to the first target adjunct that subsumes their projection, we also enforce that only non-overlapping source adjuncts may be aligned to a same target sequence, as explained in section 2.2.1.

This procedure results in a many-to-many alignment between adjuncts on either side. We distinguish several types of adjunct pairings through this alignment, which we interpret as divergent, equivalent or weakly equivalent, as described in section 2.2.2.

We perform this alignment in both source-target and target-source directions to measure the proportion of source, respectively target, adjuncts that fall in each category.

2.2.1 Adjunct pairing procedure

We define the projection of an adjunct σ as the unique tuple of maximal, non-overlapping phrases ϕ_1^n that are aligned to σ through the word alignment. Each phrase ϕ_i in this tuple is understood as being extended with possible surrounding unaligned positions–phrases are primarily identified by the aligned positions they cover. And each ϕ_i is maximal as any larger phrase distinct from ϕ_i would also include (aligned) positions not aligned to σ. Let $I(\phi_i)$ be the set of aligned positions in each ϕ_i, and $I(\phi_1^n)$ the set of aligned positions

covered by ϕ_1^n.

We align σ to the non-overlapping sequence of target adjuncts τ_1^m that has the smallest set of aligned positions while having the largest overlap with ϕ_1^n; the overlap of a projection and a target sequence is the intersection of their respective sets of aligned positions. For instance in Figure 2, the projection of $\bar{f}_4$ is maximally covered by $\bar{e}_2$, $\bar{e}_4$, and $\bar{e}_5$; we align the latter to $\bar{f}_4$ as it covers the least aligned positions. In practice, we search through the tree of target adjuncts for adjuncts that overlap with ϕ_1^n, and for each such adjunct τ we compare its overlap with ϕ_1^n to that of the sequence of its children γ_1^k to determine which (of τ or γ_1^k) should be part of the final target sequence.

We perform a similar selection on overlapping source adjuncts that point to the same target sequence. For each source adjunct σ, we determine if its target sequence τ_1^m is also aligned to adjuncts dominated by σ, in which case we compare the overlap of σ's projection with τ_1^n to that of its children in the source adjunct tree to determine which should be aligned to τ_1^m. For instance in Figure 2, $\bar{e}_4$ is aligned to $\bar{f}_2$ (when projecting from English to French), but so is $\bar{e}_2$; as $\bar{e}_2$'s projection overlaps more with $\bar{f}_2$, we discard the alignment between $\bar{e}_4$ and $\bar{f}_2$.

The final alignments for our example are represented in Table 1.

Table 1: Adjunct pairings for the alignment of Figure 2

$f \rightarrow e$		$e \rightarrow f$	
$\bar{f}_2$	$\bar{e}_2, \bar{e}_6$	$\bar{e}_2$	$\bar{f}_2$
$\bar{f}_3$	$\bar{e}_3$	$\bar{e}_3$	$\bar{f}_3$
$\bar{f}_4$	$\bar{e}_5$	$\bar{e}_4$	-
$\bar{f}_5$	$\bar{e}_6$	$\bar{e}_5$	$\bar{f}_4$
		$\bar{e}_6$	$\bar{f}_2$

2.2.2 Types of adjunct pairings

We distinguish three main classes of adjunct translation equivalence: divergent, equivalent and weakly equivalent. We further subdivide each class into two types, as shown in Table 2. Adjunct pairings fall into one of these types depending on their configuration (unaligned, one-to-one or many-to-many) and their agreement with the word alignments. Equivalent types notably differ from weakly equivalent ones by being bijectively

159

aligned; With the notations of section 2.2.1, two adjunct sequences σ_1^n and τ_1^m with respective projections $\phi_1^{n'}$ and $\psi_1^{m'}$ are translation equivalent iff $I(\phi_1^{n'}) = I(\tau_1^m)$ and $I(\psi_1^{m'}) = I(\sigma_1^n)$.

Table 2: Adjunct pairing types

divergent	
`null`	empty projection
`div`	no aligned target adjuncts
weakly equivalent	
`we-nm`	many-to-many non-bijective
`we-11`	one-to-one non-bijective
equivalent	
`eq-nm`	many-to-many bijective
`eq-11`	one-to-one bijective

In Table 1, $\bar{e}_4$'s translation is divergent as it is not aligned to any adjunct; $\bar{f}_5$ and $\bar{e}_6$ are weakly equivalent as the projection of $\bar{f}_5$ does not cover all the aligned positions of $\bar{e}_6$. The pairing from $\bar{f}_2$ to $\bar{e}_2, \bar{e}_6$ is many-to-many equivalent, and so are the pairings from $\bar{e}_2$ and $\bar{e}_6$ to $\bar{f}_2$; the remaining pairings are one-to-one equivalent.

As Table 3 shows, the divergent types `null` and `div` regroup untranslated adjuncts (Example 1) and divergent adjuncts: Examples (2) and (3) show cases of conflational divergence (Dorr, 1994), that appear in different types because of the underlying word alignments; in Example (4), the prepositional phrase *with this task* has been wrongly labelled as an adjunct, leading to a falsely divergent pairing. The weakly-equivalent types `we-nm` and `we-11` regroup both divergent and equivalent pairings: the adjuncts of Examples (5) and (8) are aligned by our method to adjuncts that are not their translation equivalent, the adjunct in Example (6) cannot be aligned to its equivalent because of a parsing error, and the equivalences in Examples (7) and (9) cannot be identified because of a word-alignment error. Finally, we show a number of equivalent pairings (`eq-nm` and `eq-11`): in Example (10), an attachment error in the French parse induces a many-to-one equivalence where there should be two one-to-one equivalences; Examples (11) to (13) show a number of true many-to-many equivalences, while Examples (14) and (15) show that adjuncts may be equivalent across a bitext while belonging to a different syntactic category and modifying a different type of phrase (15).

3 Adjunct identification

We identify adjuncts in dependency trees obtained by conversion from phrase-structure trees: we map modifier labels to adjuncts, except when the dependent is a closed-class word. For English, we use the Berkeley parser and convert its output with the *pennconverter* (Johansson and Nugues, 2007; Surdeanu et al., 2008); for French, we use the Berkeley parser and the functional role labeller of Candito et al. (2010). The pennconverter with default options and the French converter make similar structural choices concerning the representation of coordination and the choice of heads.

3.1 English adjuncts

We first identify closed-class words by their POS tag: CC, DT, EX, IN, POS, PRP, PRP$, RP, SYM, TO, WDT, WP, WP$, WRB. Punctuation marks, identified by the `P` dependency relation, and name dependencies, identified by `NAME`, `POSTHON`, or `TITLE`, are also treated as closed-class words.

Adjuncts are identified by the dependency relation: `ADV`, `APPO`, `NMOD` (except determiners, possessives and 'of' complements), `PRN`, `AMOD` (except when the head is labeled with `ADV`) and `PMOD` left of its head. Cardinals, identified by the CD POS tag, and remaining dependents are classified as arguments.

3.2 French adjuncts

Closed-class words are identified by the (coarse) POS tags: C, D, CL, P, PONCT, P+D, PRO. Auxiliary verbs, identified by the dependency relations `aux_tps` and `aux_pass`, are also included.

Adjuncts are identified by the dependency relations `mod_rel` and `mod` (except if the dependent's head is a cardinal number, identified by the `s=card` label).

3.3 Evaluation

We evaluate adjunct identification accuracy using a set of 100 English and French sentences, drawn randomly from the Europarl corpus. A single annotator marked adjuncts in both sets, identifying slightly more than 500 adjuncts in both sets. We find F scores of 71.3 and 72.2 for English and French respectively, as summarized in Table 4. We find that about a quarter of errors are related to parse attachment, yielding scores of 77.7 and 78.6 if one corrects them.

Table 3: Examples of adjunct pairing types

```
null
```
| (1) | it is **indeed** a great honour | vous me faites un grand honneur |
| (2) | the **polling** booths | les isoloirs |

```
div
```
| (3) | the **voting** stations | les isoloirs |
| (4) | to be entrusted **with this task** | en me confiant cette tâche |

```
we-nm
```
(5)	reforms **to the Canadian military**	réformes *des forces* [**armées**] [**canadiennes**]
(6)	an **even greater** country	un pays [**encore**] [**plus**] *magnifique*
(7)	**in safe communities**	[**en sécurité**] [**dans nos communautés**]

```
we-11
```
| (8) | **across the land** | *de* **tout** *le pays* |
| (9) | **strong** opinions | des opinions **bien arrêtées** |

```
eq-nm
```
(10)	a **proud** moment **for Canada**	un moment **heureux pour le Canada**
(11)	we have used the **wrong** process	nous **ne** suivons **pas** le **bon** processus
(12)	our **common** space and our **common** means	un espace et des moyens **communs**
(13)	the [**personal**] [**protected**] files	les dossiers **confidentiels et protégés**

```
eq-11
```
| (14) | the names **just announced** | les noms **que je viens de mentionner** |
| (15) | one in three **Canadian** jobs | **au Canada** , un emploi sur trois |

Table 4: Adjunct identification F scores

		prec.	recall	F
En	auto.	66.2	77.2	71.3
	corr.	72.3	84.0	77.7
Fr	auto.	68.1	76.7	72.2
	corr.	74.7	83.0	78.6

4 Experiments

4.1 Experimental set-up

We measure adjunct translation equivalence in four data sets: the manually-aligned Canadian Hansards corpus (Och and Ney, 2003), containing 447 sentence pairs, the house and senate training data of the Canadian Hansards (1.13M sentence pairs), the French-English Europarl training set (1.97M sentence pairs) and the Moses news-commentaries corpus (156k sentence pairs). Besides, we randomly selected 100 sentence pairs from the Europarl set to measure adjunct identification accuracy as reported in section 3 and adjunct correspondence with gold adjunct annotations.

All four corpora except the manual Hansards are preprocessed to keep sentences with up to 80 words, and all four data sets are used jointly to train unsupervised alignments, both with the Berkeley aligner (Liang et al., 2006) and GIZA++ (Brown et al., 1993; Och and Ney, 2003) through mgiza (Gao and Vogel, 2008), using 5 iterations of Model 1 and 5 iterations of HMM for the Berkeley aligner, and 5 iterations of Model 1 and HMM and 3 iterations of Model 3 and Model 4 for GIZA++. The GIZA++ alignments are symmetrized using the grow-diag-final heuristics. Besides, the manual Hansards corpus is aligned with Sure Only (SO) and Sure and Possible (SP) manual alignments.

4.2 Measurements with gold adjunct annotations

We compared adjunct translation equivalence of automatically identified adjuncts and gold annotations using 100 manually annotated sentence pairs from the Europarl corpus; adjuncts were aligned automatically, using the Berkeley word alignments. We also measured adjunct equivalence using automatic adjunct annotations corrected for parse attachment errors, as introduced

in section 3.3. Table 5 reports harmonic mean figures (m_h) for each adjunct projection type. For information, we also report their decomposition in the case of gold annotations, showing some dependence on the projection direction.

Table 5: Translation equivalence of automatic, rebracketed and gold adjuncts

| | auto. | corr. | gold | | |
	m_h	m_h	ef	fe	m_h
null	7.6	7.7	8.1	7.3	7.7
div	22.3	22.5	14.7	12.0	13.2
we-nm	10.8	9.6	2.7	4.6	3.4
we-11	12.5	10.8	7.4	8.5	7.9
eq-nm	3.5	2.2	2.5	3.3	2.9
eq-11	41.8	45.8	64.5	64.3	64.4

About two thirds of manually identified adjuncts form equivalent pairs, representing a gain of 20 points with regard to automatically identified adjuncts. This is accompanied by a halving of divergent pairings and of weakly equivalent ones. Further, we find that about half of the remaining weak equivalences can be interpreted as translation equivalent (to compare to an estimated third for automatically identified adjuncts), allowing us to estimate to 70% the degree of translation equivalence given Berkeley word alignments in the Europarl corpus.

4.3 Measurements with manual and automatic alignments

We aligned adjuncts in the manual Hansards corpus using all four word alignments. Table 6 presents the mean proportions for each category of adjunct projection.

Table 6: Translation-equivalence of adjuncts in the manual Hansards

	SO	SP	bky	giza
null	32.1	2.8	8.7	3.3
div	19.7	29.3	27.1	30.3
we-nm	3.4	14.6	8.5	11.4
we-11	5.7	13.8	13.5	15.3
eq-nm	4.1	7.3	4.1	4.2
eq-11	33.7	31.8	37.6	35.3

Comparing the mean proportions per type be-

tween the four alignments, we see that a third of adjuncts on either side are not aligned at all with the sure-only manual alignments. In the example of Figure 2 for instance, these alignments do not link $\bar{f}_3$ to $\bar{e}_3$. On the other hand, the sure and possible manual alignments lead to many divergent or weakly equivalent pairings, a result of their dense phrasal alignments. In comparison, the automatic alignments connect more words than the sure-only alignments, leading to a mixed result for the adjunct pairings: one gains more translation-equivalent, but also more divergent and weakly equivalent pairs. In this, the Berkeley aligner appears less noisy than GIZA++, as it captures more translation equivalent pairs and less weakly equivalent ones. This is confirmed in the other data sets too, as Table 7 shows.

Table 7: Mean proportions of adjunct-pairing types in automatically aligned data

| | hans-hst | | europarl | | news | |
	bky	giza	bky	giza	bky	giza
null	7.5	2.7	6.3	2.3	8.3	3.3
div	28.1	30.8	21.8	24.2	21.0	23.9
we-nm	10.4	12.2	11.0	12.7	10.6	12.6
we-11	13.4	15.5	12.4	14.6	11.7	14.2
eq-nm	3.2	4.0	3.2	4.0	3.1	3.8
eq-11	37.1	34.6	45.0	42.0	44.9	41.8

Comparing figures between the different data sets, we see that the Europarl and the News data have more translation-equivalent and less divergent adjuncts than the Hansards training data (hans-hst). Taking the harmonic mean for both equivalent types (eq-nm and eq-11), we find that 48.2% of adjuncts have an adjunct translation equivalent in the Europarl data (with the Berkeley aligner) and 48.0% in the News corpus, against 40.3% the Hansards training set and 41.6% in the manual Hansards set. This suggests that translations in the Hansards data are less literal than in the Europarl or the News corpus.

4.4 Effect of sentence length

We explore the relation between sentence length and translation equivalence by performing measurements in bucketed data. We bucket the data using the length of the English sentences. Measurements are reported in Table 8 for the Hansards

Table 8: Adjunct translation equivalence with the Berkeley aligner in bucketed data

	hans-man		hansard-hst				europarl			
	1-15	16-30	1-15	16-30	31-50	51-80	1-15	16-30	31-50	51-80
`null`	9.3	8.5	6.5	7.6	7.8	8.0	6.4	6.0	6.2	6.6
`div`	28.1	25.9	39.5	25.3	23.5	22.6	25.3	22.2	21.2	20.6
`we-nm`	6.1	9.4	5.3	10.1	13.6	16.7	5.0	9.3	12.5	14.9
`we-11`	11.8	14.1	12.2	13.4	14.2	14.8	10.0	11.7	13.0	13.9
`eq-nm`	3.1	4.5	2.8	3.4	3.3	3.1	3.4	3.3	3.1	2.9
`eq-11`	40.6	36.3	32.5	39.6	37.3	34.4	49.1	47.1	43.7	40.7

and the Europarl sets (the News set yields similar results to the Europarl data).

All data sets show a dramatic increase of the proportion of adjuncts involved in many-to-many, and to a lesser extent one-to-one weakly equivalent translations. This increase is accompanied by a decrease of all other adjunct-pairing types (unaligned adjuncts excepted), and is likely to result from increased word-alignment and parsing errors with sentence length.

A rather surprising result is the high proportion of divergent adjunct translations in the shorter sentences of the Hansards training set; we find the same phenomenon with the GIZA++ alignment. We attribute this effect to the Hansards set having less literal translations than the other sets. That we see this effect mostly in shorter sentences may result from translation mismatches being mostly local. As sentence length increases however, word and adjunct alignment errors are also likely to link more unrelated adjuncts, resulting in a drop of divergent adjuncts.

4.5 Simplifying alignments

We perform a simple experiment to test the effect of word-alignment simplification of adjunct translation equivalence. For this we remove alignment links between function words (as defined in section 3) on both sides of the data, and we realign adjuncts using these simplified alignments. Table 9 shows that this simplification (column '-fw') slightly decreases the proportion of weakly equivalent pairings with regard to the standard alignment ('std'), mostly to the benefit of translation-equivalent pairings. This suggests that further gains may be obtained with better alignments.

Table 9: Effect of alignment simplification on adjunct translation equivalence in the Europarl data

	bky		giza	
	std	-fw	std	-fw
`null`	6.3	7.5	2.3	3.1
`div`	21.8	21.5	24.2	24.0
`we-nm`	11.0	9.1	12.7	10.8
`we-11`	12.4	10.0	14.6	13.2
`eq-nm`	3.2	4.0	4.0	4.8
`eq-11`	45.0	47.5	42.0	43.7

5 Related work

While adjunction is a formal operation that may be applied to non-linguistic adjuncts in STAG, De-Neefe and Knight (2009) restrict it to syntactic adjuncts in a Synchronous Tree Insertion Grammar. They identify complements using (Collins, 2003)'s rules, and regard all other non-head constituents as adjuncts. Their model is able to generalize to unseen adjunction patterns, and to beat a string-to-tree baseline in an Arabic-English translation task.

Arnoult and Sima'an (2012) exploit adjunct optionality to generate new training data for a phrase-based model, by removing phrase pairs with an English adjunct from the training data. They identify adjuncts using syntactic heuristics in phrase-structure parses. They found that few of the generated phrase pairs were actually used at decoding, leading to marginal improvement over the baseline in a French-English task. They also report

figures of role preservation for different categories of adjuncts, with lower bounds between 29% and 65% and upper bounds between 61% and 78%, in automatically aligned Europarl data. The upper bounds are limited by discontinuous adjunct projections, while the estimation of lower bounds is limited by the lack of adjunct-identification means for French.

There has been a growing body of work on exploiting semantic annotations for SMT. In many cases, predicate-argument structures are used to provide source-side contextual information for lexical selection and/or reordering (Xiong et al., 2012; Li et al., 2013), without requiring cross-linguistic correspondence. When correspondence between semantic roles is required, predicates are commonly aligned first. For instance, Lo et al. (2012) use a maximum-weighted bipartite matching algorithm to align predicates with a lexical-similarity measure to evaluate semantic-role correspondence. Padó and Lapata (2009) use the same algorithm with a similarity measure based on constituent overlap to project semantic roles from English to German.

6 Conclusion

In this paper we presented the first study of translation equivalence of adjuncts on a variety of French-English parallel corpora and word alignments. We use a method based on overlap to derive many-to-many adjunct pairings, that are interpretable in terms of translation equivalence.

We found through measurements in French-English data sets that 40% to 50% of adjuncts–depending on the data–are bijectively aligned across a bitext, whereas about 25% more adjuncts align to adjuncts, albeit not bijectively. We estimate that a third of these weakly equivalent links represent true, adjunct translation equivalences.

With manually identified adjuncts, we found that about 70% have adjunct translation-equivalents in automatically aligned data. These are fairly low results if one considers that French and English are relatively close syntactically. So while they show that adjunct labelling accuracy on both sides of the data is crucial for adjunct alignment, and that applications that exploit adjunction can gain from decreasing their dependence on word alignments and idealized experimental conditions, they call for better understanding of the factors behind translation divergence.

In fact, as a remaining quarter of adjuncts have divergent translations, it would be interesting to determine, for instance, the degree to which divergence is caused by lexical conflation, or reflects non-literal translations.

References

Sophie Arnoult and Khalil Sima'an. 2012. Adjunct Alignment in Translation Data with an Application to Phrase-Based Statistical Machine Translation. In *Proceedings of the 16th Annual Conference of the European Association for Machine Translation*, pages 287–294.

Peter F. Brown, Stephen A. Della Pietra, Vincent J. Della Pietra, and Robert L. Mercer. 1993. The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(2):263–311.

M.-H. Candito, B. Crabbé, and P. Denis. 2010. Statistical French dependency parsing: treebank conversion and first results. In *Proceedings of The seventh international conference on Language Resources and Evaluation (LREC)*.

Michael Collins. 2003. Head-driven statistical models for natural language parsing. *Computational Linguistics*, 29(4):589–637.

Steve DeNeefe and Kevin Knight. 2009. Synchronous Tree Adjoining Machine Translation. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 727–736.

Bonnie J. Dorr. 1994. Machine Translation Divergences: A Formal Description and Proposed Solution. *Computational Linguistics*, 20(4):597–633.

Qin Gao and Stephan Vogel. 2008. Parallel Implementations of Word Alignment Tool. In *Software Engineering, Testing, and Quality Assurance for Natural Language Processing*, pages 49–57, Columbus, Ohio, June. Association for Computational Linguistics.

Rebecca Hwa, Philip Resnik, Amy Weinberg, and Okan Kolak. 2002. Evaluating Translational Correspondence Using Annotation Projection. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, pages 392–399.

Richard Johansson and Pierre Nugues. 2007. Extended Constituent-to-dependency Conversion for English. In *Proceedings of NODALIDA 2007*, pages 105–112, Tartu, Estonia, May 25-26.

Aravind K. Joshi, Leon S. Levy, and Masako Takahashi. 1975. Tree adjunct grammars. *Journal of Computer and System Sciences*, 10(1):136–163.

Junhui Li, Philip Resnik, and Hal Daumé III. 2013. Modeling Syntactic and Semantic Structures in Hierarchical Phrase-based Translation. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 540–549, Atlanta, Georgia.

Percy Liang, Ben Taskar, and Dan Klein. 2006. Alignment by Agreement. In *Proceedings of the Main Conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*, HLT-NAACL '06, pages 104–111.

Chi-kiu Lo, Anand Karthik Tumuluru, and Dekai Wu. 2012. Fully Automatic Semantic MT Evaluation. In *Proceedings of the Seventh Workshop on Statistical Machine Translation*, WMT '12, pages 243–252.

Franz Josef Och and Hermann Ney. 2003. A Systematic Comparison of Various Statistical Alignment Models. *Computational Linguistics*, 29:19–51.

Sebastian Padó and Mirella Lapata. 2009. Cross-lingual Annotation Projection for Semantic Roles. *Journal of Artificial Intelligence Research*, 36:307–340.

Stuart Shieber and Yves Schabes. 1990. Synchronous tree-adjoining grammars. In *Handbook of Formal Languages*, pages 69–123. Springer.

Mihai Surdeanu, Richard Johansson, Adam Meyers, Lluís Màrquez, and Joakim Nivre. 2008. The CoNLL-2008 Shared Task on Joint Parsing of Syntactic and Semantic Dependencies. In *CoNLL 2008: Proceedings of the Twelfth Conference on Natural Language Learning*, pages 159–177, Manchester, United Kingdom.

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

Deyi Xiong, Min Zhang, and Haizhou Li. 2012. Modeling the Translation of Predicate-Argument Structure for SMT. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pages 902–911.