Looking up phrase rephrasings via a pivot language

Aurélien Max
LIMSI-CNRS & Université Paris-Sud 11
Orsay, France
`aurelien.max@limsi.fr`

Michael Zock
LIF-CNRS
Marseilles, France
`michael.zock@lif.univ-mrs.fr`

Abstract

Rephrasing text spans is a common task when revising a text. However, traditional dictionaries often cannot provide direct assistance to writers in performing this task. In this article, we describe an approach to obtain a monolingual phrase lexicon using techniques used in Statistical Machine Translation. A part to be rephrased is first translated into a pivot language, and then translated back into the original language. Models for assessing fluency, meaning preservation and lexical divergence are used to rank possible rephrasings, and their relative weight can be tuned by the user so as to better address her needs. An evaluation shows that these models can be used successfully to select rephrasings that are likely to be useful to a writer.

1 Introduction

Once an initial draft of a text is ready, writers face the difficult phase of *text revision*. Changes may be made for various reasons: correcting spelling or grammatical errors, making the text locally more fluent (for example, in case it contains wordings that are literal translations from another language), avoiding close repetitions or enforcing terminological consistency, or better conveying the writer's ideas. All these changes can affect text spans of various sizes, and can globally be seen as cases of *rephrasing*. Paraphrasing involves rephrasings

that are semantically equivalent, but targets terminology and style that are more suited to the context of use of a text. In a broad sense, rephrasing may involve wordings that convey different meanings in an attempt to correct or make the writer's thoughts more precise. Research concerned with the study of changes between writers' drafts (*textual genetic criticism*) can help in understanding writers' rewriting processes, and can be supported by automatic tools (e.g. (Bourdaillet et al., 2007)).

In this work, we address the issue of how writers can be assisted in finding wordings that correspond to multi-word phrases of any nature. Given an original text span, the writer is presented with a list of rephrasings that are organized by taking into account the context of the rephrasing and user-specified preferences. Our proposal can therefore be used as a lexicon operating at the phrasal level, which can be used either when writers are faced with a tip-of-the-tongue lexical access problem, or when they are not completely satisfied with some initial wording. In the former case, they may be able to come up with some words or phrases that would be different in meaning from what they are looking for, and in the latter they may be looking for a near-synonymous wording that is more appropriate to a given context, for example to avoid close repetitions. To define such a phrase lexicon and its possible mode of use, the following questions should be considered: (a) how the lexicon entries are obtained, (b) what can be the entry points and how can one navigate in the results, and (c) how the results are displayed.

Rephrasing can be more or less complex and problematic depending on the consequences at the various levels:

- In the simplest case, replacing one element

by another does not have any consequences overall. This is often the case when a word is replaced by its synonym or a similar word.

- An entire expression or sentence is replaced by its equivalent. In this case the problem is generally to obtain a good fit with regard to the surrounding text, the replacing unit being well-formed by definition.

- The replacing element may require syntactic changes of the matrix, i.e. the text in which it is embedded. This occurs if the source word and the target word have different syntactic requirements, and this can be seen as a good reason to replace entire sentences, or at least sentence fragments. This assumes a pattern dictionary, where patterns achieving the same conceptual goal are grouped together.

In the next section, we discuss limitations of traditional dictionaries with respect to the targeted task, and describe an approach to obtain phrase rephrasings through a pivot translation into another language. In section 3, we discuss the issue of the organization of the results along various axis: fluency of rephrasings, preservation of meaning, and lexical divergence between original text spans and rephrasings. We then present an initial evaluation of our approach on French rephrasing in section 4. Related work is presented in section 5, and we finally discuss our approach and our future work in section 6.

2 Lexicon of phrase rephrasings

Dictionaries and semantic resources such as thesauri can be used to find words by following links of different kinds from a given entry point. WordNet (Fellbaum, 1998) is one such resource. For a proposal of other kinds of links and navigational aids see also (Zock and Bilac, 2004; Zock, 2006; Zock, 2007).

Words are the traditional units that people expect to find in dictionaries. Whereas some types of dictionaries can contain multiword expressions, such as compound nouns and terms, those correspond to linguistically-motivated units. In order to rephrase phrases of any type with a dictionary, a writer may have to look up several words, combine various information and validate the result using her experience of the language or throught the use of a concordancer. Moreover, dictionary lookups

are in most cases insensitive to the actual context of words in an existing text. It is therefore the responsibility of its users to ensure that a choice is appropriate for a given context, which can be quite difficult, for example when writing in a second language.

One way of obtaining phrase rephrasings is by looking at phrases that occur in similar contexts in a monolingual corpus (e.g. (Munteanu and Marcu, 2006)). In order to extract a comprehensive phrase lexicon, a very large number of sentences should be compared to extract potential rephrasings, which furthermore may often correspond to phrases that are too remotely connected. Parallel corpora provide the interesting advantage that it is reasonable to assume that elements from one side of the corpus should be aligned to elements on the other side, and that associations of elements can be reinforced by the number of times they occur in the corpus. Various approaches for word alignment from parallel corpora have been proposed (see e.g. (Och and Ney, 2003)), and the phrase-based approach to Statistical Machine Translation (Koehn et al., 2003) has led to the development of heuristics for obtaining alignments between phrases of any number of words.

Unfortunately, monolingual parallel corpora aligned at the sentence level, such as various translations of a novel in a foreign language, are resources that are extremely scarce. Using bilingual parallel corpora, a much more common resource, one can obtain various possible phrase translations for a given source phrase, as well as some estimate of the distribution of probabilities for the various translations of that phrase. Such $N \rightarrow M$ alignements can capture lexical translations (e.g. *exigeons → ask for, call for, demand, expect, request,* etc.) and phrasal literal or idiomatic translations (e.g. *un bon début → a good approach, a good first move, a good starting point, a positive initiative, an encouraging start, the right road,* etc.), but can also capture noise depending on the alignment heuristics used (e.g. *les états candidats (candidate countries) → Member States, the candidate countries were to, the accession countries have called for, candidate, the,* etc.) Different target phrases associated with a given source phrase can either represent paraphrases or phrases with different meanings. Among the limitations of this type of phrasal alignments are their inability to model non-consecutive words and to generalize the con-

tents of phrases, and the fact that their translations are not conditioned on their context.

If phrase extraction is performed in two opposite directions, then it is possible to find the possible translations of a given phrase (and their conditional probabilities), and then to translate back those phrases into the original language. In this approach proposed by (Bannard and Callison-Burch, 2005), the second language acts as a pivot, as illustrated on figure 1. Because of the nature of the possible alignments, this pivot can represent various senses, which in context can be equivalent or comparable to that of the original phrase. In turn, the same phenomena can take place when translating back from the pivot phrases to the original language, and the resulting rephrasings can be equivalent or comparable in meaning to that of the original phrase in some context, may also be incomplete and/or require other changes in the rephrased sentence.

Bannard and Callison-Burch have defined a *paraphrase probability* between two phrases p_1 and p_2 (with $p_1 \neq p_2$) that uses conditional probabilities between phrases and sums over all possible pivot phrases:

$$P(p_2|p_1) = \arg\max_{p_2 \neq p_1} \sum_{pivot} P(pivot|p_1)P(p_2|pivot)$$

(1)

(Callison-Burch, 2007) measured the importance of various factors impacting the quality of the paraphrases obtained. Using manually built alignments yields a significant improvement in paraphrase quality, showing that if better alignments are available the proposed approach can produce better paraphrases. Alignments between several languages can be used for finding pivot phrases, and using several simulateously tend to improve alignment quality and therefore paraphrases themselves. Using a language model to find paraphrases that maximize its score in the original sentencial context leads to improved fluency, but has a negative impact on meaning preservation. Lastly, restricting pivot phrases to those actually aligned in a test aligned bilingual corpus improves paraphrase quality, which illustrates the importance of disambiguating source phrases relatively to the pivot language.

The rephrasings obtained can be classified into several categories when used in context:

- A rephrasing can be a paraphrase that is valid in all contexts (e.g. *je vous donne raison* → *je suis d'accord avec vous*), in specific grammatical contexts (e.g. *pouvoir accueillir dans de bonnes conditions les pays* → *comme il se doit*) and/or pragmatic contexts (e.g. *c'est un bon début* → *nous partons du bon pied*).

- A rephrasing can contain shifts in meaning with the original phrase which might be acceptable or not (e.g. *nous voulons apporter notre contribution à ce débat* → *donner de la valeur*). Some such rephrasings reveal a natural bias towards the bilingual corpus used (e.g. *le prochain élargissement constitue la principale tâche* → *l' objectif principal*).

- A rephrasing can be ill-formed but still contain elements of interest to a writer (e.g. *ceux qui disent que ... se trompent* → *devrions à nouveau réfléchir*; here a rephrasing such as *devraient à nouveau réfléchir* could be deemed acceptable in some contexts).

- A rephrasing may introduce a contradiction in a specific context (e.g. *ce n'est pas le moment de se montrer hésitant* → *il est trop tôt pour*)

- A rephrasing may be inexploitable because it is syntactically ill-formed in context and does not contain any element of interest, or is too close to the original phrase.

The most natural entry point to such a resource is by entering a phrase or selecting it in a text under revision. Approximate search can also be of use, as done in some concordancer software, for example by allowing the user to enter word-based regular expressions mixing literal words, word lemmas, word part-of-speech or even word classes (e.g. types of named entities). Boolean queries on indexes of word lemmas can also be used to offer yet more flexibility to search the lexicon, but at the cost of more candidate results. Once results are returned, they can recursively be reused as source phrases, so as to offer a means to navigate by iterative refining.

3 Evaluation of rephrasings in context for ranking results

Each candidate phrase rephrasing for a given phrase must be evaluated in order to define a ranking order for presentation to the user, and possibly

```
                                          ce moment, il est inopportun de
              this is not the time to     ce n'est pas l'heure de
                                          ~~ce n'est pas le moment de~~
              now is not the time to      il est trop tôt pour
                                          l'heure n'est pas au
ce n'est pas le moment de                 il est à présent temps d'
              this is not time for        il n'est pas judicieux de
                                          le moment n'est pas venu de
              it is inappropriate to      il est encore trop tôt pour
                                          le moment est mal choisi
              ...                         ...
```
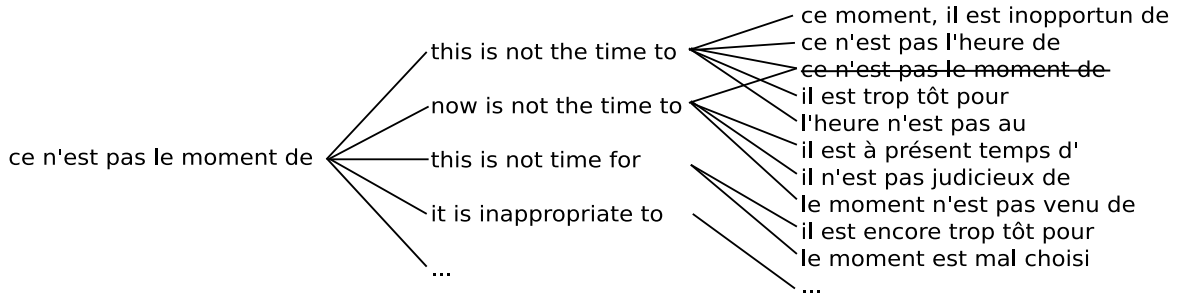
Figure 1: Example of rephrasing for the French phrase *ce n'est pas le moment de* using English as pivot.

to discard some of them. The proposed ranking should reflect as best as possible the preferences of the user for the task at hand in order to minimize reading time and maintain the user's interest in using the phrase lexicon. It is essential to give the user some control over how the results are returned depending on what is more important to her. For example, (Ferret and Zock, 2006) have proposed to present results from a dictionary enriched with topical associations in chunks to allow for categorial search. There will be cases where the user may find acceptable only grammatical results, while in other cases the user might accept agrammatical results provided they contain interesting suggestions. Moreover, it seems extremely important that result ranking can take into account the phrase substitution into the original context.

Considering how the proposed phrase lexicon is built, the pivot paraphrasing probability of equation 1 (PIV) can be used as a baseline ordering. Such a model reflects some strength of association between a rephrased phrase and the original phrase using the extracted phrases and conditional probabilities derived from a bilingual training corpus. It is therefore expected that results will be biased towards that corpus if the latter belongs to a particular genre or theme. Nonetheless, one can expect that some associations will be general enough to be of general interest.

In addition, several models that users can interpret as ranking criterion can be used simulateneously using the log-linear framework traditionally used in SMT systems. However, contrary to what is done in SMT, the weight of the models cannot be automatically optimized if we do not use an automatic evaluation of rephrasing quality, the definition of which depending heavily on the subjective appreciation of a user. Equation 2 shows how the score of a rephrasing p_2 of p_1 can be com-

puted, where M is the set of models used, h_m is the logarithm of the normalized score of a model and λ_m its weight (with $\sum_{m \in M} \lambda_m = 1$), and C is the original sentence and the placeholder for the rephrased phrase.

$$s(p_2, p_1, C) = \sum_{m \in M} \lambda_m h_m(p_1, p_2, C) \quad (2)$$

3.1 Control over fluency

As noted by (Mutton et al., 2007), the notion of sentence-level fluency is not uniformly agreed upon, and its evaluation by human judges is sometimes found subjective, but in practice judges can obtain high levels of agreement about what can be considered fluent or not. Like (Callison-Burch, 2007), we can use a language model (LM) to assess the local fluency of a sentence after a phrase has been substituted with a rephrasing. A degradation in score (with a fluent original sentence) can indicate that the rephrasing segment should be adapted to the sentence, and/or that the sentence itself should be modified in order to integrate the new phrase as is.

Syntax parsers can produce various information that can be relevant for assessing the fluency of sentences, which can be used as features from different parsers for classification that can correlate well with human judgment (Mutton et al., 2007). When substituting a part of a sentence with another phrase and if this substitution does not require other changes in the sentence, then at least the dependency relationships between words outside that phrase should be preserved. This seems coherent with our objective of focussing on the task of phrase rephrasing when it is possible to modify only a given phrase and obtain an acceptable result.

80

3.2 Control over meaning preservation

The preservation of dependency relationships outside of the rephrased phrase can also play a role in terms of meaning preservation. Dependency relationships connecting words in the phrase and words outside the phrase (i.e., whose governor is outside the phrase and dependant inside it, or the opposite) should still exist after such a substitution, but possibly with a modified dependency target in the phrase. Indeed, those relationships denote the grammatical role of the words of the phrase relative to their context, and if those are preserved then it is more likely that meaning is preserved.

We use a model based on dependency preservation (DEP) which involves relationships outside the rephrased phrase and relationships crossing a boundary of that phrase. The score is based on some proportion of the number of such dependencies found after substitution over the number of original dependencies (see (Max, 2008) for details). Another way of controlling for meaning preservation is to ensure that only the pivot phrases with the same meaning as the original phrase are kept (and then their back translations). (Callison-Burch, 2007) has shown the positive impact on paraphrase quality of using a controlled pivot present in an aligned sentence in a test bilingual corpora. Phrase disambiguation techniques have been proposed for SMT and could be applied to the problem at hand (e.g. (Stroppa et al., 2007)). In an interactive context, it makes sense to let the user the opportunity to control for phrase sense by rejecting bad pivot phrases if she wants to, which is then similar to Callison-Burch's experiment settings. This manual selection must of course be optional, but can be used when a user prefers a stricter control on meaning. Another possibly interesting use is to disambiguate in a pivot language corresponding to one's native language when writing in a foreign language.

3.3 Control over lexical divergence

There will be cases when possible rephrasings will be very close to their original phrase, differing for example by only punctuation marks or verbal forms[1]. Writers may sometimes prefer rephrasings that differ by just one word, or on the contrary rephrasings that use a set of completely different words. To account for differents words be-
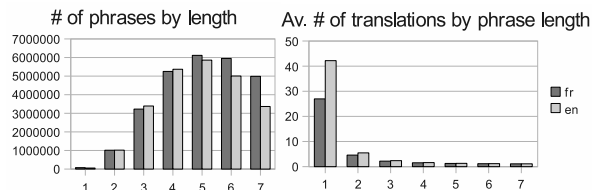

Figure 2: Bilingual phrase lexicon statistics

tween an original phrase and its rephrasing, we use a model (LEM) that returns a proportion of lemmas for full words that only belong to a rephrasing over all such lemmas for an initial phrase and its rephrasing (see (Max, 2008)).

4 Experiments and evaluation

We carried out an evaluation on the local rephrasing of French sentences, using English as the pivot language.[2] We extracted phrase alignments of up to 7 word forms using the Giza++ alignment tool (Och and Ney, 2003) and the `grow-diag-final-and` heuristics described in (Koehn et al., 2003) on 948,507 sentences of the French-English part of the Europarl corpus (Koehn, 2005) and obtained some 42 million phrase pairs for which probabilities were estimated using maximum likelihood estimation. Statistics for the extracted lexicons are reported on figure 2. Entries of the monolingual phrase lexicon are built dynamically from the entries of the monolingual lexicons.

For the LM model, we used a 5-gram language model trained on the French part of the corpus using Kneser-Ney smoothing. The robust parser for French SYNTEX (Bourigault et al., 2005) was used to obtain lemmas for word and labeled dependency relationships between words, used respectively for the LEM and DEP models. Robust parsers provide the advantage that they can provide partial analysis for correct chunks in agrammatical sentences, but they can also recover information from agrammatical chunks which can be undesirable in this case.[3]

A test corpus of 82 sentences that were not used for extracting phrase alignments and learning the

---

[1]This is particularly the case when aligning between low and highly inflected languages.

[2]The main motivation for this choice was that we could easily have access to French native speakers for manual evaluation. We plan however to start new experiments using English, as well as experiments using another highly inflected language as pivot such as Spanish.

[3]We intend to use several parsers for English implementing different approaches as in (Mutton et al., 2007), but we had access to only one parser for French.

language model was built. A human judge selected one phrase of length 3 words or more per sentence that would be a good candidate for rephrasing, and which was accepted if it belonged to the French-English lexicon[4]. We kept at most the 20 first rephrasings obtained using the baseline PIV model, and asked two French native speakers to evaluate on a 5-level scale each the 1648 reformulated sentences obtained on *fluency*, *meaning preservation*, and *authoring value*, where the latter was described in the following way: (5) the rephrasing can be directly reused for revising a text, (4) the rephrasing can be used with a minor change, (3) the rephrasing contains elements that could be used for a good rephrasing, (2) the rephrasing contains elements that could suggest a rephrasing, and (1) the rephrasing is useless.

After the judges had completed manual annotation, smoothing of the scores was done by keeping mean scores for each sentence. We measured a value of 0.59 standard deviation for score differences between judges for grammaticality, 0.7 for meaning preservation and 0.8 for authoring value. Those values can indicate a growing difficulty in judging those characteristics, and in particular that judging authoring value on the proposed scale is more dependant on personal judgment. Results of mean scores for the first rank solutions with various model combinations with uniform weights are reported on figure 3, and results for mean authoring value scores depending on the number of top results presented to the user are reported on figure 4.

Authoring value scores are lower, which can be explained by the fact that rephrasings with bad fluency and/or meaning preservation scores will penalize authoring value scores according to our scale. The best results are obtained when combining all models, which remains true when considering mean results up to at least 8 rephrasings.

The baseline PIV model seems to have the most impact, but all other models also contribute in different ways. This suggests that which model should be used (or its weight in our framework) could be chosen by a user. In the following example, the LEM model helped select a rephrasing which obtained good scores:

Original sentence: *ce que je vous propose donc,*

---

[4]This is a limitation of our evaluation, as our annotator was not strictly speaking revising a text that she wrote. We hope to be able to conduct task-based experiments in the future.

	fluency	meaning	authoring
PIV (baseline)	4.46	4.18	3.62
LM	4.28	3.62	3.45
DEP	4.35	3.68	3.43
LEM	4.05	3.21	3.28
PIV+LM	**4.65**	4.06	3.82
PIV+DEP	4.58	**4.27**	3.66
PIV+LEM	4.37	4.00	3.76
LM+DEP	4.49	3.81	3.68
LM+LEM	4.28	3.59	3.56
PIV+LM+DEP	**4.65**	4.05	3.92
PIV+LM+LEM	4.61	4.02	3.97
PIV+DEP+LEM	4.57	4.17	**4.02**
LM+DEP+LEM	4.37	3.69	3.64
PIV+LM+DEP+LEM	**4.68**	4.09	**4.05**

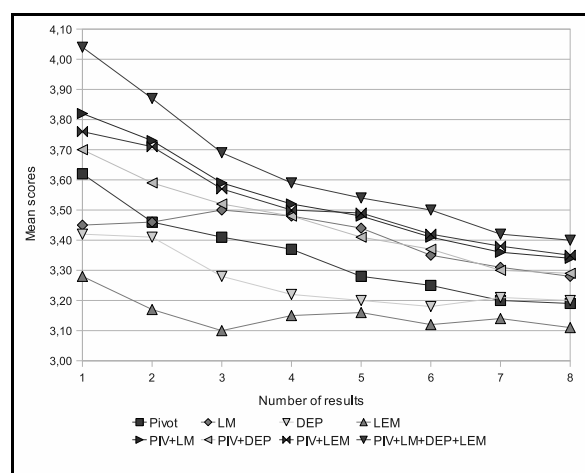Figure 3: Mean results at first rank for various model combinations (uniform weighting)

Figure 4: Mean authoring value scores depending on the number of results presented to the user

c'est de travailler dans cette direction ... (what I therefore propose is to work towards this ...)
Rephrased sentence: *ce que je vous propose donc, c'est de coopérer dans ce sens ... (work towards this goal ...)*

Figures 5 and 6 show two examples of rephrasings in French, whereby for each rephrasing the ranks given by PIV, LM and the combination of all mentioned models are shown.

5 Related work

While the traditional view of lexicons is word-based, we may as well consider larger units, including sentences. Corpus Pattern Analysis (CPA) (Hanks and Pustejovsky, 2005) is concerned with the prototypical syntagmatic patterns with which words in use are associated. For example, the meaning of *take place* is different from the mean-

Rephrasings	Ranks given by model(s)		
	PIV	LM	PIV+LM+DEP+LEM
quelques points essentiels	1	3	1
les points essentiels	19	1	2
plusieurs questions importantes	17	4	3
des points essentiels	8	6	4
deux ou trois questions importantes	5	9	5
plusieurs points importants	11	2	5
un certain nombre de questions importantes	17	7	7
certains points importants	2	5	8
un certain nombre de points importants	3	8	9
certains éléments très importants	13	11	10
une série de points importants	4	12	11
quelques accents importants	5	15	11
des choses extrêmement importantes	13	14	11
quelques remarques importantes ,	8	16	14
des points importants	12	10	15
quelques choses très importantes	13	17	16
certains points importants ,	8	13	17
quelques points essentiels sur	20	18	17
de certains éléments très importants	13	19	19
placer quelques accents importants	5	20	20

Figure 5: Examples of rephrasings for the phrase *quelques points importants* in *je voudrais mentionner* **quelques points importants** *de la directive*

Rephrasings	Ranks given by model(s)		
	PIV	LM	PIV+LM+DEP+LEM
vous avez raison	1	1	1
je suis d' accord avec vous	2	2	2
je suis d' accord	3	6	3
je conviens avec vous	6	5	4
je partage votre avis	7	4	5
vous avez raison de dire	10	3	5
je pense comme vous	7	8	7
je suis parfaitement d' accord avec vous	12	7	8
je partage votre point de vue	12	9	9
je vous rejoins	7	10	10
, je vous donne raison	3	12	11
là , je vous donne raison	3	13	12
tu as raison	16	11	12
vous avez raison de	10	14	14
je partage votre point	12	15	15
je partage votre point de	12	16	16

Figure 6: Examples of rephrasings for the phrase *je vous donne raison* in *à cet égard bien précis ,* **je vous donne raison** *, monsieur le commissaire*

ing of *take his place*, due to the possessive determiner. The actual meaning of words depends on the context in which they are used. The work done by the team of Gross on lexicon-grammar (e.g. (Gross, 1984)) showed that a relatively small set of clause patterns and syntactic constraints suffices to cover most of common French.

Comparable monolingual corpora have been used for automatic paraphrasing. Barzilay and Lee (Barzilay and Lee, 2003) learned paraphrasing patterns as pairs of word lattices, which are then used to produce sentence level paraphrases. Their corpus contained news agency articles on the same events, which allows precise sentence paraphrasing, but on a small sets of phenomena and for a limited domain. As sentential paraphrasing is more likely to alter meaning, Quirk *et al.* (Quirk et al., 2004) approached paraphrasing as a monotonous decoding by a phrase-based SMT system. Their corpus consisted of monolingual sentences extracted from a comparable corpus that were automatically aligned so as to allow aligned phrase extraction. Pang *et al.* (Pang et al., 2003) used parallel monolingual corpora built from news stories that had been independantly translated several times to learn lattices from a syntax-based alignment process.

Bannard and Callison-Burch (Bannard and Callison-Burch, 2005) proposed to use pivot translation for paraphrasing phrases. Fujita (Fujita, 2005) proposed a transfer-and-revision framework using linguistic knowledge for generating paraphrases in Japanese and a model for error detection. At the lexical level, a recent evaluation on English lexical substitution was held (McCarthy and Navigli, 2007) in which systems had to find lexical synonyms and disambiguate the context.

6   Discussion and future work

In this article, we have presented an approach for obtaining rephrasings for short text spans from parallel bilingual corpora. These rephrasings can be ranked according to user-defined preferences, and the weights of the models used can be dynamically adjusted by a user depending on what features are more important to her, for instance after an initial list of candidates has been proposed by the system. Indeed, good candidates include paraphrases, but also more generally phrases that could help a writer revise a text with some shifts in meaning, even if at the cost of some corrections to make the resulting text grammatical. Furthermore, search for rephrasings can be iteratively performed using candidate rephrasings as source phrases, and the user can have some fine-grained control if selecting or rejecting possible pivot phrases manually. Possible user interfaces to this proposed bilingual phrase lexicon could include rephrasing memory features to learn from interaction with the user, and concordancing features to display the context of use in the bilingual corpus of the segments used to build the relevant lexicon entries. In the latter case, the similarity used to select examples could take the context of the phrases into account in terms of dependency relationships.

There are several open issues to the presented work. Important issues are where the phrases can come from and the bias introduced by the resource used. Using a bilingual corpora such as the Europarl corpus with this pivot approach yields both generic and domain/genre-specific rephrasings, and it is important to be able to determine their appropriate context of use. It would also be interesting to investigate enriching this framework with phrases learnt from monolingual corpora from a given domain or genre, and to use features from the current text under revision. More generally, we would need to get some idea of the degree of possible reuse of a given rephrasing.

Another important group of issues concerns limitations due to the nature of phrases for the task at hand. As we have said, phrases as units of rephrasing are limited because they cannot model non-consecutive words and because of the rigidity of their content. Various types of entry points to the rephrasing lexicon such as using word-based regular expressions can in some way alleviate this problem, but work could be done on the lexicon itself. As shown by Callison-Burch (Callison-Burch, 2007), much can be gained by using better alignments. Alignments techniques using syntactic information could eliminate weak rephrasing candidates (i.e. increase in overall precision), but interesting phrasal alignments could be lost as well (decrease in overall recall). Furthermore, information from the context of alignments could also be used to disambiguate the source phrase and get only pivot phrases that are compatible with the context of a given rephrasing, in similar ways as recently done for SMT (Stroppa et al., 2007).

References

Bannard, Colin and Chris Callison-Burch. 2005. Paraphrasing with bilingual parallel corpora. In *Proceedings of ACL*, Ann Arbor, USA.

Barzilay, Regina and Lillian Lee. 2003. Learning to paraphrase: an unsupervised approach using multiple-sequence alignment. In *Proceedings of NAACL/HLT*, Edmonton, Canada.

Bourdaillet, Julien, Jean-Gabriel Ganascia, and Irène Fenoglio. 2007. Machine assisted study of writers' rewriting processes. In *Proceedings of NLPCS, poster session*, Madeire, Portugal.

Bourigault, Didier, Cécile Fabre, Cécile Frrot, Marie-Paule Jacques, and Sylvia Ozdowska. 2005. Syntex, analyseur syntaxique de corpus. In *Proceedings of TALN*, Dourdan, France.

Callison-Burch, Chris. 2007. *Paraphrasing and Translation*. Ph.D. thesis, University of Edinburgh.

Fellbaum, Christiane, editor, 1998. *WordNet: An Electronic Lexical Database and some of its Applications*. MIT Press.

Ferret, Olivier and Michael Zock. 2006. Enhancing electronic dictionaries with an index based on associations. In *Proceedings of COLING/ACL*, Sydney, Australia.

Fujita, Atsushi. 2005. *Automatic Generation of Syntactically Well-formed and Semantically Appropriate Paraphrases*. Ph.D. thesis, Nara Institute of Science and Technology.

Gross, Maurice. 1984. Lexicon-grammar and the analysis of french. In *Proc. of the 11th COLING*, pages 275–282, Stanford, CA.

Hanks, Patrick and James Pustejovsky. 2005. A pattern dictionary for natural language processing. *Revue Française de linguistique appliquée*, 10(2):63–82.

Koehn, Philipp, Franz Josef Och, , and Daniel Marcu. 2003. Statistical phrase-based translation. In *Proceedings of NAACL/HLT*, Edmonton, Canada.

Koehn, Philipp. 2005. Europarl: A parallel corpus for statistical machine translation. In *Proceedings of MT Summit*, Phuket, Thailand.

Max, Aurélien. 2008. Local rephrasing suggestions for supporting the work of writers. In *Proceedings of GoTAL*, Gothenburg, Sweden.

McCarthy, Diana and Roberto Navigli. 2007. Semeval-2007 task 10: English lexical substitution task. In *Proceedings of the Semeval-2007 Workshop at ACL*, Prague, Czech Republic.

Munteanu, Dragos S. and Daniel Marcu. 2006. Extracting parallel sub-sentential fragments from non-parallel corpora. In *Proceedings of COLING/ACL 2006*, Sydney, Australia.

Mutton, Andrew, Mark Dras, Stephen Wan, and Robert Dale. 2007. GLEU : Automatic evaluation of sentence-level fluency. In *Proceedings of ACL*, Prague, Czech Republic.

Och, Franz Josef and Hermann Ney. 2003. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29(1):19–51.

Pang, Bo, Kevin Knight, and Daniel Marcu. 2003. Syntax-based alignment of multiple translations: Extracting paraphrases and generating new sentences. In *Proceedings of NAACL/HLT*, Edmonton, Canada.

Quirk, Chris, Chris Brockett, and William B. Dolan. 2004. Monolingual machine translation for paraphrase generation. In *Proceedings of EMNLP*, Barcelona, Spain.

Stroppa, Nicolas, Antal van den Bosch, and Andy Way. 2007. Exploiting source similarity for smt using context-informed features. In *Proceedings of TMI*, Skvde, Sweden.

Zock, Michael and Slaven Bilac. 2004. Word lookup on the basis of associations : from an idea to a roadmap. In *Workshop on 'Enhancing and using electronic dictionaries'*, pages 29–35, Geneva. COLING.

Zock, Michael. 2006. Navigational aids, a critical factor for the success of electronic dictionaries. In Rapp, Reinhard, P. Sedlmeier, and G. Zunker-Rapp, editors, *Perspectives on Cognition: A Festschrift for Manfred Wettler*, pages 397–414. Pabst Science Publishers, Lengerich.

Zock, Michael. 2007. If you care to find what you are looking for, make an index: the case of lexical access. *ECTI, Transaction on Computer and Information Technology*, 2(2):71–80.