Semi-compositional Method for Synonym Extraction of Multi-Word Terms

Amir Hazem, Béatrice Daille

Semi-compositional Method for Synonym Extraction of Multi-Word Terms

Amir HAZEM and Béatrice DAILLE

Laboratoire d'Informatique de Nantes-Atlantique (LINA)
Université de Nantes, 44322 Nantes Cedex 3, France
Amir.Hazem@univ-nantes.fr,Beatrice.Daille@univ-nantes.fr

Abstract

Automatic synonyms and semantically related word extraction is a challenging task, useful in many NLP applications such as question answering, search query expansion, text summarization, etc. While different studies addressed the task of word synonym extraction, only a few investigations tackled the problem of acquiring synonyms of multi-word terms (MWT) from specialized corpora. To extract pairs of synonyms of multi-word terms, we propose in this paper an unsupervised semi-compositional method that makes use of distributional semantics and exploit the compositional property shared by most MWT. We show that our method outperforms significantly the state-of-the-art.

Keywords: Synonyms, Multi-word terms, Compositionality, Distributional semantics, Unsupervised methods

1. Introduction

Identifying synonyms and more generally semantically related words is currently based on the following paradigms: distributional, lexicon-based and multilingual approaches.
The former that is the most popular relies on the distributional hypothesis that puts forward the idea that words with similar meaning tend to occur in similar contexts (Harris, 1954). Hence, semantically related words are ranked according to the similarity of their contexts. There have been many studies based on distributional paradigm (Hindle, 1990; Grefenstette, 1994; Lin, 1998; Hagiwara, 2008; Ferret, 2010). Lin (1998) for instance, introduced the idea that words sharing more syntactic relations are more likely to be semantically related. This idea has been extended to the syntactic path level by (Hagiwara, 2008) in order to account for less direct syntactic dependencies. The lexicon-based approach exploits the word definitions of general-language dictionaries or terminology banks. Synonyms are extracted according to the similarity between vertices in graphs (Blondel and Senellart, 2002). The last approach to address synonym extraction adopts a multilingual scenario under the assumption that words with similar translational contexts tend to be semantically related (Wu and Zhou, 2003; van der Plas and Tiedemann, 2006; Andrade et al., 2013).
Beyond all these paradigms, it is not obvious to make a clear distinction between synonyms and other semantically related words (Lin et al., 2003; van der Plas and Tiedemann, 2006). Concerning the distributional method, Resnik (1993:18) states that «It would seem that the information captured using distributional methods is not precisely syntactic, nor purely semantic - in some sense the only word that appears is *distributional*. ». If we want to characterize this semantic proximity, we will encounter classical lexical semantic relationships such as synonymy, antonymy, hyperonymy, co-hyponymy and meronymy, and non-classical lexical semantic relationships such as action/agent (Morris and Hirst 2004).
All these approaches were applied to the detection of single word synonyms and multi-word unit synonyms were mostly ignored. Dealing with specialized languages and romance languages, multi-word terms that are composed of multi-word strings are by far the more frequent terms. It is thus important to identify synonymic variants of multi-word terms.
This paper reports on an ongoing investigation dealing with the synonym extraction of multi-word terms (MWT) in specialized monolingual corpora. We adopt the principle of compositionality to generate MWT synonyms, but unlike (Hamon and Nazarenko, 2001), we do not use a dictionary to propose synonyms of the parts of the MWT but we exploit the semantically-related words that are provided by a distributional method, adding a few constraints. We will demonstrate that this semi-compositional method is able to find in a monolingual corpora MWT synonyms whereas a dictionary-based method will not.
The remainder of this paper is organized as follows. Section 2. describes the principle of compositionality and synonymy. Section 3. presents our approach. Section 4. describes the different linguistic resources used in our experiments. Section 5. evaluates the contribution of all the approaches on the quality of MWT synonym extraction through different experiments. Sections 6. and 7. present our discussion and conclusion.

2. Compositionality and synonymy

A general admitted definition of composition proposed by Partnee et al. 1990 is that *"a compound expression is composional if its meaning is a function of the meaning of the parts and of the syntactic rule by which they are combined*".
Hammon and Nazarenko (2001) assumed that synonymy of MWT is compositional if their parts are synonyms. They defined three rules to detect synonymy relations. Given the complex candidate terms $CCT_1 = (T_1, E_1)$ and $CCT_2 = (T_2, E_2)$ and $syn(CT_1, CT_2)$ a synonym relation between the candidate terms CT_1 and CT_2, the following inferences rules are used:

- R_1: $T_1 = T_2 \wedge syn(E_1, E_2) \supset syn(CCT_1, CCT_2)$

- R_2: $E_1 = E_2 \wedge syn(T_1, T_2) \supset syn(CCT_1, CCT_2)$
- R_3: $syn(T_1, T_2) \wedge syn(E_1, E_2) \supset syn(CCT_1, CCT_2)$

Rule R_1 means that the heads are identical and the expansions are synonymous (*collecteur général* (*general collector*)/*collecteur commun* (*common collector*)). Synonym extraction of single-word terms such as *général* in this example is carried out using a dictionary of synonyms.

Kracht (2007) notes that expressions and their parts are usually ambiguous and that a meaning can only be assigned to their analyses. We illustrate this remark with the analysis of multi-word term synonyms that are recorded in terminology banks. Examining the parts of synonymic MWTs of the wind-energy domain sharing at least one common content part, we face miscellaneous relationships:

- Synoyms: *energy output*/*energy production* given by Termium[1] where *output*/*production* are synonyms;
- Hyperonyms: *turbine noise*/*turbine sound* given by Grand dictiontionnaire terminologique[2] where *sound* is an hyperonym of *noise* or *implantación de las máquinas* /*implantación de aerogeneradores* given by the Lexique panlatin where *máquina* 'machine' is an hyperonym of *aerogenerador*;
- Undefined: *nuclear plant*/*nuclear station* given by Termium with no relation between *plant* and *station* or *arbre lent*/*arbre primaire* 'low speed shaft' given by Terminalf with no relation between *lent* 'slow' and *primaire* 'primary'.

Our hypothesis is that distributional semantics that is provided by semantic-related words could be of some help to discover MWT synomyms. We need for this task to adapt the compositional method proposed by (Hamon and Nazarenko, 2001).

3. Semi-Compositional Method

Our method is inspired by the work of (Morin and Daille, 2012). The authors improve the alignment of equivalent terms from comparable corpora by using compositional method and context-based projection. We apply the same idea to the task of MWT synonym extraction. We start from the assumption that MWT and their synonyms follow the principle of compositionality. For example, the synonym of *énergie renouvelable (renewable energy)* can be obtained by first extracting each part of the MWT; then, finding the semantically related words of *énergie (energy)* and/or *renouvelable (renewable)* with distributional methods; finally, filtering all expressions using monolingual specialized corpora. Our semi-compositional method differs from the method of (Hamon and Nazarenko, 2001) in two points: (i) the way to extract synonyms of single-word terms and (ii) the length of MWTs.

[1] http://www.btb.termiumplus.gc.ca
[2] http://www.oqlf.gouv.qc.ca/

3.1. Distributional Method

Instead of using a dictionary that will provide synonyms of each lexical element of the MWT, our method exploits distributional relationships. We follow the assumption that two words are more likely to be semantically related if they share the same lexical contexts. For a source word w_i^s, we first build its context vector $v_{w_i^s}$. The vector $v_{w_i^s}$ contains all the words that co-occur with w_i^s within window of n words. Let us denote by $occ(w_i^s, w_j^s)$ the co-occurrence value of w_i^s and a given word of its context w_j^s. The process of building context vector is repeated for all words of the monolingual specialized corpus. An association measure such as the point-wise mutual information (noted MI) (Fano, 1961), the log-likelihood (noted LL) (Dunning, 1993) or the discounted odds-ratio (noted DOR) (Laroche and Langlais, 2010) is used to score the strength of correlation between a word and all the words of its context vector. Finally, a similarity measure is used to score each target word w_i^t with respect to the target context vector, $v_{w_i^s}^t$. Various vector similarity measures can be used, for instance the cosine similarity (Salton and Lesk, 1968) (noted COS) or the weighted Jaccard index (noted JAC) (Grefenstette, 1994). The candidate synonyms of the word w_i^s are the target words ranked following their similarity score.

Hereafter the used association and similarity measures:

	j	$\neg j$
i	$a = occ(i, j)$	$b = occ(i, \neg j)$
$\neg i$	$c = occ(\neg i, j)$	$d = occ(\neg i, \neg j)$

Table 1: Contingency table

$$\begin{aligned} LL(i,j) = \quad & a\log(a) + b\log(b) + c\log(c) + d\log(d) \\ & +(N)\log(N) - (a+b)\log(a+b) \\ & -(a+c)\log(a+c) - (b+d)\log(b+d) \\ & -(c+d)\log(c+d) \end{aligned} \tag{1}$$

with $N = a + b + c + d$.

$$\text{MI}(i,j) = log\frac{a}{(a+b)(a+c)} \tag{2}$$

$$DOR(i,j) = \log\frac{(a+\frac{1}{2}) \times (d+\frac{1}{2})}{(b+\frac{1}{2}) \times (c+\frac{1}{2})} \tag{3}$$

$$Cosinus_{v_l}^{v_k} = \frac{\sum_t assoc_t^l \, assoc_t^k}{\sqrt{\sum_t {assoc_t^l}^2}\sqrt{\sum_t {assoc_t^k}^2}} \tag{4}$$

$$Jaccard_{v_l}^{v_k} = \frac{\sum_t min(assoc_t^l, assoc_t^k)}{\sum_t max(assoc_t^l, assoc_t^k)} \tag{5}$$

with $assoc_t^l$ for instance that refers to a given association measure (LL, MI or DOR) between the two words t and l.

3.2. Length of MWTs

In contrast to (Hamon and Nazarenko, 2001) method, our semi-compositional method is not limited to MWT composed of two elements. It can be applied to MWT of any length as long as they follow the rules R_1 and R_2 (Hamon and Nazarenko, 2001). We extend the rules from (Hamon and Nazarenko, 2001) to the semantic level and generalize them to MWT of any length. We extend the compositional rules R_1 and R_2 by replacing $syn(CCT_1, CCT_2)$ which means synonym relation between CCT_1 and CCT_2 by $sem(CCT_1, CCT_2)$, which means semantic relation between CCT_1 and CCT_2. R_1^G corresponds to the generalized rule R_1 (respectively, R_2^G corresponds to the generalized rule R_2) and T_1, T_2, E_1, E_2 can be MWTs. In addition, we remove the rule R_3 relying on the results of (Hamon and Nazarenko, 2001) where they have shown that is the less productive and reliable rule. We obtain the following rules:

- $R_1^G : T_1 = T_2 \wedge sem(E_1, E_2) \supset sem(CCT_1, CCT_2)$
- $R_2^G : E_1 = E_2 \wedge sem(T_1, T_2) \supset sem(CCT_1, CCT_2)$

Table 2 illustrates some MWT synonym examples in French, English and Spanish from the *wind energy* corpus that are attested MWT synonyms in wind energy terminology banks. Most of the examples of Table 2 follow the rule that one of the parts of the MWT and its synonym remain unchanged (wind turbine/wind machine, power supply/energy supply, énergie renouvelable/énergie durable, etc.). For each MWT, we alternatively fix the left part and extract the semantically related words of the right part (and vice-versa, fix the right part and extract the semantically related words of the left part). This corresponds to rules R_1^G and R_2^G. We filter the obtained MWTs according to the n-grams extracted from the specialized corpora. One drawback of this method is the impossibility to treat synonyms that do not follow the above cited rules, for instance, the MWT *moulin à vent* 'windmill' and its synonym *éolienne* 'wind turbine'. We do not address this particular case in this paper.

English term synonyms	
wind turbine	wind machine
power supply	energy supply
power plant	electricity plant
savonius model	savonius type
energy output	energy production
sea wind farm	offshore wind farm
wind farm	wind power plant
wind turbine	aeroturbine
French term synonyms	
énergie renouvelable	énergie durable
centrale électrique	centrale éolienne
unité de stockage	dispositif de stoskage
arbre primaire	arbre lent
force du vent	vitesse du vent
éolienne	moulin à vent
Spanish term synonyms	
ángulo de paso	ángulo de calaje
extremo de la pala	punta de la pala
mapa de vientos	mapa eólico
coeficiente de potencia	coeficiente de rendimiento
implantación de las máquinas	implantación de aerogeneradores
aerogenerador	torre eólica

Table 2: Examples of English/French/Spanish synonyms of multi-word terms recorded in terminology banks of the wind energy domain

4. Experimental Setup

In this section, we will describe the data and the different parameters used in our experiments.

4.1. Corpora and reference lists

The experiments have been carried out on the French/English/Spanish specialized corpus from the domain of wind energy of 400,000 words[3]. The corpus has been normalized through the following linguistic preprocessing steps: tokenization, part-of-speech tagging, and lemmatization. To build our reference lists, we selected the French MWT pairs from the Terminalf[4] linguistic resource. From 84 MWTs of the wind energy domain, we obtained 34 French MWT synonyms as a result of filtering out SWT synonyms and after checking that the MWT synonyms occur in the specialized corpora. For English, we selected the MWT pairs from the glossary of wind energy from the online book (Gipe, 2004) and from the linguistic resource Termium [5]. As a result of filtering and of corpus projection, we obtained 20 English MWT pairs. For Spanish, we found 64 MWT listed in the Lexique panlatin de l'énergie éolienne[6]. Applying the same filtering process than French and English, we obtained 26 Spanish MWT pairs. The small size of the reference lists can be explained by the fact that the term should follow the principle of monosemy and mononymy recalled by Bowker and Hawkins (2006:83): a term should be applied to a single concept, and a concept should be designed by only one term. So, synonyms of terms are rare phenomena. Another reason is that small specialized corpora contain a limited set of specialized terms.

4.2. Dictionary-based method

We used as baseline the method proposed by (Hamon and Nazarenko, 2001). To extract French synonyms of single-

[3]Wind energy corpora can be downloaded from the website http://www.lina.univ-nantes.fr/?Ressources-linguistiques-du-projet.html

[4]http://www.terminalf.scicog.fr

[5]http://www.btb.termiumplus.gc.ca/tpv2alpha/alpha-eng.html?lang=eng

[6]http://www.bt-tb.tpsgc-pwgsc.gc.ca/btb.php?lang=fra&cont=s_219

word terms we used the on-line dictionary DES [7]. DES Contains 49,168 entries and 201,511 synonym relations. The initial database has been constructed from seven dictionaries. The extraction of English synonyms has been conducted using the lexical database WordNet [8]. WordNet contains aproximatly 117,000 synsets. The main relation among words in WordNet is synonymy.

4.3. Distributional Method Settings

Using the *distributional method*, three major parameters need to be set:

1. The size of the window used to build the context vectors (Morin et al., 2007; Gamallo, 2008)
2. The association measure (the log-likelihood (Dunning, 1993), the point-wise mutual information (Fano, 1961), the discounted odds-ratio (Laroche and Langlais, 2010),...)
3. The similarity measure (the weighted Jaccard index (Grefenstette, 1994), the cosine similarity (Salton and Lesk, 1968),...)

To build the context vectors we chose a 7-window size. We used MI, LL and DOR as association measures and COS and JAC as similarity measures.
We use the mean average precision *MAP* (Manning et al., 2008) to evaluate the quality of the system.

$$MAP = \frac{1}{|W|} \sum_{i=1}^{|W|} \frac{1}{Rank_i} \quad (6)$$

where $|W|$ corresponds to the size of the evaluation list, and $Rank_i$ corresponds to the ranking of a correct synonym candidate i.

5. Results

In this section we present the experimental results conducted on the French/English/Spanish wind energy corpus for the extraction of MWTs synonyms. We refer to the method of (Hamon and Nazarenko, 2001) as "baseline" and refer to our method as "Semi-Comp".

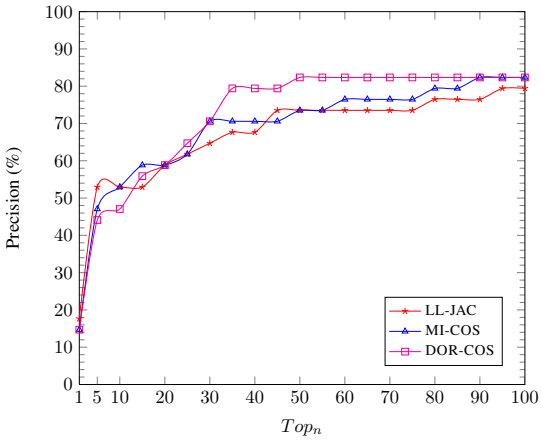
Method	French	English	Spanish
Baseline	0.25	3.63	8.09
Semi-Comp (MI-COS)	27.4	32.6	**24.7**
Semi-Comp (DOR-COS)	26.8	27.2	14.2
Semi-Comp (LL-JAC)	**31.4**	**36.1**	18.8

Table 3: Results on the wind energy corpus (MAP%)

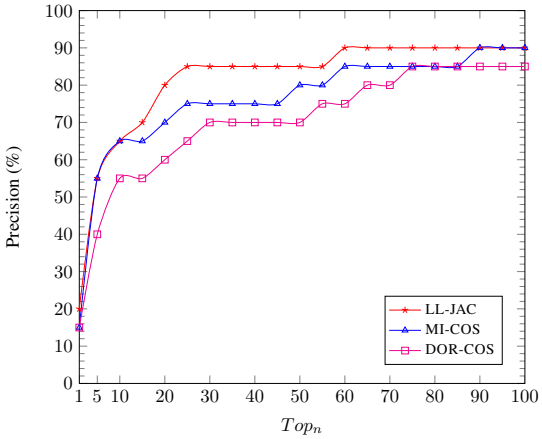
We can see from Table 3 that the Semi-Compositional approach outperforms significantly the baseline for all languages. The Semi-Compositional approach obtains 31.4% of MAP for the French corpus (LL-JAC configuration), 36.1% of MAP for the English corpus (LL-JAC configuration) and 24.7% of MAP for the Spanish corpus (MI-COS configuration) while the baseline could only obtain 0.25%, 3.63% and 8.09% of MAP for the three corpora.

[7] http://www.crisco.unicaen.fr/des/synonyms

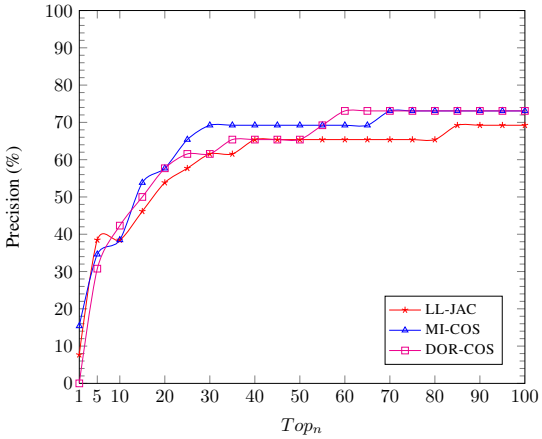
[8] http://wordnetweb.princeton.edu/perl/webwn/



(a) French corpus



(b) English corpus



(c) Spanish corpus

Figure 1: Comparison (Top_n) of different configurations of the Semi-Compositional approach on the French, the English and the Spanish corpus

Figure 1 shows the behavior of the Semi-Compositional approach according to different configurations. For the French corpus, we can note that the configuration LL-JAC gives the best precision for top_1 and top_5. It is then significantly outperformed by the configuration DOR-COS (from top_{35}) and slightly outperformed by the configuration MI-COS (from top_{60}). The results of MI-COS in comparison with LL-JAC from top_{20} to top_{55} are more

contrastive.

Concerning the English corpus, the results are quite different comparing to the results of the French corpus. As we can see, the LL-JAC configuration always outperforms the two other configurations except from top_5 to top_{10} and for the last tops where the MI-COS configuration shows the same precision. We can also see that the DOR-COS configuration is always outperformed except for top_1 and from top_{75} to top_{85} where it shows the same results as MI-COS. The results are more contrastive for the Spanish corpus. We can nonetheless observe that the MI- COS configuration performs the best in general.

It is not straightforward to choose the best configuration according to the results of Figure 1. That said, we can suppose that a tuning process or a combination of the different configurations should be a suitable alternative.

There are terms of which the synonym appears at the first position of the synonym list such as for English: *lattice construction/lattice tower*, *drive train/power train*, *nominal speed nominal power*, for French: *éolienne offshore/éolienne en mer*, *générateur synchrone/machine synchrone machine à induction générateur à induction*, *site éolien/parc éolien*, for Spanish: *torre eólica/turbina eólica*, *coeficiente de rendimiento/ coeficiente potencia implantación máquina/implantación aerogeneradores*. Some of them such as *lattice construction/lattice tower* are collocations in the sense that *lattice* accepts only a right context either *construction* or *tower*. When the synonym is closed to the first position, we notice that the candidate synonyms that appears before are either flexional or derivational variants such as for *generador asíncrono*, the morphological variants *generador sincrónicos* and *generador sincrónicos* appear just before the recorded synonym *generador a inducción*, or other synomyms such as for *frein à disque*, the recorded synonym *frein mécanique* appears at Top_3 whether *frein aérodynamique* which appears before is a valid synonym too. Of course, the distributional analysis fails for a few MWTs when the recorded synonym appears too far in the list or when it is simply not found. If we decide to filter out the MWT synonym candidates that are not sharing the same syntactic structure than the term, we can obtain the recorded synonym at the first position: for the term *viento aparente*, the recorded synonym *viento relativo* of N A structure appears at Top_{14} and all the candidate MWT synomyns that appear before share other grammatical structures *viento fuerza/ N N*, *viento denominar/N V*.

6. Discussion

Few approaches to extract synonyms of MWTs have been proposed so far. To our knowledge, only (Hamon and Nazarenko, 2001) addressed this task following the principle of compositionality. The weak results obtained by the baseline can be explained by two facts: (i) in specialized corpora, dictionaries of synonyms are often not available; (ii) relying on lexicons from the general domain can lead to irrelevant or off-topic MWTs as shown in our experiments. The main contribution of our approach is the use of distributional analysis instead of dictionaries. Distributional analysis allows us to identify semantically related words that will be used to discover MWT synonyms using the compositionality principle. We also highlight the fact that MWT synomyms are not always composed of the synonyms of its elements, and that distributionally related words is suitable for this task. Finally, our method is not limited to MWTs composed of two elements. It can be applied to any MWT that follows the rules R_1^G and R_2^G. If we have demonstrated that MWT synonyms could be discovered by distributional analysis, we still need to improve the precision in order to obtain valid synonym candidates at the first positions. Several filters and ranking methods should be proposed to improve the ranking and to filter out wrong synonym candidates.

7. Conclusion

We presented in this paper a semi-compositional method for synonym extraction of multi-word terms. Based on the principal of compositionality and distributional semantics, our method has shown significant improvements while compared to the state-of-the-art approach for the discovering of MWT synomyms. If more investigation is certainly needed for particular cases, the encouraging results lend support the idea that combining compositionality with distributional semantics is a relevant way to handle MWT synonym extraction. For future work, we will explore synonym extraction of MWT that are not compositional (SWT) and conversely, explore synonyms of SWT that are MWTs.

Acknowledgments

The research leading to these results has received funding from the French National Research Agency under grant ANR-12-CORD-0020 (CRISTAL project).

8. References

Daniel Andrade, Masaaki Tsuchida, Takashi Onishi, and Kai Ishikawa. 2013. Synonym acquisition using bilingual comparable corpora. In *International Joint Conference on Natural Language Processing (IJCNLP'13)*, Nagoya, Japan.

Vincent D. Blondel and Pierre Senellart. 2002. Automatic extraction of synonyms in a dictionary.

Lynne Bowker and Shane Hawkins. 2006. Variation in the organization of medical terms - exploring some motivations of term choice. *Terminology*, 12(1):79–110.

Ted Dunning. 1993. Accurate Methods for the Statistics of Surprise and Coincidence. *Computational Linguistics*, 19(1):61–74.

Robert M. Fano. 1961. *Transmission of Information: A Statistical Theory of Communications*. MIT Press, Cambridge, MA, USA.

Olivier Ferret. 2010. Testing semantic similarity measures for extracting synonyms from a corpus. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *LREC*.

Otero Gamallo. 2008. Evaluating two different methods for the task of extracting bilingual lexicons from comparable corpora. In *Proceedings of LREC 2008 Workshop on Comparable Corpora (LREC'08)*, pages 19–26, Marrakech, Marroco.

P. Gipe. 2004. *Wind power: renewable energy for home, farm, and business*. Chelsea Green Pub. Co.

Gregory Grefenstette. 1994. *Explorations in Automatic Thesaurus Discovery*. Kluwer Academic Publisher, Boston, MA, USA.

Masato Hagiwara. 2008. A supervised learning approach to automatic synonym identification based on distributional features. In *Proceedings of the ACL-08: HLT Student Research Workshop*, pages 1–6, Columbus, Ohio, June. Association for Computational Linguistics.

Thierry Hamon and Adeline Nazarenko. 2001. Detection of synonymy links between terms: experiment and results. In *Recent Advances in Computational Terminology*, pages 185–208. John Benjamins.

Zellig S Harris. 1954. Distributional structure. *Word*.

Donald Hindle. 1990. Noun classification from predicate-argument structures. In *ACL*, pages 268–275.

Marcus Kraft. 2007. Compositionality: The very idea. *Research on Language and Computation*, 5(3):287–308.

Audrey Laroche and Philippe Langlais. 2010. Revisiting Context-based Projection Methods for Term-Translation Spotting in Comparable Corpora. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING'10)*, pages 617–625, Beijing, China.

Dekang Lin, Shaojun Zhao, Lijuan Qin, and Ming Zhou. 2003. Identifying synonyms among distributionally similar words. In *IJCAI-03, Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence*, Acapulco, Mexico.

Dekang Lin. 1998. Automatic retrieval and clustering of similar words. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 2*, ACL '98, pages 768–774, Stroudsburg, PA, USA. Association for Computational Linguistics.

D Christopher Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. Introduction to information retrieval. Cambridge University Press.

Emmanuel Morin and Béatrice Daille. 2012. Revising the compositional method for terminology acquisition from comparable corpora. In *24th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, Coling'12*, pages 1797–1810, Mumbai, India.

Emmanuel Morin, Béatrice Daille, Koichi Takeuchi, and Kyo Kageura. 2007. Bilingual Terminology Mining – Using Brain, not brawn comparable corpora. In *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL'07)*, pages 664–671, Prague, Czech Republic.

J. Morris and G. Hirst. 2004. Non-classical lexical semantic relations. In *HLT-NAACL Worlshop on Computational Lexical semantics (CLS'04)*, pages 46–51. ACL.

Barbara H. Partnee, Alice ter Meulen, and Robert E. Wall. *Mathematical Methods in Linguistics*, volume 30 of *Studies in Linguistics and Philosophy*. Dordrecht, kluwer edition.

P. Resnik. 1993. *Selection and Information: A class-based approach to lexical relationships*. Ph.D. thesis, University of Pennsylvania.

Gerard Salton and Michael E. Lesk. 1968. Computer evaluation of indexing and text processing. *Journal of the Association for Computational Machinery*, 15(1):8–36.

Lonneke van der Plas and Jörg Tiedemann. 2006. Finding synonyms using automatic word alignment and measures of distributional similarity. In *21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics ACL'06*, Sydney, Australia.

Hua Wu and Ming Zhou. 2003. Optimizing synonym extraction using monolingual and bilingual resources. In *In Proceedings of the second international workshop on Paraphrasing*, page 72.