Proceedings of the Globalex Workshop on Linked Lexicography @LREC2022, pages 19–25
Marseille, 20 June 2022

TIAD 2022
The Fifth Translation Inference Across Dictionaries Shared Task

Jorge Gracia[1], Besim Kabashi[2], Ilan Kernerman[3]
[1] Aragon Institute of Engineering Research (I3A), University of Zaragoza, Spain, jogracia@unizar.es
[2] Corpus and Computational Linguistics, University of Erlangen-Nuremberg, Germany, besim.kabashi@fau.de
[3] K Dictionaries - Lexicala, Tel Aviv, Israel, ilan@kdictionaries.com

Abstract
The objective of the Translation Inference Across Dictionaries (TIAD) series of shared tasks is to explore and compare methods and techniques that infer translations indirectly between language pairs, based on other bilingual/multilingual lexicographic resources. In this fifth edition, the participating systems were asked to generate new translations automatically among three languages - English, French, Portuguese - based on known indirect translations contained in the Apertium RDF graph. Such evaluation pairs have been the same during the four last TIAD editions. Since the fourth edition, however, a larger graph is used as a basis to produce the translations, namely Apertium RDF v2. The evaluation of the results was carried out by the organisers against manually compiled language pairs of K Dictionaries. For the second time in the TIAD series, some systems beat the proposed baselines. This paper gives an overall description of the shard task, the evaluation data and methodology, and the systems' results

Keywords: TIAD, translation inference, lexicographic data, dictionary, Apertium RDF

1. Introduction

A number of methods and techniques have been explored in the past with the aim of automatically generating new bilingual and multilingual dictionaries based on existing ones. For instance, given a bilingual dictionary containing translations from one language L1 to another language L2, and another dictionary with translations from L2 to L3, a new set of translations from L1 to L3 is produced. The intermediate language (L2 in this example) is called pivot language, and it is possible to use multiple pivots for this purpose. When using intermediate languages, it is necessary to discriminate wrong inferred translations caused by translation ambiguities. The method proposed by Tanaka and Umemura (Tanaka and Umemura, 1994) in 1994, called One Time Inverse Consultation (OTIC), identified incorrect translations when constructing bilingual dictionaries intermediated by a third language. This was a pioneering work and it still constitutes a baseline that is hard to beat, as the previous TIAD editions demonstrated. The OTIC method has been further adapted and evolved in the literature, for instance by Lim et al. (Lim et al., 2011), who grounded on it for their method for multilingual lexicon creation. From a different perspective, other works were proposed that relied on cycles and graph exploration to validate indirectly inferred translations, such as the SenseUniformPaths algorithm by Mousam et al. (Mausam et al., 2009), the CQC algorithm by Flati et al. (Flati and Navigli, 2013) or the exploration based on cycle density by Villegas et al. (Villegas et al., 2016).

However, previous work on the topic of automatic bilingual/multilingual dictionary generation was usually conducted on different types of datasets and evaluated in different ways, applying various algorithms that are often not comparable. In this context, the objective of the Translation Inference Across Dictionaries (TIAD) shared task is to support a coherent experiment framework that enables reliable validation of results and solid comparison of the processes used. In addition, this initiative aims to enhance further research on the topic of inferring translations across languages.

The TIAD first edition[1] took place in Galway (Ireland) in 2017, co-located with the LDK'17 conference. The second edition[2] in 2019 was co-located with LDK'19 in Leipzig (Germany), and the third one was planned at LREC'20 in Marseille (France) as part of the Globalex Workshop on Linked Lexicography[3]. Although the workshop of the third edition did not take place because of the COVID-19 crisis, the evaluation was run and the results published[4]. Participants in the 3rd edition had the opportunity to present their systems jointly with the contributors to the 4th TIAD edition[5], during the whorkshop that took place in Zaragoza (Spain) at LDK'21. The fifth edition of TIAD was held in conjunction to the GLOBALEX 2022 – Linked Lexicography workshop[6] at the 13th Language Resources and Evaluation Conference (LREC 2022)[7] in Marseille (France) on June 20, 2022. In this paper, we give an overall description of the shard task, the evaluation data and methodology, and the system results of TIAD 2022.

[1] https://tiad2017.wordpress.com/
[2] https://tiad2019.unizar.es
[3] https://globalex2020.globalex.link/globalex-workshop-lrec2020-about-globalex-lrec2020/
[4] https://tiad2020.unizar.es
[5] https://tiad2021.unizar.es
[6] https://globalex2022.globalex.link/lrec2022/
[7] https://lrec2022.lrec-conf.org/en/

The remainder of this paper is organised as follows. In Section 2, an overall description of the shared task is given. Section 3 describes the evaluation data and Section 4 explains the evaluation process. In Section 5 the system results are reported, and conclusions are summarised in Section 6.

2. Shared task description

The objective of TIAD shared task is to explore and compare methods and techniques that infer translations indirectly between language pairs, based on other bilingual resources. Such techniques would help in auto-generating new bilingual and multilingual dictionaries based on existing ones.

In this fifth edition, the participating systems were asked to generate new translations automatically among three languages: English, French, and Portuguese, based on known translations contained in the Apertium RDF v2.0 graph[8]. As these languages (EN, FR, PT) are not directly connected in this graph, no translations can be obtained directly among them there. Based on the available RDF data, the participants had to apply their methodologies to derive translations, mediated by any other language in the graph, between the pairs EN/FR, FR/PT and PT/EN.

Participants could also make use of other freely available sources of background knowledge (e.g. lexical linked open data and parallel corpora) to improve performance, as long as no direct translation among the studied language pairs were available. Beyond performance, participants were encouraged to consider also the following issues in particular:

1. The role of the language family with respect to the newly generated pairs
2. The asymmetry of pairs, and how translation direction affects the results
3. The behavior of different parts of speech among different languages
4. The role that the number of pivots plays in the process

The evaluation of the results was carried out by the organisers against manually compiled pairs of K Dictionaries (KD), extracted from its Global Series[9], which were not accessible to the participants. A validation data set was made available to participants, upon request, in particular a 5% of randomly selected translations for each language pair. The goal of this validation data is to allow participants to analyse the nature of the data, to run some validation tests, and to analyse negative results.

3. Evaluation data

In this section we briefly describe the input data source that has been proposed in the shared task as a source of known translations, i.e., Apertium RDF, as well as the Global series data used as golden standard, from KD.

3.1. Source data

As mentioned above, the shared task relies on known translations contained in Apertium RDF, which were used to infer new ones. Apertium RDF is the linked data counterpart of the Apertium dictionary data. Apertium (Forcada et al., 2011) is a free open-source machine translation platform. The system was initially created by Universitat d'Alacant and is released under the terms of the GNU General Public License. In its core, Apertium relies on a set of bilingual dictionaries, developed by a community of contributors, which covers more than 40 languages pairs.

Apertium RDF (Gracia et al., 2018) is the result of publishing the Apertium bilingual dictionaries as linked data on the Web. The result groups the data of the (originally disparate) Apertium bilingual dictionaries in the same graph, interconnected through the common lexical entries of the monolingual lexicons that they share. An initial version of 22 language pairs was developed by Universidad Politécnica de Madrid and Universitat Pompeu Fabra[10]. A later conversion of the Apertium data into RDF, which we call Apertium RDF v2 in the following, was made by Goethe University Frankfurt and University of Zaragoza (Gracia et al., 2020). It contains 44 languages and 53 language pairs, with a total number of 1,540,996 translations between 1,750,917 lexical entries. In the second and third TIAD editions, the first version of Apertium RDF was used, while in the fourth and fifth editions we moved to the larger and richer Apertium RDF v2 graph.

In its first version, Apertium RDF was modeled using the *lemon* model (McCrae et al., 2012) jointly with its translation module (Montiel-Ponsoda et al., 2011), while Apertium RDF v2 uses the Ontolex lemon core model to represent the data (McCrae et al., 2017), jointly with the lemon vartrans module[11].

Each original Apertium bilingual dictionary was converted into three different objects in RDF: source lexicon, target lexicon, and translation set. As a result, two independent monolingual lexicons per dictionary were published as linked data on the Web, along with a set of translations that connects them. Note that the naming rule used to build the identifiers (URIs) of the lexical entries allows to reuse the same URI per lexical entry across all the dictionaries, thus explicitly connecting them. For instance the same URI is used for

[8] https://tiad2021.unizar.es/images/ApertiumRDFv2.0_graph.png

[9] https://www.lexicala.com/

[10] http://linguistic.linkeddata.es/apertium/

[11] https://www.w3.org/2016/05/ontolex/#variation-translation-vartrans

the English word *bench* as a noun[12]: throughout the Apertium RDF graph, no matter if it comes from, e.g., the EN-ES dictionary or the CA-EN one. More details about the generation of Apertium RDF based on the Apertium data can be found at (Gracia et al., 2018).

Figure 1: The Apertium RDF v2 graph. The nodes in the figure represent the monolingual lexicons and the edges are the translation sets between them. The darker the colour, the more connections a node has. We have highlighted the three languages of this evaluation campaign: PT, FR, and EN.

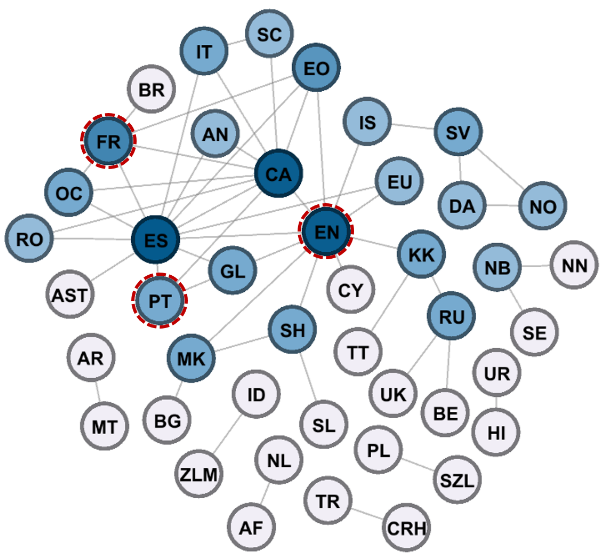

Figure 1 illustrates the Apertium RDF v2 unified graph. The nodes in the figure are the languages and the edges are the translation sets between them. All the datasets are available in Zenodo[13]. There is a plan to store the data in a permanent triplestore and expose it through a SPARQL endpoint in the near future, as part to the Prêt-à-LLOD project[14].

There were several ways in which the evaluation data was available to the participants: (i) through the data dumps available in Zenodo, which need to be loaded in a local triplestore, e.g., Apache Fuseki, and queried locally; (ii) through a testing SPARQL endpoint[15], and (iii) in a ZIP file in comma separated values (CSV) format[16], for those not acquainted with semantic web technologies. More details on how to access the data are available in the TIAD 2022 website[17].

3.2. Gold standard

The evaluation of the results was carried out by the organisers against manually compiled language pairs of K Dictionaries, extracted from its Global series, particularly the following pairs: BR-EN, EN-BR, FR-EN, EN-FR, FR-PT, PT-FR. The translation pairs extracted from these dictionaries served as a golden standard and remained blind to the participants. Notice that the Brazilian Portuguese variant was used for the translations to/from English (whereas the European Portuguese variant was used with French), which might introduce a bias; however its influence should be equivalent to every participant system thus still allowing for a valid comparison.

Given the fact that the coverage of KD is not the same as Apertium, we took the subset of KD that is covered by Apertium to build the gold standard and allow comparisons, i.e., those KD translations for which the source and target terms are present in both Apertium RDF source and target lexicons.

Table 1 shows the size (in number of translations) of the different language pairs in the gold standard. This number might differ from previous TIAD editions because since TIAD'20 the golden standard data have been curated with respect to the initial version in several aspects (see (Kernerman et al., 2020)) and, further, the use of a larger Apertium graph since TIAD'21 might have slightly changed the overlap degree between Apertium lexica and KD data.

Table 1: Number of translations per language pair in the gold standard.

Language pair	Size
EN-FR	12,453
EN-PT	10,151
FR-EN	16,103
FR-PT	7,982
PT-EN	12,219
PT-FR	6,589

4. Evaluation methodology

The participants run their systems locally, using the Apertium RDF data as known translations, to infer new translations among the three studied languages: FR, EN, PT. Once the output data (inferred translations) were obtained, they loaded the results into a file per

[12] http://linguistic.linkeddata.es/id/apertium/lexiconEN/bench-n-en

[13] https://tinyurl.com/apertiumrdfv2

[14] https://pret-a-llod.eu/

[15] Hosted by the University of Frankfurt at http://dbserver.acoli.cs.uni-frankfurt.de:5005/dataset.html. The queries should be restricted to this graph: http://linguistic.linkeddata.es/id/apertium-ud. Since this is for testing purposes, there is no guarantee of a quick and efficient response, and the link may not be persistent long after the evaluation campaign. See an example query at https://ndownloader.figshare.com/files/26321950

[16] https://tiad2021.unizar.es/data/TransSets_ApertiumRDFv2_1_CSV.zip

[17] See the "how to get the data source" section at https://tiad2022.unizar.es/task.html

language pair in TSV format, containing the following information per row (tab separated):

"source written representation"
"target written representation"
"part of speech"
"confidence score"

The confidence score takes float values between 0 and 1 and is a measure of the confidence that the translation holds between the source and target written representations. If a system does not compute confidence scores, this value had to be put to 1.

4.1. Evaluation process

The organisers compared the obtained results with the gold standard automatically. This process was followed for each system results file and per language pair:

1. Remove duplicated translations (if any).
2. Filter out translations for which the source entry or the target entry are not present in the golden standard (otherwise we cannot assess whether the translation is correct or not). We call *systemGS* the subset of translations that passed this filter, and *GS* the whole set of gold standard translations, in the given language pair.
3. Translations with confidence degree under a given threshold were removed from $systemGS$. In principle, the used threshold is the one reported by participants as the optimal one during the training/preparation phase.
4. Compute the coverage of the system with respect to the gold standard, i.e., how many gold standard entries in the source language were effectively translated by the system (no matter if they were correct or wrong ones).
5. Compute precision as P = (#correct translations in systemGS) / |systemGS|
6. Compute recall as R = (#correct translations in systemGS) / |GS|
7. Compute F-measure as $F = 2 * P * R / (P + R)$

The precision/recall metrics calculated after applying steps 1 to 3 correspond to what in (Goel et al., 2021) is defined as *both-word precision* and *both-word recall*. The idea is to reduce the penalization to a system for inferring correct translations that are missing in the golden standard dictionary because human editors might have overlooked them when elaborating the dictionary. Note that in TIAD editions previous to TIAD'21 we only filtered out translations for which the source entry was not present in the translation (step 2), which led to computing the so-called one-word precision/recall, thus only partially covering such a goal.

4.2. Baselines

We have run the above evaluation process with results obtained with two baselines, to be compared with the participating systems' results:

4.2.1. Baseline 1 - Word2Vec

The method uses Word2Vec (Mikolov et al., 2013) to transform the graph into a vector space. A graph edge is interpreted as a sentence and the nodes are word forms with their POS tag. Word2Vec iterates multiple times over the graph and learns multilingual embeddings (without additional data). We used the Gensim[18] Word2Vec implementation. For a given input word, we calculated a distance based on the cosine similarity of a word to every other word with the target-POS tag in the target language. The square of the distance from source to target word is interpreted as the confidence degree. For the first word the minimum distance is 0.6^2, for the others it is 0.8^2. Therefore multiple results are only in the output if the confidence is not extremely weak. In our evaluation, we applied an arbitrary threshold of 0.5 to the confidence degree[19].

4.2.2. Baseline 2 - OTIC

In short, the idea of the One Time Inverse Consultation (OTIC) method (Tanaka and Umemura, 1994) is to explore, for a given word, the possible candidate translations that can be obtained through intermediate translations in the pivot language. Then, a score is assigned to each candidate translation based on the degree of overlap between the pivot translations shared by both the source and target words[20]. In our evaluation, we applied the OTIC method using Spanish as pivot language, and using an arbitrary threshold of 0.5.
Note that since the TIAD'21 edition, the Word2Vec baseline, although based on the same principles, was re-implemented and re-trained to be adapted to the new Apertium RDF v2 dataset, thus leading to different (generally better) results than in the previous TIAD editions. The OTIC baseline, although it does not need re-training, was also re-run for TIAD'21 to be adapted to the new Apertium RDF v2 dataset (the new baseline results remain valid for TIAD'22). The results are generally worse than in TIAD'20 (with the smaller Apertium RDF v1 graph).
Strictly speaking, these are not baselines as they are conceived in other shared tasks, meaning naive approaches with a straightforward implementation, but state-of-the-art methods to solve the task.

5. Results

In this section we review the participating systems in TIAD 2022 and their evaluation results.

[18] https://radimrehurek.com/gensim/

[19] The code can be found at https://github.com/kabashi/TIAD2022_word2vec

[20] You can find the code at https://gitlab.com/sid_unizar/otic

5.1. Participating systems

Two teams participated in this edition of the shared task, contributing with four systems or system variants. Table 2 lists the participant teams and systems.

The first team, L. Dranca from Centro Universitario de la Defensa (CUD), Spain, developed three variants of a system that was based on the use of FastRP (Chen et al., 2019). The algorithm, generates embeddings form a graph node (in this case words) based on the neighbourhood information, in this case translations into other languages. Thus, words with similar translations will have similar FastRP embeddings. They use ES as a pivot language, or both ES and CA. Note that we cannot refer to a detailed description of the system because the author decided not to publish their system description paper, nor to participate in the workshop. We still include their result here for completeness.

The second team, Y. Bestgen (Bestgen, 2022) from Universite catholique de Louvain, Belgium, presented a system that combines a classical machine learning technique such as logistic regression with the use of pivot languages to obtain inferred translations.

5.2. Evaluation results

The complete evaluation results per system and per language pair are accessible in the TIAD 2022 website[21]. In order to give an overview of the results, we include here Table 3, which shows the averaged results, evaluated by using the confidence threshold that every participant reported as optimal according to their internal tests. Since the evaluation setup was identical as in TIAD 2021, we combine in the table the results of both evaluation campaigns.

5.3. Discussion

As can be seen in Table 3, two of the four systems obtained better results than both baselines in terms of F-measure. This continues a trend started in TIAD 2021 when some systems were able to beat both baselines, since in previous TIAD editions there was no system beating both baselines. Interestingly, the OTIC method, based on purely graph exploration and dated back to 1994, systematically outperformed more contemporary methods based on word embeddings and distributional semantics, which gives an idea of the difficulty of the task. The last two years' results confirm our intuition that OTIC was not an upper bound and that there were still much room for improvement for more methods.

Note that the precision values shown in Table 3 are conservative since there is a small but undefined number of false negatives (correct translations that are not present in the gold standard) that can be found in the results. For example, from the EN→FR set of translations are as follows: "wizard"→"sorcier" (noun), "abandon"→"quitter" (verb) and the "dump"→"vider" (verb).

[21]Cf. `https://tiad2022.unizar.es/results.html` under the section "Evaluation results".

6. Conclusions

In this paper we have given an overview of the 5th Translation Inference Across Dictionaries (TIAD) shared task, and a description of the results obtained by the four participating systems and two baselines, compared also with the results of the previous campaign. In this edition, the participating systems were asked to generate new translations automatically among English, French, Portuguese, based on known indirect translations contained in the Apertium RDF graph. Same as in the previous edition, a new larger version of the data graph was used, that is Apertium RDF v2. The evaluation of the results was carried out by the organisers against manually compiled pairs of K Dictionaries. The results are good (two systems beat the baselines), are along the lines of the previous edition, and illustrate improvement in the area of translation inference across dictionaries despite the difficulty of the task. However, we consider that the task is far from being solved, with much room for improvement and other aspects and languages to be explored.

7. Acknowledgements

We would like to thank Miguel López-Otal (University of Zaragoza) for his assistance with the baselines and evaluation process. This work has been supported by the European Union's Horizon 2020 research and innovation programme through the projects Prêt-à-LLOD (grant agreement No 825182) and Elexis (grant agreement No 731015). This work is also based upon work from COST Action CA18209 – NexusLinguarum "European network for Web-centred linguistic data science", supported by COST (European Cooperation in Science and Technology). It has been also partially supported by the Spanish project PID2020-113903RB-I00 (AEI/FEDER, UE), by DGA/FEDER, and by the *Agencia Estatal de Investigación* of the Spanish Ministry of Economy and Competitiveness and the European Social Fund through the "Ramón y Cajal" program (RYC2019-028112-I).

8. Bibliographical References

Bestgen, Yves. (2022). Creating bilingual dictionaries from existing ones by means of pivot-oriented translation inference and logistic regression. In *Proc. of Globalex 2022*.

Chen, Haochen and Sultan, Syed Fahad and Tian, Yingtao and Chen, Muhao and Skiena, Steven. (2019). Fast and accurate network embeddings via very sparse random projection. In *Proc. of International Conference on Information and Knowledge Management*, pages 399–408. Association for Computing Machinery, nov.

Flati, Tiziano and Navigli, Roberto. (2013). The CQC Algorithm: Cycling in Graphs to Semantically Enrich and Enhance a Bilingual Dictionary (Extended Abstract). In *Proce. of the 23th International Joint*

Table 2: Participant systems.

Team	System
Lacramiora Dranca (Centro Universitario de la Defensa, Spain)	FastRP_ES FastRP_fasttext_ES FastRP_fasttext_ES_CA
Yves Bestgen (Université catholique de Louvain, Belgium)	SATLab

Table 3: Averaged system results, ordered by F-measure in descending order. In **bold** the baselines, and in *italics* you can find the participants of TIAD 2022 (the rest are participants of TIAD 2021)

System	Precision	Recall	F-measure	Coverage
PivotAlign-R	0.71	0.58	0.64	0.77
PivotAlign-F	0.81	0.51	0.62	0.68
SATLab	0.86	0.48	0.62	0.70
ACDcat	0.75	0.53	0.61	0.75
TUANWEsg	0.81	0.47	0.59	0.76
TUANWEcb	0.81	0.47	0.59	0.76
ULD_graphSVR	0.70	0.49	0.57	0.69
fastRP_fastText_ES_CA	0.85	0.28	0.42	0.43
PivotAlign-P	0.86	0.24	0.37	0.33
baseline-Word2Vec	0.69	0.23	0.33	0.40
ULD_MUSE	0.29	0.41	0.33	0.65
baseline-OTIC	0.78	0.18	0.29	0.28
fastRP_fastText_ES	0.83	0.15	0.25	0.25
fastRP_ES	0.83	0.15	0.25	0.25
ULD_onetaSVR	0.76	0.10	0.17	0.14
TUANMUSEca	0.86	0.10	0.16	0.16
TUANMUSEes	0.87	0.08	0.13	0.14
ULD_oneta2	0.64	0.07	0.13	0.11
ULD_vecmap	0.36	0.01	0.01	0.02
ULD_mbert	0.00	0.00	0.00	0.11

Conference on Artificial Intelligence, IJCAI '13, pages 3151–3155. AAAI Press.

Forcada, Mikel L. and Ginestí-Rosell, Mireia and Nordfalk, Jacob and O'Regan, Jim and Ortiz-Rojas, Sergio and Pérez-Ortiz, Juan Antonio and Sánchez-Martínez, Felipe and Ramírez-Sánchez, Gema and Tyers, FrancisM. (2011). Apertium: a free/open-source platform for rule-based machine translation. *Machine Translation*, 25(2):127–144.

Goel, Shashwat and Gracia, Jorge and Forcada, Mikel Lorenzo. (2021). Bilingual dictionary generation and enrichment via graph exploration. *Semantic Web Journal*.

Gracia, Jorge and Villegas, Marta and Gómez-Pérez, Asunción and Bel, Núria. (2018). The apertium bilingual dictionaries on the web of data. *Semantic Web*, 9(2):231–240.

Gracia, Jorge and Fäth, Christian and Hartung, Matthias and Ionov, Max and Bosque-Gil, Julia and Veríssimo, Susana and Chiarcos, Christian and Orlikowski, Matthias. (2020). Leveraging Linguistic Linked Data for Cross-Lingual Model Transfer in the Pharmaceutical Domain. In Bo Fu et al., editors, *Proc. of 19th International Semantic Web Conference (ISWC 2020)*, pages 499–514. Springer.

Kernerman, Ilan and Krek, Simon and McCrae, John P. and Gracia, Jorge and Ahmadi, Sina and Kabashi, Besim. (2020). Introduction to the proceedings of globalex 2020 workshop on linked lexicography. In *Proc. of Globalex 2020 Workshop on Linked Lexicography – Twelfth International Conference on Language Resources and Evaluation (LREC 2020), Marseílle, France*. European Language Resources Association (ELRA).

Lim, Lian T and Ranaivo-Malançon, Bali and Tang, Enya K. (2011). Low Cost Construction of a Multilingual Lexicon from Bilingual Lists. *Polibits*, 43:45–51.

Mausam and Soderland, Stephen and Etzioni, Oren and Weld, Daniel S and Skinner, Michael and Bilmes, Jeff. (2009). Compiling a Massive, Multilingual Dictionary via Probabilistic Inference. In *Proc. of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 1 - Volume 1*, ACL '09, pages 262–270, Stroudsburg, PA, USA. Association for Computational Linguistics.

McCrae, John and Aguado-de-Cea, Guadalupe and Buitelaar, Paul and Cimiano, Philipp and Declerck, Thierry and Gómez-Pérez, Asunción and Gracia, Jorge and Hollink, Laura and Montiel-Ponsoda, Elena and Spohr, Dennis and Wunner, Tobias. (2012). Interchanging lexical resources on the Semantic Web. *Language Resources and Evaluation*, 46:701–719.

McCrae, John P and Bosque-Gil, Julia and Gracia, Jorge and Buitelaar, Paul and Cimiano, Philipp. (2017). The OntoLex-Lemon Model: Development and Applications. In *Electronic lexicography in the 21st century. Proc. of eLex 2017 conference, in Leiden, Netherlands*, pages 587–597. Lexical Computing CZ s.r.o., sep.

Mikolov, Tomas and Chen, Kai and Corrado, Greg and Dean, Jeffrey. (2013). Efficient Estimation of Word Representations in Vector Space. In *Proc. of International Conference on Learning Representations (ICLR)*.

Montiel-Ponsoda, E. and Gracia, J. and Aguado-De-Cea, G. and Gómez-Pérez, A. (2011). Representing translations on the semantic Web. In *Proc. of the 2nd International Workshop on the Multilingual Semantic Web (MSW) at ISWC '11*, volume 775. CEUR Press.

Tanaka, Kumiko and Umemura, Kyoji. (1994). Construction of a Bilingual Dictionary Intermediated by a Third Language. In *COLING*, pages 297–303.

Villegas, Marta and Melero, Maite and Bel, Núria and Gracia, Jorge and Bel, Núria. (2016). Leveraging RDF Graphs for Crossing Multiple Bilingual Dictionaries. In *Proc. of 10th Language Resources and Evaluation Conference (LREC'16) Portorož (Slovenia)*, pages 868–876, Paris, France, may. European Language Resources Association (ELRA).