An Unsupervised Approach for Bootstrapping Arabic Sense Tagging

Mona T. Diab
Stanford University
Stanford, CA 94305, USA
mdiab@stanford.edu

Abstract

To date, there are no WSD systems for Arabic. In this paper we present and evaluate a novel unsupervised approach, SALAAM, which exploits translational correspondences between words in a parallel Arabic English corpus to annotate Arabic text using an English WordNet taxonomy. We illustrate that our approach is highly accurate in $\leq 90.1\%$ of the evaluated data items based on Arabic native judgement ratings and annotations. Moreover, the obtained results are competitive with state-of-the-art unsupervised English WSD systems when evaluated on English data.

1 Introduction

Word Sense Disambiguation (WSD) is the process of resolving the meaning of a word unambiguously in a given natural language context. Within the scope of this paper, it is the process of marking text with an explicit set of sense tags or labels from some predefined tag set. It is well established that in order to obtain best quality sense annotations of words in running text, one needs a wide coverage lexicon and a trained lexicographer to annotate the words manually with their appropriate senses. Such a task is very tedious, expensive, and, by many standards, daunting to the people involved, even when all the required resources are available (Fellbaum et al., 2001). The problem becomes ever more challenging when dealing with a language with virtually no automated knowledge resources or tools. Like the majority of natural languages, the Arabic language happens to fall in this category of languages with minimal automated resources.

The focus of this paper is the sense disambiguation of Modern Standard Arabic which is the language used in formal speech and writing in the Arab world; Moreover, the script is shared with Urdu, Farsi, Dari and Pashtu. To our knowledge, there are no Arabic WSD systems reported in the literature.

Arabic is a Semitic language with rich templatic morphology. An Arabic word in text or speech may be composed of a stem, plus affixes and clitics. The affixes include inflectional markers for tense, gender, and/or number. The clitics include some (but not all) prepositions, conjunctions, determiners, possessive pronouns and pronouns. The stems consist of an underlying consonantal root and a template. The root could be anywhere from two to four consonants devoid of vocalization. Typically text in Modern Standard Arabic is written in the stem surface form with the various affixes. However, most Arabic dictionaries list the entries in terms of roots rather than surface forms.

In this paper, we present an approach, SALAAM (Sense Annotations Leveraging Alignments And Multilinguality), to bootstrap WSD for Arabic text presented in surface form. The approach of SALAAM is based on work by (Diab and Resnik, 2002) but it goes beyond it in the sense of extending the approach to the tagging of Arabic as a target language.(Diab, 2003) SALAAM uses cross-linguistic correspondences for characterizing word meanings in natural language. This idea is explored by several researchers, (Resnik and Yarowsky, 1998; Chugur et al., 2002; Ide, 2000; Dyvik, 1998). Basically, a word meaning or a word sense is quantifiable as much as it is uniquely translated in some language or set of languages. SALAAM is an empirical validation of this very notion of characterizing word meaning using cross-linguistic correspondences. Since automated lexical resources are virtually non-existent for Arabic, SALAAM leverages sense ambiguity resolution for Arabic off of existing English lexical resources and an Arabic English parallel corpus, thereby providing a bilingual solution to the WSD problem.

The paper is organized as follows: Section 2 describes the SALAAM system; Section 3 presents an evaluation of the approach followed by Section 4 which discusses the chosen sense inventory in relation to the Arabic data; We conclude with a summary and some final remarks in Section 6.

2 Approach

SALAAM exploits parallel corpora for sense annotation. The key intuition behind SALAAM is that when words in one language, *L1*, are translated into the same word in a second language, *L2*, then the *L1* words are semantically similar. For example, when the English — *L1* — words *bank, brokerage, mortgage-lender* translate into the Arabic — *L2* — word *bnk* (بنك) in a parallel corpus,[1] where the *bank* is polysemous, SALAAM discovers that the intended sense for the English word *bank* is the *financial institution* sense, not the *geological formation* sense, based on the fact that it is grouped with *brokerage* and *mortgage-lender*. Two fundamental observations are at the core of SALAAM:

- **Translation Distinction Observation (TDO)**

  Senses of ambiguous words in one language are often translated into distinct words in a second language.

  To exemplify **TDO**, we consider a sentence such as *I walked by the bank*, where the word *bank* is ambiguous with n senses. A translator may translate *bank* into *Dfp* (ضفة) corresponding to the *GEOLOGICAL FORMATION* sense or to *bnk* (بنك) corresponding to the *FINANCIAL INSTITUTION* sense depending on the surrounding context of the given sentence. Essentially, translation has distinctly differentiated two of the possible senses of *bank*.

- **Foregrounding Observation (FGO)**

  If two or more words are translated into the same word in a second language, then they often share some element of meaning.

  FGO may be expressed in quantifiable terms as follows: if several words $(w_1, w_2, \ldots, w_x)$ in *L1* are translated into the same word form in *L2*, then $(w_1, w_2, \ldots, w_x)$ share some element of meaning which brings the corresponding relevant senses for each of these words to the foreground. For example, if the word *Dfp* (ضفة), in Arabic, translates in some instances in a corpus to *shore* and other instances to *bank*, then *shore* and *bank* share some meaning component that is highlighted by the fact that the translator chooses the same Arabic word for their translation. The word *Dfp* (ضفة), in this case, is referring to the concept of *LAND BY WATER SIDE*, thereby making the corresponding senses in the English words more salient. It is important to note that the foregrounded senses of *bank* and *shore* are not necessarily identical, but they are quantifiably the closest senses to one another among the various senses of both words.

Given observations **TDO** and **FGO**, the crux of the SALAAM approach aims to quantifiably exploit the translator's implicit knowledge of sense representation cross-linguistically, in effect, reverse engineering a relevant part of the translation process.

SALAAM's algorithm is as follows:

- SALAAM expects a word aligned parallel corpus as input;
- *L1* words that translate into the same *L2* word are grouped into clusters;
- SALAAM identifies the appropriate senses for the words in those clusters based on the words senses' proximity in WordNet. The word sense proximity is measured in information theoretic terms based on an algorithm by Resnik (Resnik, 1999);
- A sense selection criterion is applied to choose the appropriate sense label or set of sense labels for each word in the cluster;
- The chosen sense tags for the words in the cluster are propagated back to their respective contexts in the parallel text. Simultaneously, SALAAM projects the propagated sense tags for *L1* words onto their *L2* corresponding translations.

The focus of this paper is on the last point in the SALAAM algorithm, namely, the sense projection phase onto the *L2* words in context. In this case, the *L2* words are Arabic and the sense inventory is the English WordNet taxonomy. Using SALAAM we annotate Arabic words with their meaning definitions from the English WordNet taxonomy. We justify the usage of an English inventory on both empirical and theoretical grounds. Empirically, there are no automated sense inventories for Arabic; Furthermore, to our knowledge the existing MRDs for Arabic are mostly root based which introduces another layer of ambiguity into Arabic processing

[1] We use the Buckwalter transliteration scheme for the Arabic words in this paper. http://www.ldc.org/aramorph

since Modern Standard Arabic text is rendered in a surface form relatively removed from the underlying root form. Theoretically, we subscribe to the premise that people share basic conceptual notions which are a consequence of shared human experience and perception regardless of their respective languages. This premise is supported by the fact that we have translations in the first place. Accordingly, basing the sense tagging of *L2* words with corresponding *L1* sense tags captures this very idea of shared meaning across languages and exploits it as a bridge to explicitly define and bootstrap sense tagging in *L2*, Arabic.

3 Evaluation

In order to formally evaluate SALAAM for Arabic WSD, there are several intermediary steps. SALAAM requires a token aligned parallel corpus as input and a sense inventory for one of the languages of the parallel corpus. For evaluation purposes, we need a manually annotated gold standard set.

3.1 Gold Standard Set

As mentioned above, there are no systems that perform Arabic WSD, therefore there exist no Arabic gold standard sets as such. Consequently, one needs to create a gold standard. Since SALAAM depends on parallel corpora, an English gold standard with projected sense tags onto corresponding Arabic words would serve as a good start. A desirable gold standard would be generic covering several domains, and would exist in translation to Arabic. Finding an appropriate English gold standard that satisfies both attributes is a challenge. One option is to create a gold standard based on an existing parallel corpus such as the Quran, the Bible or the UN proceedings. Such corpora are single domain corpora and/or their language is stylistic and distant from everyday Arabic; Moreover, the cost of creating a manual gold standard is daunting. Alternatively, the second option is to find an existing English gold standard that is diverse in its domain coverage and is clearly documented. Fortunately, the SENSEVAL2 exercises afford such sets.[2] SENSEVAL is a series of community-wide exercises that create a platform for researchers to evaluate their WSD systems on a myriad of languages using different techiques by constantly defining consistent standards and robust measures for WSD.

Accordingly, the gold standard set used here is the set of 671 Arabic words corresponding to the correctly sense annotated English nouns from the SENSEVAL2 English All Words Task. SALAAM achieved a precision of 64.5% and recall of 53% on the English test set for that task. SALAAM ranks as the best unsupervised system when compared to state-of-the-art WSD systems on the same English task. The English All Words task requires the WSD system to sense tag every content word in an English language text.

3.2 Token Aligned Parallel Corpora

The gold standard set corresponds to the test set in an unsupervised setting. Therefore the test set corpus is the SENSEVAL2 English All Words test corpus which comprises three articles from the Wall Street Journal discussing religious practice, medicine and education. The test corpus does not exist in Arabic. Due to the high expense of manually creating a parallel corpus, i.e. using human translators, we opt for automatic translation systems in a fashion similar to (Diab, 2000). To our knowledge there exist two off the shelf English Arabic Machine Translation (MT) systems: Tarjim and Almisbar.[3] We use both MT systems to translate the test corpus into Arabic. We merge the outputs of both in an attempt to achieve more variability in translation as an approximation to human quality translation. The merging process is based on the assumption that the MT systems rely on different sources of knowledge, different dictionaries in the least, in their translation process.

Fortunately, the MT systems produce sentence aligned parallel corpora.[4] However, SALAAM expects token aligned parallel corpora. There are several token alignment programs available. We use the GIZA++ package which is based on the IBM Statistical MT models.[5] Like most stochastic NLP applications, GIZA++ requires large amounts of data to produce reliable quality alignments. The test corpus is small comprising 242 lines only; Consequently, we augment the test corpus with several other corpora. The augmented corpora need to have similar attributes to the test corpus in genre and style. The chosen corpora and their relative sizes are listed in Table 1.

BC-SV1 is the Brown Corpus and SENSEVAL1 trial, training and test data. SV2-LS is the SENSEVAL2 English Lexical Sample trial, training and test data. WSJ is the Wall Street Journal. Finally SV2AW is SENSEVAL2 English All Words test corpus.

[2] http://www.senseval.org

[3] http://www.Tarjim.com, http://www.almisbar.com

[4] This is not a trivial problem with naturally occurring parallel corpora.

[5] http://www.isi.edu/och/GIZA++.html

Corpora	Lines	Tokens
BC-SV1	101841	2498405
SV2-LS	74552	1760522
WSJ	49679	1290297
SV2AW	**242**	**5815**
Total	*226314*	*5555039*

Table 1: Relative sizes of corpora used for evaluating SALAAM

The three augmenting corpora, BC-SV1, SV2LS and WSJ are translated into Arabic using both MT systems, AlMisbar and Tarjim. All the Arabic corpora are transliterated using the Buckwalter transliteration scheme and then tokenized. The corpora are finally token aligned using GIZA++. Figure 1 illustrates the first sentence of the SV2AW English test corpus with its translation into Arabic using AlMisbar MT system followed by its transliteration and tokenization, respectively.[6]

The art of change-ringing is peculiar to the English, and, like most English peculiarities, unintelligible to the rest of the world.

إن فن تغيير الدقاق خاص بالأنجليز، ومثل أكثر الخواص
الأنجليزية، غير واضح إلى بقية العالم

{n fn tgyyr AldqAq xAS bAl{njlyz, wmvl Akvr AlxwAS Al{njlyzyp, gyr wADH Ila bqyp AlEAlm.

{n fn tgyyr Al dqAq xAS b Al {njlyz , w mvl Akvr Al xwAS Al {njlyzyp , gyr wADH Ila bqyp Al EAlm .

Figure 1: First sentence in test corpus SV2AW and its Arabic translation, transliteration and tokenization

3.3 Sense Inventory

The gold standard set is annotated using the WordNet taxonomy, WN1.7pre, for English. Like previous WordNet editions (Fellbaum, 1998), WN17pre is a computational semantic lexicon for English. It is rapidly becoming the community standard lexical resource for English since it is freely available for academic research. It is an enumerative lexicon in a Quillian style semantic network that combines the knowledge found in traditional dictionaries (Quillian, 1968). Words are represented as concepts, referred to as synsets, that are connected via different types of relations such as hyponymy, hypernymy, synonymy, meronymy, antonymy, etc. Words are represented as their synsets in the lexicon. For example, the word *bank* has 10 synsets in WN17pre corresponding to 10 different senses. The concepts are organized taxonomically in a hierarchical structure with the more abstract or broader concepts at the top of the tree and the specific concepts toward the bottom of the tree. For instance, the concept *FOOD* is the hypernym of the concept *FRUIT*, for instance.

Similar to previous WordNet taxonomies, WN17pre comprises four databases for the four major parts of speech in the English language: nouns, verbs, adjectives, and adverbs. The nouns database consists of 69K concepts and has a depth of 15 nodes. The nouns database is the richest of the 4 databases. Majority of concepts are connected via the IS-A identity relation. The focus of this paper is exclusively on nouns.[7]

3.4 Experiment and Metrics

We conducted two experiments.

3.4.1 Experiment 1

In the first experiment a native speaker of Arabic with near native proficiency in English is asked to pick the appropriate meaning definition of an Arabic word — given in its Arabic context sentence in which it appears in the corpus — from the list of WN1.7pre definitions. They are allowed to pick more than one definition for each item. Or alternatively, the annotator has the option to choose *NONE* where none of the definitions is appropriate for the Arabic word given the Arabic context sentence; Or *MISALIGNMENT* where the Arabic word is not a translation of the English word whose meaning definitions appear in the list that follows, or it is simply a misalignment. The results from this experiment are illustrated in Table 2.

Category	Num. of items	%
Agreement	605	90.1
Disagreement	21	3.1
None	1	0.14
Misalignment	44	6.55

Table 2: Human Annotator agreement scores with SALAAM automatic annotations.

It is worth noting the high agreement rate between the annotator and the SALAAM annotations

[6] All the Arabic sentences in this paper are output from one of the MT systems used.

[7] SALAAM, however, has no inherent restriction on part of speech.

which exceed 90%. The only case that is considered a ”NONE” category is for the word *bit* which is translated as the past tense of *to bite* as عض. It should have been translated as قطعة meaning a *morsel/piece*.

3.4.2 Experiment 2

In this experiment, the Arabic words annotated with English WN1.7pre tags are judged on a five point scale metric by three native speakers of Arabic with near native proficiency in English. The experiment is run in a form format on the web. The raters are asked to judge the accurateness of the chosen sense definition from a list of definitions associated with the translation of the Arabic word. The Arabic words are given to the raters in their respective context sentences. Therefore the task of the rater is to judge the appropriateness of the chosen English sense definition for the Arabic word given its context. S/he is required to pick a rating from a drop down menu for each of the data items. The five point scale is as follows:

- **Accurate**: This choice indicates that the chosen sense definition is an appropriate meaning definition of the Arabic word.
- **Approximate**: This choice indicates that the chosen sense definition is a good meaning definition for the Arabic word given the context yet there exists on the list of possible definitions a more appropriate sense definition.
- **Misalignment**: This choice indicates that the Arabic word is not a translation of the English word due to a misalignment or the word being rendered in English in the Arabic sentence, i.e. the English word was not translated by either of the Arabic MT systems.
- **None**: This choice indicates that none of the sense definitions listed is an appropriate sense definition for the Arabic word.
- **Wrong**: This choice indicates that the chosen sense definition is the incorrect meaning definition for the Arabic word given its context.

3.5 Results

Table 3 illustrates the obtained results from the three raters.

The inter-rater agreement is at a high 96%. They all deemed on average more than 90% of the data items to be accurately tagged by SALAAM. The most variation seemed to be in assessing the APPROXIMATE category with Rater 1, R1, rating 19 items as APPROXIMATE and R2 rating 10 items as APPROXIMATE and R3 rating 14 data items as APPROXIMATE.

Type	R1	R2	R3
Accurate	90.3	90.4	91.4
Approximate	2.8	2	1.5
Misalignment	5.6	5.9	5.9
None	0	0	0
Wrong	1.2	1.3	1.2

Table 3: Rater judgments on the Arabic WSD using meaning definitions from the English WN1.7pre

An example of a data item that is deemed APPROXIMATE by the three raters is for the word *AltjmE* (التجمع) in the following sentence:

تدق فرقة جديدة كليا كل يوم فى تورنجتون العظيمة،
عدة من أعضاء التجمع

transliterated as

tdq frqp jdydp klyA kl ywm fy twrnjtwn AlEZymp, edp mn AED’ AltjmE

which means

In Great Torington, a brand new band plays everyday comprising members of the congregation

The word *AltjmE* (التجمع) is a translation of *congregation* which has the following sense definitions in WN1.7pre:

- congregation: an assemblage of people or animals or things collected together; "a congregation of children pleaded for his autograph"; "a great congregation of birds flew over"
- congregation, fold, faithful: a group of people who adhere to a common faith and habitually attend a given church
- congregation, congregating: the act of congregating

SALAAM favors the last meaning definition for *congregation*.

An example of a MISALIGNMENT is illustrated in the following sentence:

القولون والرئه وسرطان الثدى أكثر الأشكال القاتله للمرض

transliterated as

Alqwlwn wAlr'p wsrTAn Alvdy Akvr AlA$kAl AlqAtlp llmrD...

which is a translation of

Cancer of the Colon, Breast and Lungs are the most deadly forms of the disease...

The words *srTAn* (سرطان), meaning *cancer*, and *lungs* were aligned leading to tagging the Arabic word with the sense tag for the English word *lungs*. Finally, the following is an example of a WRONG data item as deemed by the three raters. The definition for the word *Alywm* (اليوم) in the following sentence:

يعيش الأخرون اليوم فى مكان آخر

transliterated as

yEy$ AlAxrwn Alywm fy mkAn Axr...

which means

The others live today in a different place...

where the word equivalent to *today* is the target word with the following sense definitions:

- today: the day that includes the present moment (as opposed to yesterday or tomorrow); "Today is beautiful"; "did you see today's newspaper?"
- today: the present time or age; "the world of today"; "today we have computers"

SALAAM chooses the first meaning definition while the raters seem to favor the second.

None of the raters seemed to find data items that had no corresponding meaning definition in the given list of English meaning definitions. It is interesting to note that the single item considered a ”NONE” category in experiment 1 was considered a misalignment by the three raters.

If we calculate the average precision of the evaluated sense tagged Arabic words based on the total tagged English nouns of 1071 nouns in this test set, we obtain an absolute precision of 56.9% for Arabic sense tagging. It is worth noting that the average precision on the SENSEVAL2 English All Words Task for any of the unsupervised systems is in the lower 50% range.

4 General Discussion

It is worth noting the high agreement level between the rating judgments of the three raters in experiment 2 and the human manual annotations of experiment 1. The obtained results are very encouraging indeed but it makes the implicit assumption that the English WordNet taxonomy is sufficient for meaning representation of the Arabic words used in this text. In this section, we discuss the quality of WN1.7pre as an appropriate sense inventory for the Arabic task.

With that intent in mind, we evaluate the 600 word instances of Arabic that are deemed correctly tagged using the English WN17pre.[8] We investigate three different aspects of the Arabic English correspondence: Arabic and English words are equivalent; Arabic words correspond to specific English senses; And English words do not sufficiently correspond to all possible senses for the Arabic word. The three aspects are discussed in detail below.

- **Arabic and English words are equivalent**

  We observe that a majority of the ambiguous words in Arabic are also ambiguous in English in this test set; they preserve ambiguity in the same manner. In Arabic, 422 word tokens corresponding to 190 word types, are at the closest granularity level with their English correspondent;[9] For instance, all the senses of *care* apply to its Arabic translation *EnAyA* (عنايه); the sense definitions are listed as follows:

  - care, attention, aid, tending: the work of caring for or attending to someone or something; "no medical care was required"; "the old car needed constant attention"
  - caution, precaution, care, forethought: judiciousness in avoiding harm or danger; "he exercised caution in opening the door"; "he handled the vase with care"

[8] The overlapping number of Arabic words rated ACCURATE by the three annotators of experiment 1 and those accurate items from experiment 1.

[9] This means that all the English senses listed for WN17pre are also senses for the Arabic word.

- concern, care, fear: an anxious feeling; "care had aged him"; "they hushed it up out of fear of public reaction"
- care: a cause for feeling concern; "his major care was the illness of his wife"
- care, charge, tutelage, guardianship: attention and management implying responsibility for safety; "he is under the care of a physician"
- care, maintenance, upkeep: activity involved in maintaining something in good working order; "he wrote the manual on car care"

It is worth noting that the cases where ambiguity is preserved in English and Arabic are all cases where the polysemous word exhibits regular polysemy and/or metonymy. The instances where homonymy is preserved are borrowings from English. Metonymy is more pragmatic than regular polysemy (Cruse, 1986); for example, *tea* in English has the following metonymic sense from WN1.7pre:

- a reception or party at which tea is served; "we met at the Dean's tea for newcomers"

This sense of *tea* does not have a correspondent in the Arabic *$Ay* (شاى). Yet, the English *lamb* has the metonymic sense of *MEAT* which exists in Arabic. Researchers building EuroWordNet have been able to devise a number of consistent metonymic relations that hold cross linguistically such as *fabric/material, animal/food, building/organization* (Vossen et al., 1999; Wim Peters and Wilks, 2001). In general, in Arabic, these defined classes seem to hold, however, the specific case of *tea* and *party* does not exist. In Arabic, the English sense would be expressed as a compound *tea party* or *Hflp $Ay* (حفلة شاى).

- **Arabic word equivalent to specific English sense(s)**

In this evaluation set, there are 138 instances where the Arabic word is equivalent to a subsense(s) of the corresponding English word. The 138 instances correspond to 87 word types. An example is illustrated by the noun *ceiling* in English.

- ceiling: the overhead upper surface of a room; "he hated painting the ceiling"
- ceiling: (meteorology) altitude of the lowest layer of clouds
- ceiling, cap: an upper limit on what is allowed: "they established a cap for prices"
- ceiling: maximum altitude at which a plane can fly (under specified conditions)

The correct sense tag assigned by SALAAM to *ceiling* in English is the first sense, which is correct for the Arabic translation *sqf* (سقف). Yet, the other 3 senses are not correct translations for the Arabic word. For instance, the second sense definition would be translated as {*rtfAE* (إرتفاع) and the last sense definition would be rendered in Arabic as *Elw* (علو). This phenomenon of Arabic words corresponding to specific English senses and not others is particularly dominant where the English word is homonymic. By definition, homonymy is when two independent concepts share the same orthographic form, in most cases, by historical accident. Homonymy is typically preserved between languages that share common origins or in cases of cross-linguistic borrowings. Owing to the family distance between English and Arabic, polysemous words in Arabic rarely preserve homonymy.

- **English word equivalent to specific Arabic sense**

40 instances, corresponding to 20 type words in Arabic, are manually classified as more generic concepts than their English counterparts. For these cases, the Arabic word is more polysemous than the English word. For example, the English noun *experience* possesses three senses in WN17pre as listed below.

- experience: the accumulation of knowledge or skill that results from direct participation in events or activities; "a man of experience"; "experience is the best teacher"
- experience: the content of direct observation or participation in an event; "he had a religious experience"; "he recalled the experience vividly"
- experience: an event as apprehended; "a surprising experience"; "that painful experience certainly got our attention"

All three senses are appropriate meanings of the equivalent Arabic word *tjrbp* (تجربة) but they do not include the *SCIENTIFIC EXPERIMENT* sense covered by the Arabic word.

From the above points, we find that 63.9% of the ambiguous Arabic word types evaluated are conceptually equivalent to their ambiguous English translations. This finding is consistent with the observation of EuroWordNet builders. Vossen, Peters, and Gonzalo (1999) find that approximately 44-55% of ambiguous words in Spanish, Dutch and Italian have relatively high overlaps in concept and the sense packaging of polysemous words (Vossen et al., 1999). 29.3% of the ambiguous Arabic words correspond to specific senses of their English translations and 6.7% of the Arabic words are more generic than their English correspondents.

5 Acknowledgements

I would like to thank Philip Resnik and Daniel Jurafsky for their insightful comments. I would like to thank two annonymous reviewers for their detailed comments. This work is supported, in part, by NSF Award #IIS-0325646.

6 Conclusions

We presented, SALAAM, a method for bootstrapping the sense disambiguation process for Arabic texts using an existing English sense inventory leveraging translational correspondence between Arabic and English. SALAAM achieves an absolute precision of 56.9% on the task for Arabic. Of the 673 correctly tagged English tokens for the SENSEVAL2 English All Words Task, approximately 90% of the Arabic data is deemed correctly tagged by 3 native speakers of Arabic. Therefore, SALAAM is validated as a very good first approach to Arabic WSD. Moreover, we perform a preliminary investigation with very promising results into the quality of the English sense inventory, WN1.7pre, as an approximation to an Arabic sense inventory.

References

Irina Chugur, Julio Gonzalo, and Felisa Verdejo. 2002. Polysemy and sense proximity in the senseval-2 test suite. In *Proceedings of Word Sense Diasmbiguation: Recent Successes and Future Directions*, University of Pennsylvania, Pennsylvania, July.

D. Cruse. 1986. *Lexical Semantics*. Cambridge University Press.

Mona Diab and Philip Resnik. 2002. Word sense tagging using parallel corpora. In *Proceedings of 40th ACL Conference*, Pennsylvania, USA.

Mona Diab. 2000. An unsupervised method for multilingual word sense tagging using parallel corpora: A preliminary investigation. In *SIGLEX2000: Word Senses and Multi-linguality*, Hong Kong, October.

Mona Diab. 2003. Word sense disambiguation within a multilingual framework. In *PhD Thesis*, University of Maryland, College Park.

Helge Dyvik. 1998. Translations as semantic mirrors.

Christiane Fellbaum, Martha Palmer, Hoa Trang Dang, Lauren Delfs, and Susanne Wolff. 2001. Manual and Automatic Semantic Annotation with WordNet. In *Proceedings of the NAACL Workshop on WordNet and Other Lexical Resources: Applications, Customizations*, Carnegie Mellon University, Pittsburg, PA.

Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. MIT Press. http://www.cogsci.princeton.edu/˜wn [2000, September 7].

Nancy Ide. 2000. Cross-lingual sense discrimination: Can it work? *Computers and the Humanities*, 34:223–34.

M.R. Quillian. 1968. Semantic Memory. In M. Minsky, editor, *Semantic Information Processing*. The MIT Press, Cambridge, MA.

Philip Resnik and David Yarowsky. 1998. Distinguishing Systems and Distinguishing Senses: New Evaluation Methods for Word Sense Disambiguation. *Natural Language Engineering*, 1(1):1–25.

Philip Resnik. 1999. Disambiguating Noun Groupings with Respect to WordNet Senses. In S. Armstrong, K. Church, P. Isabelle, S. Manzi, E. Tzoukermann, and D. Yarowsky, editors, *Natural Language Processing Using Very Large Corpora*, pages 77–98. Kluwer Academic, Dordrecht.

P. Vossen, W. Peters, and J. Gonzalo. 1999. Towards a Universal Index of Meaning. pages 1–24.

Louise Guthrie Wim Peters and Yorick Wilks. 2001. Cross-linguistic discovery of semantic regularity.