A Large Coverage Verb Taxonomy For Arabic

Jaouad Mousser
Universität Konstanz
Fachbereich Sprachwissenschaft
78457 Konstanz, Germany
Jaouad.Mousser@uni.konstanz.de

Abstract
In this article I present a lexicon for Arabic verbs which exploits Levin’s verb-classes (Levin, 1993) and the basic development procedure used by (Schuler, 2005). The verb lexicon in its current state has 173 classes which contain 4392 verbs and 498 frames providing information about verb root, the deverbal form of the verb, the participle, thematic roles, subcategorisation frames and syntactic and semantic descriptions of each verb. The taxonomy is available in XML format. It can be ported to MYSQL, YAML or JSON and accessed either in Arabic characters or in the Buckwalter transliteration.

1. Introduction
Class-based approaches to verb lexicon that provide key elements information about the syntax and semantic enjoy a popularity in a variety of natural language tasks including machine translation, document classification (Klavans and Kan, 1998), semantic role labeling (Gildea and Jurafsky, 2002), sense disambiguation (Dang, 2004), and subcategorisation acquisition (Korhonen and Briscoe, 2004). A prominent large scale lexical resource that uses the notion of verb classes is Verbnet for English. Despite the universality postulate asserting that verb-classes can be identified across languages (Jackendoff, 1990), only few languages dispose of standardized collections of verb-classes or a verb lexicon.

Most theoretical work on verb classes for Arabic does not result in a verb lexicon, not least because the approaches are still controversial. The contentious issue is the correlation between morphological basic forms (root, stem, etymon) and the meaning as well as the organisation in the lexicon. The common approach to verb classes in Arabic is a root based approach, which claims that the core meaning of a verb is carried by a root consisting of 2-4 consonants in a specific order which guarantee the semantic relation to other verbs in the lexicon (McCarthy, 1981). This approach is indeed useful for grouping entries related to the same root and derivation is hard to find. An alternative model presented in (Ehret, 1995) and especially in (Bohas, 1991) claims that verb meaning resides in a 2-consonants root-like form called etymon $\epsilon$. An etymon is a combination of two consonants presented as a matrix $\mu$ of primitive phonetical features (dental, labial ..) that are assumed to have a semic value shared by all words built on the basis of the etymon. For example the feature matrix in (1) produces etymons like $f\bar{a}$ or $f\bar{h}$ which have a semic nucleus described in Table 1. The verbs built on the basis of this etymons are listed in Table 2.

\[
(1) \quad \mu = \{ [+\text{consonantal}] \}
\]

\[
\{ [+\text{consonantal}] \} \\
\{ [-labbial] \} \\
\{ [+\text{continuant}] \}
\]

Table 1: Semic nucleaus of the features matrix $\mu$

| Semic | Meaning |
|------|---------|
| $f\bar{a}$ | “to blow on something” |
| $f\bar{h}$ | “to hiss, to wheeze while sleeping” |
| $f\bar{h}a$ | “to wheeze while sleeping” |
| $f\bar{a}$ | “to perfume meals with aromas” |
| $f\bar{a}$ | “to blow (said of a warm wind)” |

Despite the fact that the two models (root model and etymon model) cannot be generalized over the whole lexicon, they make no attempt to associate the semantic meaning of verbs with their syntactic structures.

A more NLP oriented lexicon providing semantic description of verbs is Arabic Wordnet (Elkateb et al., 2006). Despite the fact that the lexicon contains only 1400 verbs (a school conjugation manual of Arabic contains 10000 verbs), it inherits all shortcomings of the English Wordnet pointed out by (Kipper et al., 2000), such as listing to many fine-grained sense distinctions and the lack of explicit syntactic information like predicate-argument structures.

(Diab and Snider, 2006) attempted to apply clustering techniques to induce verb classes for Arabic from a corpus using features like subcategorization frames, verb patterns, subject animacy, LSA semantic vectors. They reported that their clustering method performed well with respect to a gold standard produced by a noisy translation of English verbs from the Levin classes. However no information about the number or the natures of classes produced by the clustering was reported.

On the other hand, the English verb lexicon Verbnet provides the advantage of treating syntactic and semantic properties of verbs in a compact way by profiting from Levin’s detailed work on English verb classes. In this lexicon,
verbs are grouped into classes according to shared syntactic and semantic properties. Other verbs are grouped into subclasses according to restrictions on the thematic participants or to semantic predicates separating them from the prototypical class member. The main assumption is that diathesis alternations are meaning preserving (Levin, 1993). This assumption served as basis for the building of Verbnet for English (Schuler, 2005).

Computational linguists may encounter many difficulties when they try to generalize the claim to other languages in perspective of building similar lexicon as in English. This article presents some of the issues encountered during the building process of a Verbnet for Arabic.

2. Basic approach

The building of a large coverage verb lexicon for Arabic is a challenging task. Unlike the developers of the English Verbnet, we do not dispose of a collection of verb classes like the one provided by Levin (Levin, 1993). The main work that has to be done essentially is of the collection and classification of verbs. In this context two approaches can be used given the available data:

1. the first approach departs from a given set of classes trying to give them a computer readable form and to populate them automatically or semi automatically with members sharing their properties. This approach uses mostly the results of years of theoretical works such those compiled in (Levin, 1993).

2. the second approach works on corpora using algorithms for supervised or unsupervised clustering to automatically induce classes of verbs on the basis of some shared syntactic and grammatical features (Stevensons and Joanis, 2002). This approach is grounded in theoretical considerations assuming that verbs with similar meaning components can be detected according to surface properties like the ability to alternate in the same syntactic structures. In this work I opted for the first approach and used Levin’s verb classes and some of the novel classes of (Korhonen and Briscoe, 2004) assuming that with some adaptations to its properties (syntactic structure, morphological structure, etc.) these classes can also be used for Arabic.

Each Verbnet class is represented by the set of properties carried by its prototypical verb. The relevant semantic information for class mapping are:

1. the kind and number of thematic roles the prototypical verb selects.
2. the selectional restrictions on its participants.
3. the core semantic meaning of the prototypical verb.

The information related to last two points have been made possible by the compositional semantics added to each verb frame in Verbnet (Dang et al., 2000). Table 3 shows the relevant information for mapping the hit class in Verbnet to Arabic.

Members of each Verbnet class are translated to Arabic and expanded according to relations like synonymy, hyponymy, hyperonymy, etc. using different dictionaries including dictionaries of classical Arabic like lissan Al Arab or Al qaamus al muhit. The last ones turns out to be useful for clarifying the etymological background of verbs, which, some times, is a nececcary step for inducing root forms from infinitival forms of verbs. A prototypical verb is selected under the set of verbs produced in the translation and the expansion and put into all its possible frames. Each frame is described by an example sentence, a syntactic structure reflecting the subcategorization information of the verb and a semantic structure including information about its temporal aspects. The rest of the verbs are then added to the new class according to whether they shares properties of the prototype verb. Other are added to subclasses when they diverge from the main class in some not central points.

For example trying to apply this building process on the class manner-speaking, which in English includes verbs like whisper, babble and cluck results in two groups of verbs as shown in Table 4. The prototype verb waswasa of the fist group attests the most of the meaning aspects of the Verbnet class manner-speaking as the three thematic roles agent, topic and recipient and the main semantic predicates cause and transfer_info which can be composed to: cause(agent, transfer_info(topic, recipient)). However the second group consists of verbs which do not share these properties, since they lack the thematic roles topic and recipient, which are two important roles for the meaning transfer_info.

| Group 1                          | Groupe 2                        |
|---------------------------------|---------------------------------|
| unwhad ‘reveal’                  | tamtama ‘mumble’                |
| hashasa ‘swish’                  | walwala ‘make a howl’           |
| hassa ‘murmur’                   | ga’gâ ‘growl’                   |
| waswasa ‘murmur’                 | damdama ‘burr’                  |
| wašwaš ‘whisper’                 | etc.                            |
| hamasa ‘whisper’                 |                                |

Table 4: Two verb groups resulting from translating Verbnet class manner-speaking

Diathesis alternations are indeed a good indicator for mapping English verb classes into Arabic verb classes when they are detected, but they are not a condition since they are not expected to be always similar in two different languages. So for example the conative alternation in (2) may be crucial in distinguishing some classes in English, but it plays no role in Arabic since it simply does not exist.

(2) Conative Alternation

a. Paula hit the fence.

b. Paula hit at the fence.

The same can be said about the temporal information of verbs (they are also included in the semantic descriptions of frames in Verbnet). Whereas the temporal aspect plays a
crucial role in characterizing verbs of the class *hit*, it plays no role in building a class like *smell_emission* in Arabic. Verbs of the class *smell_emission* in English like *reek, smell, stink* are static as the semantic description in (3) shows, that is, they do not entail a start or an end of the event, whereas in Arabic the same verbs express a change in the smell property of the theme with a start point and an end point (example (4)).

(3) emit(during(E), Theme, Odor)

(4) not(emit(start(E), Theme, Odor)) ∧ emit(end(E), Theme, Odor)

In future work corpora and information extraction algorithms will be used to

a. examine the affiliation of verbs to the classes they was manually sorted into
b. extend the available classes and their members
c. enrich the available classes with more frames and semantic information.

3. Verb entries

The main verb entry in the Arabic verb lexicon is a diacritized infinitival form which corresponds to the perfective of the third person masculine singular, as it is common for referring to verbs in Arabic.

Every verb entry is a node that contains four child nodes representing the verb itself, its root, a deverbal noun, and a participle. The last two derived forms can have multiple entries since one verb can bear more than one deverbal nouns or participle.

The containment relation between the node element and its children does not make any claim about the organization of these elements in the lexicon. It is just a way to organize an entry and the elements, which are morphologically and semantically related to it.

The motivation behind adding the deverbal noun and the participle is the fact that they inherit all the semantic and syntactic properties of the verb they are derived from. The relation between verb and the two derived forms is by no way a one-to-one relation, but at this stage in which polysemy arises.

Classes of the Arabic verb lexicon can be accessed not only through the main verb entry but also through the deverbal and the participle or the root.

Polysemy in the verb level is resolved through cross listing entries in different classes. Hence the verb *ğara∫a*, which has the meaning of ‘to run’, ‘to meander’ or ‘to occur’ is listed in four classes: the class *run, meander, occurrence* and finally the class *mode of being*.

4. Diathesis Alternations

A preliminary study about diathesis alternations in Arabic was required to determine whether the alternation building the basis of Levin’s classes are also available in Arabic. Table 5 shows some of the alternations.1

Surprisingly 65% of all alternations available in English are also available in Arabic. The specificity of Arabic alternations lay in the fact that they often engage morphological operations. This is the case with transitivity alternations and causative alternations like the *middle alternation* and the causative/Inchoative alternation etc.

The crucial point here is the distribution of this alternation in classes and their relevance for the class building process. Whereas the *spread/load* alternation is relevant for building the *spray class*, the *congative object contraction* alternation has in constrast a marginal significance since almost all verbs in Arabic (transitive as well as intransitive) can appear in this construction (example (5)).

(5) a. ə çıhbəhəa əcba’n
   *I love her deeply.*
   *1sg-love-3sg-Fem-Acc INDEF-love-Acc əmiyq’a’n*  .
   deep-Mas-Acc
   ‘I love her deeply.’

b. ə ǧara∫a alwaladu uzbrahu
   *The child cut her thumb deeply.*
   cut-Mas-SG DEF-child-NOM thumb-his
   əgürh’a əmiyq’d’a’n .
   INDEF-cut-acc deep-Mas-Acc
   ‘The child cut her thumb deeply.’

c. ə tasarrufa ta∫arrufan
   *He behaves irresponsibly.*
   beahave-Mas-SG INDEF-behavior-acc əlməs-wənla’n .
   deep-Mas-Acc
   ‘He behaves irresponsibly.’

Since the diathesis alternations are assumed to be meaning preserving, the question to ask here is: can we still

---

1A more comprehensive is to find in: http://ling.uni-konstanz.de/pages/home/mousser/alternations

| Thematic roles and Restrictions | Semantic predicates |
|---------------------------------|---------------------|
| Agent(+int,control)            | Cause(agent, E)     |
| Patient(+solid)                | Contact(during(E), patient) |
| Instrument(+solid)             | degradation_material_integrity(result(E), Patient) |
|                                | pysical_form(result(E), From, Patient) |
The morphology plays a significant role in the expression of event structures of verbs in Arabic. Since morphological changes of derived verbs technically produce new lexical entries, a decision has to be made about whether derived verbs (and derived verb classes) still belong to the original class and whether they should therefore be distinguished by building a separate subclass inside the original class—since they share most of their properties—, or they should build a new class—since they are lexically autonomous entries?

I opted for a mixed approach which can be summarized as follows:

- If there is a class which shares exactly the same properties as the derived verbs, the derived verbs are transported to this class.
- If the derived verbs do not fit any existing class and the effect of the derivation is only a valency changing effect, a sibling class is created and the verbs linked to the verbs of the original class.
- If the derived verbs correspond to one of the existing classes, but adds additional semantic predicates separating them form the meaning of the original class, the verbs build a new class.

6. Thematic roles

I use the same set of thematic roles used in the English Verbnet. It consists of 23 thematic roles mapping verb arguments for all classes. It includes commonly used roles like agent, patient, theme and specific roles like patient2, theme1, theme2. Selectional restrictions such as concrete, abstract, location, state etc. are applied on thematic roles in order to get finer underspecifications. Some restrictions split into more precise restrictions. For example the restriction location splits into region for cases like min taht aţlīf wilah 'from under the table' and place for cases like fy a-lriibaţ in Rabat'. Language specific restrictions like dual

| Alternation                              | Occurrence | Example                                                                 | Translation |
|------------------------------------------|------------|------------------------------------------------------------------------|-------------|
| Middle + causative alternation           | Prefix+    | - qatuwa ʿālfalābūhū ʿālsīgharāh                                      | - The farmer cuts the tree |
|                                          |            | - ʿinqaṭarı ʿālsīgařarāh bīsīhuwalh                                     | - The farmer cuts the tree |
| Induced action alternation               | Duplication+ | - qaftaḍ ʿalmaṭ ʿālḥīsaīna                                           | - Salma jumped the horse |
|                                          |            | - qaftaḍ ʿalḥīsaīnu                                                   | - The horse jumped |
| Unspecified object alternation           | yes        | - ṭakala ʿīṣā̀mūn aḏl̩kaḥ                                          | - Issam ate the cake |
|                                          |            | - ṭakala ʿīṣā̀mūn                                                   | - Issam ate |
| Substance/source alternation              | Prefix+    | - ṭanbāṯt ʿaılıḥaraaṛatū mīna ʿālṣ̩mānsi                             | - Heat radiates from the sun |
|                                          |            | - ṭalvaṭt ʿaılıḥamsu ʿaılıḥaraaṛatū                                   | - The sun radiate heat |
| Undescribed body-part object alternation | yes        | - ṭamṣīṭū ʿsr̩āḍān ᵈ̱sr̩āḥāḥ                                 | - Suwad brushed her hair |
|                                          |            | - ṭamṣīṭū ᵈ̱sr̩āḍān                                                   | - Suwad brushes |
| Undescribed reflexive object alternation | no         |                                                                         |             |
| Undescribed reciprocal object alternation| yes        | - ṭilaṭq̩ū ʿalṣīmūn ʿīṣ̩āmūn                                          | - Salim met Issam |
|                                          |            | - ʿalṣīmūn waw ʿīṣ̩āmūn ṭilaṭq̩ūyāḥ                                | - Salim and Issam (both) met |
| PRO-arb object alternation               | yes        | - ʿyāḍīmū ʿalṭ̩īḻmū ʿalmaṣ̩aḥāḥidīṽān                             | - The movie shocks the public |
|                                          |            | - ʿalṭ̩īḻmū ʿyāḍīmū                                               | - The movie shocks |
| Characteristic property of instrument alternation | yes | - qataṭa ʿalḥakaba ʿbīaḥāṁīnaari                                    | - I cut the piece of wood with the saw |
|                                          |            | - ʿaḥāṁīnaari ʿyaq̩ṭ̱a                                               | - The saw cuts |
| Benefitative instrument alternation       | no         |                                                                         |             |
| Preposition drop alternation              | yes        | - ṭasallaq̩a ʿalṣīmūn ṣālāḏ ʿalḡ̱abālī                           | - Salim climbed up the mountain |
|                                          |            | - ṭasallaq̩a ʿalṣīmūn aḏl̩g̱abala                                   | - Salim climbed the mountain |
| Dative alternation                       | yes        | - bawā ʿamīṇyūn ʿalṣ̩yāṭaṭū ṭalām̱iyy̱ū ṭalām̱iyy̱ū            | - Jamil sold a car to Hamid |
|                                          |            | - bawā ʿamīṇyūn ṭalām̱iyyūn ʿalṣ̩yāṭaṭū                             | - Jamil sold Hamid a car |
| Spray/load alternation                   | yes        | - ṭaṣ̱aṭṭū ʿaḥṣ̱ibāq̱ṯa ʿalāḏ ʿaḥṣ̱ibāḏ̱ārī                          | - I sprayed paint on the wall |
|                                          |            | - ṭaṣ̱aṭṭū ʿaḥṣ̱ibāḏ̱ārī bī ʿaḥṣ̱ibāq̱ṯa                          | - I sprayed the wall with paint |
| Clear alternation (intr.)                | yes        | - ʿīq̱aṭarı ʿaḥṣ̱uḥāḥū mīna ʿaḥṣ̱amaṭāĩ                          | - Clouds cleared from the sky |
|                                          |            | - ʿīq̱aṭarı ʿaḥṣ̱uḥāḥū                                          | - The sky cleared |

Table 5: Occurrence of some Alternations in Arabic
are introduced for cases like traṣṣaṣa ʿalḏārainqu ‘The neighbours (both) correspond (with each other)’.  Disjunction of restrictions are expressed through the boolean operator or such as in orf(+dual,+plural) such as the the case in reciprocal verbs.

7. Frame Description and syntactic structures

Arabic Verbnet uses descriptive constructs that allow to call each frame precisely and to distinguish it from other frames. They consist of a primary description which is more general and reflects the surface syntactic structure of the frame and a secondary description which is more specific and reflects the kind of alternation used in the frame. For example the description in example (6) describes the frame of a sentence whose main verb subcategorizes three arguments and where the third argument has the thematic role beneficiary. The secondary description (the part after the two slashes) specifies properties of the alternation (preposition used, restrictions, thematic roles pointed out etc.)

(6)  NP V NP PP/beneficiary // NP-PPfor-PP

The frame descriptions are adapted to Arabic such that:

- The prepositions of obliques correspond to the preposition inventory of Arabic, for example NP-PP//EalaY-PP
- English specific structures like infinite clauses (described with ING-POSSING) are omitted.

As in Verbnet we use a LTAG (Lexicalized Tree-Adjoining Grammars) as framework to describe the surface syntactic structures of each frames. LTAGs consists of a finite set of initial and auxiliary trees and two operations to combine them namely Adjunction and substitution. Every tree is associated with a lexical item and sets directly specific syntactic constraints, such as selectional restrictions for the plural. LTAGs tree are adapted to Arabic such that:

They reflect the syntactic structure of Arabic which is generally VSO

Arabic specific frames like V NP bi COMP or NP V mataY S are described.

8. Semantic Structure

As in Verbnet for English, a compositional semantics is employed in relation to the LTAGs tree to describe regu-
lar senses of verbs. The compositional semantics has the advantage of allowing verb sense extension or modification resulting from auxiliary trees (adjuncts). Semantic predicates such as cause, emotional state, motion, made_of are associated with each tree to describe the key component meaning of the verb as well as the relation between participant and event structures. Each event $E$ is presented as a tripartite structure according to (Moens and Steedman, 1988), which describe the temporal aspect of the event. The semantic predicates are associated with time functions specifying the part of time in which the event is true. This functions are: the begin of the event (start($E$)), the preparatory stage (during($E$)), the end stage (end($E$)), and the consequent (result($E$)). Example (7) shows the semantic structure of sentence (8). The predicates are: cause which has as argument the agent and the event $E$, b. and state, which take another predicate as argument namely result and use, which describe the use of an instrument to fulfill the action. The time functions are: endstate describing the state of the patient at the end of the action and (during($E$)) which means that the use of the instrument was during the event.

(7) cause(Agent, $E$), state(result($E$), Endstate, Patient), use(during($E$), Agent, Instrument)

(8) kasara 'alwaladu 'alzuzuţaţa biîlmitraqt. ‘The child broke the window with the hammer.’

9. Introducing sibling classes

In the effort of building a Verbnet for Arabic the problem of class overgeneration arises because of the morphologically overt realisation of many semantic predicates of verbs. To be able to alternate in some syntactical structures verbs in Arabic use productive morphological operations (affixation and prosodical stem mutation etc.). These operations have an impact on either the thematic arity of the verb Table 9 or of the semantic predicates building its meaning. The derived verbs diverge significantly from the base verbs, but they inherit their core semantic meaning. The divergence produce a new lexical entry and expels the derived verb from the class of the base verb. However the fact that the derived class contains a part of the morphologically marked diathesis alternations of the verbs make reconnecting the main classes with the derivative classes a reasonable task since it reflects the natural connection between verbs and their derived forms in the lexicon.

Those all classes morphologically related to each other and sharing the same meaning are linked bidirectionally to each other. In addition each class member (verb) is linked separately to the verb it is derived from. The linking establish no hierarchial relation between two classes (such as parent child relation). However only the classes resulting from a valancy changing operations are linked to their sibling classes since this operations are more regular (affect all classes with a particular semantic meaning) and complete (affect all members of the concerned class).

Linking morphologically related classes and their members helps taking all morphological manifestations of alternations under account. It should help showing and predicting the relation between morphology and semantics and how morphology contributes to express meaning aspects of verbs.

10. New classes

By applying Levin’s class inventory on Arabic it turns out that many classes does not exist in Arabic like the class debone since it has morphological implication in English (the negation prefix de-) due to the Latin origin of the verbs. Other exist with a small amount of members like the class coil or vehicle (one member). Some classes are integrated into other classes since the properties which may make them to autonomous classes does not exist in Arabic like the class gobble whose members are integrated in the class devour since Arabic unlike English does not make a distinction between gobble-verbs and devour-verbs.

In addition, Levin’s class inventory does not describe event structures of some Arabic verbs. Verbs like sīhada which can be paraphrased as ‘to make a statement of belief’ does not match any class of communication in Levin’s collection. They lexically incorporate the proposition or formula used in the statement using acronymy. Those a new class with the name of the prototype verb sāhada is created to contain all verbs which share the same semantic meaning and show the same syntactic behaviour. Similarly verbs like saraa ‘to walk in the night’, which belong to the verbs of motion and inherits all properties of the class run build a new class since they add a new predicate describing the part of day in which the walking event takes place.

11. Mapping to other lexical resources

The verb lexicon was mapped to Arabic Wordnet (Elkateb et al., 2006), a version of Wordnet for Arabic developed
| Operation      | Basis Verb       | Derived verb                                      |
|----------------|------------------|--------------------------------------------------|
| Causativation  | waqa ‘fell’      | → awqa ‘Cause X to fall’                          |
| Decausativation| salla¯a ‘Cause X to amuse’ | → tasalla¯a ‘amuse’                                |
| Reciprocalisation | annaqa ‘hugge’ | → tava¯dnaqa ‘hugge each other’                    |
| Reflexivization| gahhaza ‘equip’  | → ta¯gahhaza ‘equip his/herself’                   |

Table 9: Thematic arity changing operations

by a group of researchers from different universities following the development process of Princeton Wordnet and Euro Wordnet and using a suggested upper merged ontology that links the lexicon to English Wordnet. Associating Arabic Verbnet with Arabic Wordnet synsets may provide the verb lexicon with richer semantic descriptions, but unfortunately Arabic Wordnet covers only a small part of the most frequent verbs in Arabic (about 1400 verbs).

12. Conclusion and future work

I have presented a class-based lexicon for Arabic using Levin’s classes and the building procedure of English Verbnet which provides the advantage of associating syntax and semantic in describing verbs. Transferring class information from one language to another requires adapting this information to the properties of the target language. This reveals itself be a possible task due to the compositional character of the syntactic and semantic descriptions provided in English Verbnet, which allow to remove or add predicates describing language specific meaning aspects flexibly. Diathesis alternations engage often overt morphological operations. The risk of losing the connection between main classes and derived classes while building a lexicon is countered by linking them and their verb members together. Evaluating the coverage of the actual lexicon and expanding its classes and members is part of the next step. I expect to reach a similar coverage and accuracy as provided by Verbnet for English.

13. References

G. Bohas. 1991. Le pco, la composition des racines et les conventions d’association. Bulletin des Etudes Orientales, XLIII.
H. T. Dang, K. Kipper, and M. Palmer. 2000. Integrating compositional semantics into a verb lexicon. In GOLING-2000 Eighteenth International Conference on Computational Linguistics, Saarbrücken.
T. H. Dang. 2004. Investigations into the Role of lexical Semantics in Word Sense Disambiguation. Ph.D. thesis, CIS, University of Pennsylvania.
M. Diab and N. Snider. 2006. Unsupervised induction of modern standard arabic verb classes. In NAACL.
C. Ehret. 1995. Reconstructing Proto-Afroasiatic (Proto-Afrosian) Vowels, Tone, Consonant and Vocabulary, University of Calefornia Press, Berkley and Loas Angeles.
S. Elkateb, W. Black, P. Vossen, and C. Fellbaum. 2006. Arabic wordnet and the challenges arabic language. In International Conference at The British Computer Society (BCS), London.
G. Gildea and D. Jurafsky. 2002. Automatic labeling if semantic roles. Computational Linguistics.
R. Jackendoff. 1990. Semantic Structures. MIT Press, Cambridge, MA.
K. Kipper, T. H. Dang, and M. Palmer. 2000. Class-bases construction of a verb lexicon. In AAAI-2000 Seventeenth National Conference on Artificial Intelligence, Austin.
J. Klavans and M. Y. Kan. 1998. Role of verbs in document analysis. In Proceedings of the 17th International Conference on Computational Linguistics (COLING/ACL).
A. Korhonen and T. Briscoe. 2004. Extended lexical-semantic classification of english verbs. In The HLT/NACCL workshop on computational lexical semantics.
B. Levin. 1993. English Verb Classes and Alternations. A Preliminary Investigation. The University of Chicago Press, Chicago and London.
J. McCarthy. 1981. A prosodic theory of nonconcatenative morphology. Linguistic Inquiry.
M. Moens and M. Steedman. 1988. Temporal ontology and temporal reference. Computational Linguistics, 14.
K. Schuler. 2005. VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon. Ph.D. thesis, University of Pennsylvania.
S. Stevensons and E. Joanis. 2002. Semi-supervised verb class discovery using noisy features. In The conference On Computational Natural Language Learning.