Enriching Lexical Transfer With Cross-Linguistic Semantic Features
or
How to Do Interlingua without Interlingua

Alexis Nasr, Owen Rambow
CoGenTex, Inc.,
840 Hanshaw Road, Suite 11
Ithaca, NY 14850-1589
nasr/owen@cogentex.com

Martha Palmer, Joseph Rosenzweig
Department of Computer and Information Science
University of Pennsylvania
Philadelphia PA 19104-6389
mpalmer/josephr@linc.cis.upenn.edu

1 Introduction

In this paper, we propose an alternative to interlingua which can capture the analyses and generalizations that interlinguas can express, but which uses cross-linguistic semantic features rather than a separate level of representation. This alternative we call *lexico-structural transfer*. Lexico-structural transfer relies on the expressive power of a *lexicalized* syntactic representation (or "lexicalized grammar" for short). In a lexicalized grammar, lexemes are associated with syntactic structure; in the transfer lexicon, we do not simply relate words (or context-free rewrite rules) from one language to words (or context-free rewrite rules) from another language. Instead, we relate lexemes along with relevant syntactic structure (essentially, their syntactic projection along with syntactic and lexical-semantic features). Several different lexicalized grammar formalisms have been proposed in the past, including notably Tree Adjoining Grammar [4], Lexical-Functional Grammar [5] and various dependency grammars. We will present our work using a transfer formalism based on a dependency grammar, namely Mel'čuk's Meaning Text Theory (MTT) [7], specifically the "Deep Syntactic Level". This level of representation is similar in crucial respects to the derivation structures of TAG [11] and to the f-structure of LFG. There are two main reasons why we may want to investigate an alternative to the use of an interlingua:

- It is theoretically interesting to understand the exact contribution of an interlingua. More precisely, if we can represent interlingua-based analyses without interlingua but using certain key elements of the interlingua analysis, then we have made important strides towards understanding why interlingua-based analyses are successful.
- There are practical advantages to tying machine translation to surface-oriented representations as much as possible: the availability of large bilingual corpora has made the exploitation of stochastic approaches a crucial element in the practical success of MT, and such approaches are by nature oriented towards the surface form.

This paper is structured as follows. In Section 2, we introduce some the standard interlingua-based analysis for a well-known case of "structural divergence", a difficult MT challenges [2]. In Section 3, we present our MT system and its lexico-structural transfer formalism in particular. In Section 4, we discuss the semantic analysis that underlies our approach. In Section 5, we present a simple transposition of the interlingua approach to lexico-structural transfer. In Section 6, we show how we can eliminate the most unmotivated features of this analysis by using lexical functions, a language-internal device for relating lexemes. We conclude in Section 7.

2 Structural Divergences

A classic problem in machine translation is the translation of motion verbs from English to French. In English the manner of motion can be incorporated into the matrix verb, with the direction of the motion being adjoined on by a prepositional phrase, as in *John swam across the lake.* In many cases, this is not allowed in French. In French, the direction becomes incorporated into the matrix verb, and the manner is adjoined on as an adverbial or a prepositional phrase, as in *Jean a traversé le lac à la nage* - John crossed the lake by swimming. This type of translation mismatch has been termed *structural divergence*, and is often cited as strong evidence for the advantages of an interlingual representation [1, 2].

Dorr uses Jackendoff's Lexical Conceptual Structures (LCS) as an interlingua which handles these types of divergence problems. Based on the assumption that motion and manner of motion are conflated in a matrix verb like *swim*, the use of LCS allows Dorr to tease apart the three separate concepts of motion, direction, and manner of motion in sentences such as *John swam across the lake.* Each one is represented separately in the interlingua representation, as GO, PATH and MANNER, respectively.

$[_{\text{Event}}$ CAUSE ($[_{\text{Person}}$ JOHN],
$[_{\text{Event}}$ GO ($[_{\text{Person}}$ JOHN],
$[_{\text{Path}}$ ACROSS ($[_{\text{Position}}$ AT ($[_{\text{Thing}}$ JOHN], $[_{\text{Thing}}$ LAKE])])])]
$[_{\text{Manner}}$ SWIMMINGLY])]

Then the French linking rules can allow these components to be recombined in a different order, this time incorporating GO and DIRECTION into *traverser* and leaving MANNER by itself as an adjunct.

As is well known, not all manner-of-motion verbs show this divergence. For example, *Marie-France swam to Calais* is probably best translated literally, namely as *Marie-France a nagé jusqu'à Calais*. In the interlingua-based approach, the proper choice of target language realization falls entirely on the generation component, which must be able to match the meaning components from the LCS against the lexicon and choose lexemes which together cover the entire LCS. (This task can be computationally costly.)

While an interlingua analysis can perspicuously handle such translations, this substantial difference in how the languages incorporate information is troublesome for transfer-based systems. This is because there are many possible direction prepositions that can be adjoined on, and even more different manners of motion. One might at first think that each possible unique combination of manner and direction must be spelled out explicitly and associated with its appropriate French equivalent.

In the following sections, we will show how we can use the key insight from the interlingua analysis — namely, the decomposition of meaning into components — and incorporate into a lexical transfer approach.

3 A Transfer Formalism Based on the Meaning Text Theory

In this section, we describe the tranfer formalism we have developed as part of the TransLex MT project [12].

3.1 The Deep Syntactic Level

The Deep Syntactic Structure (DSyntS) specifies the syntactic organization of a sentence in terms of a dependency tree. Such trees are composed of nodes labeled by generalized lexemes and directed arcs (dependencies) labeled with Deep Syntactic Relations. A generalized lexeme is a full lexeme, a multilexemic unit (idiom), or a lexical function (see Section 6 for details). Semantically empty lexemes, such as governed prepositions or auxiliaries are not represented at this level. Generalized lexemes can be enriched with meaning bearing morphological features such as number in nouns and tense and/or aspect in verbs. Syntactically conditioned morphological features, induced by syntactic rules (such as agreement rules), are not represented in the Deep Syntactic Structure.

The Deep Syntactic Relations labeling the dependencies are supposed to be language independent. Each stands for a family of specific syntactic constructions of a particular language. There are different arc labels for the different arguments ('I' for subject, 'II' for direct object, 'III' for indirect object, and so on); label 'ATTR' covers all adjuncts. Two additional labels handle coordination and parentheticals and related constructions.

The Deep Syntactic Structure is closely related to a specialized lexicon, the Explanatory Combinatorial Dictionary [8] which, among other information, defines the subcategorization frames of predicative lexemes, as well as the prepositions introducing the actants of a predicate and the lexical functions linking the different lexical entries.

For the sake of illustration, a Deep Syntactic Structure, corresponding to the sentence *Temperature will remain seasonable in the remainder of the province*, is represented below:

```
remain [class:verb tense:future]
   (II seasonable  [class:adjective]
    I temperature  [class:common_noun number:sg det:def]
    ATTR in  [class:preposition]
       (II remainder  [class:common_noun number:sg det:def]
          (II province  [class:common_noun number:sg det:def])))
```

Through the use of generalized lexemes and deep syntactic relations, the abstraction achieved at the Deep Syntactic level concerns lexical as well as syntactic aspects of utterances. In the framework of the MTT, this level of abstraction allows generation, from a single Deep Syntactic Structure, of several semantically equivalent Deep Syntactic Structures which will themselves give rise to a large number of paraphrases. We use the same characteristics of the Deep Syntactic level in our lexicalized transfer formalism which elates semantically equivalent lexemes or groups of lexemes of different languages. The level of abstraction at which these relations are defined releases the author of the transfer lexicon from monolingual constraints (linear order, agreement, function words) which are not relevant for transfer, leading to a reduced number of transfer rules.

3.2 The transfer formalism

The transfer formalism relates Deep Syntactic Subtrees, anchored by lexemes of different languages. Two related subtrees respectively anchored by the two lexemes l_1 and l_2 represent a context in which l_1 is translated by l_2. Transfer is carried out by replacement of a subtree by another to which it is linked. In the simplest cases, the subtrees are reduced to a single node: the root of the tree. The following example shows a relation between the three lexemes *manger* (French), *eat* and *akl* (Arabic).

```
@TRANS_CORR @FR manger
            @EN eat
            @AR akl
```

The simplicity of this relation is due to the fact that these three verbs share a common subcategorization frame. When applying such a rule on a DSyntS, the nodes that are not represented in the rule will remain unchanged after application of the rule. This is the reason why the arguments of the verbs have not been represented in the rule.

When the translation of a lexeme (or group of lexemes) results in a syntactically divergent structure in the target language, this divergence is represented in the transfer lexicon by subtrees anchored by the lexemes. The following correspondence exhibits a case of argument shifting: the first actant of the English verb *like* (represented by the variable X) becomes the second actant of the French verb *plaire* and the English second actant becomes the French first actant. The preposition *à* for the second actant is introduced during French generation by reference to the monolingual French lexicon.

```
@TRANS_CORR @EN like (I  X
                      II Y)
            @FR plaire (I   Y
                        II  X)
```

In accordance with the complexity of the structural changes induced, the subtrees can get larger and can involve several lexical items. The translation of a support verb structure in one language into a simple verb in another language, for example, necessitates the introduction of two lexical items (the support verb and the predicative noun) in a subtree:

```
@TRANS_CORR @FR se_suicider
            @EN commit (II suicide)
```

Other constraints can be introduced in the subtrees by means of features on the nodes. The simplest case of feature introduction is represented by constraining the category of lexemes. The verb *rain*, in English will be translated by *pleuvoir* in French while the noun *rain* is translated by *pluie*:

```
@TRANS_CORR @EN rain [cat:verb]
            @FR pleuvoir

@TRANS_CORR @EN rain [cat:noun]
            @FR pluie
```

It can be noted that the entries in the transfer lexicon are not restricted to a single language pair and are also non directional. This is why we prefer to call a collection of such entries a Multi Linguistic Base rather than a transfer lexicon. Actual directed transfer lexicons, in a computationally efficient format, are automatically extracted from Multi Linguistic Bases.

4 Motion Verbs and Other Path-Denoting Expressions

One of the most difficult questions in lexical semantics is whether or not a prepositional phrase constitutes an argument to a verb or an adjunct. For instance, a verb's semantics may require that a path be specified, to complete the description of the eventuality the verb refers to. The prepositional phrase, if it denotes a path, may then be interpreted as the argument filling this slot. In other cases however, a path-denoting prepositional phrase is juxtaposed with a verb whose semantics does not seem to necessarily involve any paths, suggesting that the relationship between the semantic forms is more complicated. This situation is exemplified by sentences like:

(1) The bottle floated into the cave.
(2) The train roared through the station.

The verbs in these sentences can also occur without a syntactically related prepositional phrase, in a simple intransitive frame, where they do not seem to make reference to events involving a path:

($1'$) The bottle floated.
($2'$) The train roared.

This suggests that the semantic component contributed by the verbs in (1) and (2) does not contain a placeholder for a path, so it must be adjoined.

On the other hand, it has been argued that the verbs in ($1'$) and ($2'$) are in fact different from the verbs in (1) and (2), despite sharing the same vocable [13, 6]. That is, there are actually two verbs, *float* and *float′*, and two verbs *roar* and *roar′*. *Float* and *roar* take a prepositional phrase argument and have a semantic form with a slot that can be filled by a path. The other two verbs, *float′* and *roar′*, differ syntactically in taking no internal arguments, and semantically in not involving any paths.

If this *polysemy* argument is adopted, (1) and (2) are standard models of predicate-argument composition and no special mechanisms are required to derive their semantics. However, one would still like to account for the strong similarity between the semantics of *float* and *roar* on the one hand, and *float′* and *roar′* on the other.

Indeed, the eventualities described in (1) and (2) seem to entail those in ($1'$) and ($2'$), suggesting that the semantic forms of the intransitive verbs are in some way contained in the semantic forms of the verbs which require a prepositional argument. If this is so, the situations described in (1) and (2) may be viewed as *event complexes*, since they comprise other, possibly simpler, events.

Lexical rules, mapping from verbs like *float′* to verbs like *float*, have been proposed by Levin and Rappaport Hovav to formalize the relationship between the respective semantic forms, and capture the generalizations that would be lost if these verbs were treated as unrelated in the lexicon [6]. Formulating such rules though poses many of the same problems as those which arise if the verbs *float* and *roar* are taken to be *monosemous*, that is, the same in both their intransitive and prepositional-phrase frames.

Under both hypotheses, the event involving a path must be formally related to the event not involving a path. If the vocables are viewed as polysemous, a lexical rule must operate on the semantic form of the verb sense that does not have a path argument, and which denotes the event not involving a path, yielding a semantic form that does take a path argument, and that denotes the event which does involve a path. If on the other hand the vocables are viewed as monosemous, then there must be an operation which combines the semantic form of the verb, which by itself evokes an event not involving a path, with the semantics of the path-denoting phrase, producing a composed semantic form that now does evoke an event in which a path is involved.

We have adopted the latter hypothesis, and view the path prepositional phrase as carrying with it the information that it is part of a directed motion event. If the verb phrase it co-occurs with is not already typed as a directed-motion event, but does belong to a set of events such as manner of motion events or certain kinds of sound emission events, then the adjunction of a path-denoting prepositional phrase causes the verb phrase event type to be extended to include the notion of a directed motion event. For a longer discussion of this class of verbs and a detailed description of their treatment in Tree-Adjoining Grammars, see [9].

5 First Implementation: Semantic Features

By adopting the hypothesis discussed in the previous section, i.e., that in English the directed-motion event type is associated with the prepositional phrase that is adjoined on rather than with the matrix verb, we can tease apart in the syntax the same separate semantic components that are explicitly represented in an LCS interlingua. For *swim across the lake*, *swim* is associated purely with manner-of-motion, whereas path and direct-motion (GO) are both associated with *across the lake*. Each of these components is represented directly as a feature value for a semantic feature. Then in French, the verb *traverser* is marked as being both directed-motion and containing a path which is supplied by the verb object. The manner-of-motion in this case is supplied by the adjunct, *à la nage*. The end result is the same, namely that the event type of the sentence in both cases is a directed-motion event with a particular path and a particular manner-of-motion. If these components are all expressed as features on nodes, then the transfer lexicon can contain a general entry that abstracts away from individual lexical items, and simply enforces that these three features receive the same value for both the source language and the target language.

The interlingua analysis described in Section 2 is implemented in a straightforward manner in our formalism through the introduction of semantic features in the transfer correspondences as just discussed. These correspondences abstract away from the actual lexical items, and instead deal with classes of lexical items indicated by semantic features. In the pair (*swim across, traverser à la nage*), the presence of all of the lexical items in the rule would lead to a proliferation of the number of rules. Instead, we can represent the transfer correspondence simply as follows:

```
@TRANS_CORR @EN V1 [cat:verb manner:M ]
```

```
        (ATTR Y [cat:preposition path:P event:go]
           (II N))
@FR V2 [cat:verb path:P event:go]
        (II N
         ATTR Z [manner:M])
```

The three semantic features present in the correspondence (*manner, event, path*) are distributed differently in the French and the English subtrees. In the English one, the feature *manner* is supported by the verb while the features *path* and *event* are supported by the preposition. In the French part, the verb supports *path* and *event* and the adjunct (which could be an adverbial or a prepositional phrase, or an adjunct clause) handles *manner*. During the application of the rule to a source Deep Syntactic Structure, the two features *manner* and *path* are instantiated and their value copied into the target Deep Syntactic Structure.

The introduction of the lexeme *swim* in the English part of the correspondence will instantiate the manner-of-motion feature, M, with the value *swimmingly* and restrict the rule to this particular lexeme. The introduction of *run, float, jog, etc.* has the same effect, instantiating the manner-of-motion values accordingly. The introduction of the path (a locative prepositional phrase) will instantiate the path:P with the appropriate feature value, i.e., across for *across*, and so on.

The following source structure:

```
swim [cat:verb manner:swimmingly event:go ]
    (ATTR across [cat:preposition path:across]
      (II river [cat:noun]))
```

will give rise to the following target structure:

```
X [cat:verb path:across event:go]
  (II riviere [cat:noun]
   ATTR Z [cat:adverb manner:swimmingly])
```

The target Deep Syntactic Structure contains two uninstantiated lexemes. Their values are computed during generation by searching the lexicon for entries with the appropriate features (in a manner similar to generation from LCS).

However, this analysis implements the interlingua analysis so faithfully that it simply represents the interlingua in the two representations for source and target language. For example, the entry for *traverser* in the French monolingual dictionary must contain the features `path:across event:go`. While `event:go` is a reference to a *class* of verbs (motion verbs) which can be motivated within French by similar syntactic behavior (see Section 4), feature `path:across` does not refer to a meaningful class of verb but simply identifies a conceptual meaning element of the verb. The problem becomes even more apparent for the adjunct which is to express the manner in French, since here the set of possible values of the conceptual meaning feature `manner` is not, in principle, bounded.

It is therefore quite implausible to claim that all we need to do is to relate two independently motivated monolingual analyses. Instead, we need to first elaborate an interlingua, and then map the interlingua to the monolingual representations. While there may be some computational advantage to eliminating the interlingua as a separate level of representation, there is as yet no methodological advantage.

6 Second Implementation: Lexical Functions

In order to arrive at a solution in the lexical transfer approach which truly relies only on two independently motivated monolingual analyses, we propose to use the notion of *lexical function* (LF). Lexical functions were first introduced in the sixties by Igor Mel'čuk and his colleagues in the context of Meaning-text Theory. Since then, LFs have received wide attention in linguistics, lexicography, and NLP. For a recent overview over the issue, we refer to [14].

Lexical functions capture paradigmatic and syntagmatic co-occurrence restrictions of lexemes. Consider the well-known example of intense liquid precipitation, which in English is referred to as *heavy rain*, while in French it is *pluie forte* 'strong rain'. While the meanings of *heavy* and *forte* appear to be essentially the same ("intense"), we cannot say **strong rain* in English nor **pluie lourde* in French. From this and many other examples discussed in the literature it becomes clear that the choice of adjective marking intensity is conventionalized in language, and that competence in a language includes knowledge about such conventions.

LFs are typically written as the first four letters of a Latin word relating to its special meaning. Here are some of the standard examples:

Example	Explanation
Syn(to telephone) = to phone	Synonymy
$\mathbf{Conv}_{13}$(to sell) = to buy	Conversive: *John sold a car to Mary* is synonymous with *Mary bought a car from John*
Magn(rain) = heavy	Intensification
$\mathbf{Oper}_1$(resistance) = to put up	Support verb for subject: *John put up resistance against Mary*
$\mathbf{Adv}_0$(*to rush*) = *fast*	Relates a verb to an adjunct phrase with the same meaning

It is clear that LFs are an important part of lexicographic description for a given language. In addition, LFs also play a crucial role in relating the lexicons of two different languages in a transfer component. For example, if *heavy* is related to French *lourd* and we transfer lexeme by lexeme, we will not obtain the correct translation for *heavy rain*, but rather **pluie lourde*. We can include *heavy rain* in the transfer component as a complex entry, but this is unappealing because the meaning of the expression is in fact compositional and the phrase shows none of the signs of idioms (restricted syntactic variability and lexical compositionality and so on). Instead, we can, using the English monolingual lexicon, transform *heavy rain* into **Magn**(*heavy*) *rain* and then transfer to **Magn**(*pluie*) *pluie* (note that the transfer of the modifier now involves no transfer lexicon lookup at all). We then use the monolingual French lexicon to generate the correct *pluie forte*. (We observe that an interlingua cannot contribute to solving this problem, since the problem then would simply be displaced into the mapping from interlingua to target language. The same remark holds for all syntagmatic LFs.)

We claim that LFs are a tool we can use in order to eliminate the open-class conceptual features used in our first attempt at replacing an interlingua with lexico-structural transfer. The crucial insight here is that the adjunct that expresses manner of motion is related in a systematic manner to a manner of motion verb. In fact, it is related in a systematic manner to the direct translation of the head of the English sentence. The relation in question is the standard LF $\mathbf{Adv}_0$. For our example, we have $\mathbf{Adv}_0$(*nager*) = *'a la nage* (actually, a proper DSyntS for *'a la nage*).

However, we cannot use a LF to relate a preposition to a motion verb expressing motion on a path (as expressed by the preposition). Such an LF cannot exist, since the relation is not sufficiently systematic, and since the shared meaning component is too small. Therefore, we do lexicalize the entry with the preposition (in English) and the path-of-motion verb in French, as follows:

```
@TRANS_CORR @EN X [cat:verb class:manner-of-motion]
                  (ATTR across [cat:preposition]
                     (II N))
            @FR traverser [cat:verb]
                  (II N
                   ATTR Adv0(X) )
```

Since the number of possible paths is constrained by the number of prepositions in English (as opposed to the number of possible manners, which is unconstrained), we have adequately limited the size of the translation dictionary. In addition, we can easily account for prepositions such as *to* (as in the example *Marie-France swam to Calais* given above) which typically do not trigger structural divergence in translation: for these prepositions, we simply do not have a complex entry of the type given above, and the regular mechanism of lexico-structural transfer results in a non-divergent translation. Thus the apparent lexical idiosyncrasy of this translation problem also speaks directly in favor of the analysis above.

7 Conclusion

The lexically-based transfer approach presented here has been implmented as part of the TransLex MT system [12]. It does not use a level of representation which would be intended as a truly language-independent representation of the meaning of the sentence and would traditionally be termed an "interlingua". Instead, our lexicalized grammar approach provides us with a unified syntactic and semantic representation for each lexical item. The dependency relations we derive during the parsing process represent directly the predicate-argument structure; the DSyntS can in addition be richly annotated with semantic features from the lexicon. By including appropriate cross-linguistic semantic features, and coindexing them in the transfer lexicon, we can capture the same generalizations that are traditionally associated with an interlingua approach.

The principal advantage of a lexically-based transfer approach is that it is still fairly close to the surface structure. This allows us to exploit statistical techniques for analyzing corpora and for extracting information from them (including translation lexicons). We have extracted large parts of the translation lexicons for our subdomains automatically, and we have trained two different parsers on the syntactic structures in the corpora to improve their performance.

However, mapping between language-specific predicate argument structures, although providing more flexibility and facilitating the use of statistical techniques, does not obviate the need for making semantic generalizations across several languages. These generalizations must now be captured by imposing structure on the lexicon of a single language. Then, independently motivated structures for different languages must be related. The structures in questions are either classifications (e.g., verb classes) or systematic relations between lexemes (i.e., lexical functions).[1] Specifying language-specific classifications and relations and finding mappings between them that can facilitate translation is indeed an extremely challenging task on which depends the success of either approach, whether it is transfer-based or interlingua based [10, 3].

References

[1] Dorr, B. J. (1993). *Machine Translation: A View from the Lexicon.* MIT Press, Boston, Mass.

[2] Dorr, B. J. (1994). Machine translation divergences: A formal description and proposed solution. *Computational Linguistics*, 20(4):597–635.

[3] Hoa Trang Dang, Joseph Rosenzweig, M. P. (1997). Verb classifications as a basis for machine translation. In *submitted to the Interlingua Workshop at the MT Summit*, San Diego, California.

[4] Joshi, A. K. (1987). An introduction to Tree Adjoining Grammars. In Manaster-Ramer, A., editor, *Mathematics of Language*, pages 87–115. John Benjamins, Amsterdam.

[5] Kaplan, R. M. and Bresnan, J. W. (1982). Lexical-functional grammar: A formal system for grammatical representation. In Bresnan, J. W., editor, *The Mental Representation of Grammatical Relations.* MIT Press, Cambridge, Mass.

[1]We observe that, formally, a LF is really a relation, which is a subset of the cross-product of the lexicon with itself, while a verb class is simply a subset of the lexicon.

[6] Levin, B. and Hovav., M. R. (1995). *Unaccusativity. At the Syntax-Semantics Interface.* MIT Press.

[7] Mel'čuk, I. A. (1988). *Dependency Syntax: Theory and Practice.* State University of New York Press, New York.

[8] Mel'čuk, I. A. and Polguère, A. (1987). A formal lexicon. *Computational Linguistics*, 13(3–4):13–54.

[9] Palmer, M. and Rosenzweig, J. (1996). Capturing motion verb generalizations with synchronous TAGs. In *Proceedings of AMTA-96*, Montreal, Quebec.

[10] Palmer, M., Rosenzweig, J., Dang, H., and Xia, F. (1997). Capturing syntactic/semantic generalizations in a lexicalized grammar. presentation in working session of Semantic Tagging Workshop, ANLP-97.

[11] Rambow, O. and Joshi, A. (1996). A formal look at dependency grammars and phrase-structure grammars, with special consideration of word-order phenomena. In Wanner, L., editor, *Current Issues in Meaning-Text Theory.* Pinter, London.

[12] Rambow, O., Nasr, A., Palmer, M., Bleam, T., Collins, M., Kipper, K., Melamed, D., Park, J., Rosenzweig, J., Schuler, W., and Srinivas, B. (1997). Machine translation of battlefield messages using lexico-structural transfer. Technical report, CoGenTex, Inc.

[13] Talmy, L. (1991). Path to realization: A typology of event conflation. In *Proceedings of the Berkeley Linguistics Society Parasession*, Berkeley, CA.

[14] Wanner, L., editor (1996). *Lexical Functions in Lexicography and Natural Language Processing.* John Benjamins, Amsterdam and Philadelphia.