CAPTURING MOTION VERB GENERALIZATIONS IN SYNCHRONOUS TREE ADJOINING GRAMMARS

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

Abstract

This paper describes the use of verb class memberships as a means of capturing generalizations about manner-of-motion verbs in Synchronous Tree Adjoining Grammars, STAGs, [20, 21, 22]. This approach allows STAGs, which are essentially transfer-based, to take advantage of the same types of generalizations which are generally thought of as wholly the domain of interlingua systems - without giving up any of the lexical specificity unique to transfer-based systems. In this way a machine translation system based on STAGs can respond with seamless flexibility to a wide spectrum of phenomena being presented for translation ranging from idioms and idiosyncratic lexical items to well-behaved verbs that follow lexical rules.

1 Introduction

This paper describes the use of verb class memberships as a means of capturing generalizations about manner-of-motion verbs in Synchronous Tree Adjoining Grammars, STAGs, [20, 21, 22]. This approach allows STAGs, which are essentially transfer-based, to take advantage of the same types of generalizations which are generally thought of as wholly the domain of interlingua systems - without giving up any of the lexical specificity unique to transfer-based systems. In this way a machine translation system based on STAGs can respond with seamless flexibility to a wide spectrum of phenomena being presented for translation ranging from idioms and idiosyncratic lexical items to well-behaved verbs that follow lexical rules.

This paper begins with a review of the role played by lexical semantics in natural language processing and machine translation. Lexical Conceptual Structure (LCS) as an independent level of lexical semantic representation are introduced, with reference to their implementation in text processing systems and as an interlingua. The use of linking rules as a means of communication between syntactic and semantic levels of representation is discussed and then Lexicalized Tree-Adjoining Grammars (LTAGs) are described. The inherent advantages of lexical trees for the representation of lexical semantic information, as compared with Lexical Conceptual Structures, are emphasized. An LTAG-based approach to machine translation (MT) is presented, and its handling of cross-linguistic lexical divergences is contrasted with an interlingual approach based on Lexical Conceptual Structures. A comparison is also made with the ACQUILEX MT framework, which uses typed feature structures and unification to represent a categorial grammar.

2 Lexical-semantic representation formalisms

A primary task of lexical semantics is to find correct correspondences between the underlying semantic representation of a verb and its alternative syntactic realizations. Given any particular syntactic realization, the appropriate semantic representation is then available. A lexical semantic representation formalism should encode the constraints governing these syntax-semantics correspondences transparently and efficiently, enabling applications such as machine translation systems to retrieve semantic information from syntactic structure as quickly and directly as possible.

2.1 Lexical Conceptual Structures

Lexical Conceptual Structures (LCS), Jackendoff's seminal contribution to the representation of lexical semantic information, have been utilized in a number of natural language processing systems. One application of the LCS formalism, the PUNDIT/KERNEL text processing system [16], makes use of LCS as a way of encoding the predicate-argument structure and selectional preferences of verbs. LCS allows verbs to be expressed as complex predicates which select for typed arguments. Linking rules are used to map to alternative syntactic realizations. In PUNDIT, the explicit syntactic realization of an AGENT or an INSTRUMENT results in substantial changes to the semantic representation that is produced. A second application, the Unitran machine-translation system [7], employs a modified version of the LCS formalism as an *interlingua.* To translate a sentence, this system first converts it into a language-independent structure (an interlingua based on the LCS representation). This structure is then used to generate the target sentence in the other language. The *linking rules of LCS,* which relate argument positions in a semantic predicate with syntactic positions in a parse tree, provide a framework for representing syntax-semantics correspondences. Since Jackendoff's semantic structures are intended to capture cross-linguistic regularities, they offer the promise of providing a language-independent meaning representation, which would be useful in machine translation. An example of an LCS interlingua representation is given in Section 3.1.

To make use of LCS, a implementation must include an independent syntactic parser which produces parse trees for its input. Then, a semantic interpreter which encodes the linking rules of the LCS component is run on the parse tree, resulting in the production of instantiated LCS predicate-argument structures. In the next section, we describe a lexicalized approach to grammar writing which allows the semantic generalizations captured by an LCS to be encoded directly in the syntactic structure of the lexical entry, allowing much closer interaction between syntactic and semantic processing.

2.2 Lexicalized Tree-Adjoining Grammars

Lexicalized Tree-Adjoining Grammars (LTAGs) are tree rewriting systems [12], [20, 19]. The primitive elements of LTAGs are called elementary trees and are of two types: initial trees and auxiliary trees. The minimal, non-recursive linguistic structures of a language, such as a verb and its complements, are captured by initial trees. Recursive structures of a language, such as prepositional modifiers which result in syntactically embedded NPs, are represented by auxiliary trees.

Elementary trees are combined by the operations of substitution and adjunction. Every tree is associated with a lexical item (or non-compositional phrasal form) of the language, called the *anchor* of the tree. The tree represents the domain over which the lexical item can directly specify syntactic constraints, such as subject-verb number agreement, or semantic constraints, such as selectional restrictions, all of which are implemented as features in *Feature-Based* LTAG, FB-LTAG [24]. Alternative syntactic realizations of a lexical item are grouped together into *tree families,* and the semantic constraints automatically apply to the the same arguments in the alternative trees.

There are critical benefits to lexical semantics that are provided by the extended domain of locality of the lexicalized trees. Each lexical entry corresponds to a tree that will be used to build the parse of the sentence that item appears in. If the lexical item is a verb, the corresponding tree is a skeleton for an entire sentence with the verb already present, anchoring the tree as a terminal symbol. The other parts of the sentence, the noun phrases that occur as subject and object, will be substituted or adjoined in at appropriate places in the skeleton tree in the course of the derivation. The composition of trees during parsing is recorded in a *derivation tree.* The derivation tree nodes correspond to lexically anchored initial trees, and the arcs are labelled with information about how these trees were combined to produce the parse. Since each lexically anchored initial tree corresponds to a semantic unit, the derivation tree closely resembles a semantic-dependency representation such as those used in the MT system of [18].

Since a unique tree structure is associated with each lexical item, verb-specific idiosyncrasies can also be captured at the lexical level. Idioms and light verb constructions, such as *kick the bucket* and *take advantage of,* appear in the lexicon as single trees, with multiple lexical items anchoring the tree. However, the tree composition operations of the LTAG formalism allow appropriate lexical items to be inserted between words in such a phrase when necessary, allowing the handling of variations such as *kick the proverbial bucket* or *take unfair advantage of.*

2.2.1 An example

For example, in the case of the related transitive and intransitive trees for verbs such as *break,* the common information for the tree family may specify that the verb has two semantic arguments, say, *agent* and *theme.* The transitive tree in the family is specified as having its subject NP linked to the *agent* argument and its object to the *theme* argument. For the intransitive tree, the subject NP is linked to the *theme,* and the *agent* argument is left unrealized. This linking may be specified by *coindexing* among the feature structures on the nodes of the tree (see Figure 1).

We are currently investigating the types of tree families that can be useful for an English grammar, building on work done by [4] on incorporating such generalizations into a TAG.

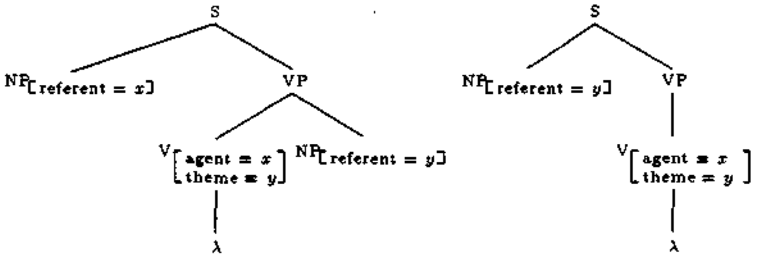

Figure 1: Linking rules in a tree family

2.2.2 Adding semantic features to LTAGs

Selectional restrictions on a verb's arguments can be stated once for all the trees anchored by the verb, irrespective of where the arguments are realized syntactically. The individual trees, which reflect the verb's syntactic variability, each specify, by means of feature-structure co-indexing, where the selectional restrictions are enforced. For example, if one sense of *break* requires its *theme* to have the physical property *+brittle,* this can be noted as a constraint on the feature structure associated with the verb in the lexicon: *[theme = [brittle = +]].* Because of the co-indexing on the *theme* feature structure in the verb tree, this constraint will be propagated to the syntactic position corresponding to the *theme* argument in any tree anchored by *break,* including both the transitive and intransitive trees in Figure 1.

Moreover, two different senses of *break,* separate-into-pieces *break* and non-functional *break,* may be differentiated by appropriate selectional restrictions on the verb object: *+brittle* and *+mechanical-device,* respectively (see Figure 2). If the sentence is *John broke the window, window* will only be able to substitute into the verb tree with the *+brittle* feature, thus selecting the appropriate sense of *break.* It is often the case that a particular language will distinguish

Information	*LCS*	*FB-LTAG*
Syntactic Realizations	Linking Rules	Tree Families
Predicate-Argument Structure	Lexical Conceptual Structure	Elementary Trees
Selectional Restrictions	Added to LCS as Constraints	Features in Trees
Semantic Components	Predicates of LCS	Features for Class Membership

Table 1: Contributions to Lexical Semantics - LCS

senses based on subtle shadings in meaning that are not relevant to other languages. For instance, Mandarin has different expressions for *breaking* events depending on the characteristics of the objects involved. *Da-duan* is used for objects shaped like line segments, *da-po* for objects that *break* into irregularly shaped pieces, and *da-sui* for objects that break into many small pieces, [17]. These features can be associated with the correct expressions in the Mandarin FB-LTAG, without needing to make the same distinctions in other languages. When attempting translations, it is only necessary to pinpoint the range of possible translation, and the selectional restriction features will ensure that inappropriate combinations are not produced [8], [11].

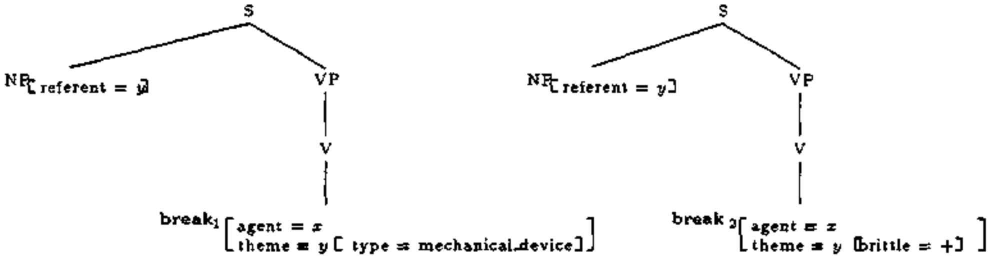

Figure 2: Selectional restrictions for two senses of *break*

Semantic features may also be used to represent verb-class membership. For instance, *change-of-state* verbs such as *dry* can be marked by a feature in the lexicon.[1] Selectional restrictions may be generalized over entire verb classes by specifying feature-structure equations using the verb-class membership features. Also, such features may be used to constrain or facilitate adjunctions in trees anchored by verbs in a particular class. For instance, intransitive *manner-of-motion* verbs such as *walk* and *float* can readily take *path* prepositional phrases such as *to the shore,* creating a directed motion event. Features can be used to associate the appropriate verb class with a lexical item, thus constraining the types of adjunctions that would be most suitable. One possible interpretation of this is that in the FB-LTAG implementation the *path* prepositional phrases constitute constructions, as discussed in Section 3.2.

The elementary trees of LTAGs offer an elegant framework for the direct encoding of lexical semantic constraints, which allow a transfer-based MT system to provide a general treatment for cross-linguistic divergences (see below).

2.3 Summary

The table below summarizes the mechanisms that are used by Lexical Conceptual Structures for *transparent* and *efficient* representation of the requisite lexical semantic information:

We have argued that the extended domain of locality in FB-LTAGs provides an especially suitable environment for representing lexical semantic information. Examining the same criteria

---

[1] The feature may be merely an atomic value, or it may be a more complicated, compositional structure depicting the structure of events which the verb denotes.

that we used to evaluate lexical conceptual structures, we find that FB-LTAGs are indeed quite comparable. They have a significant advantage, however, when it comes to actual processing. Since the syntactic and lexical semantic information are included in the same structure, the application of syntactic and semantic constraints is performed in a single unification step during the parsing process, allowing simultaneous feedback from syntax and lexical semantics on possible linguistic structures.

3 Translating with Synchronous TAGs

The effectiveness of using semantic features in an FB-LTAG is most clearly illustrated by looking at translation examples which heretofore seemed to necessitate interlingua solutions.

FB-LTAGs can be applied to MT using the Synchronous Tree-Adjoining Grammar (STAG) formalism, which is uniquely suited to combining the strengths of both the transfer method and the interlingua method [2], [1], [21, 22], [15]. In STAG-based MT, LTAGs are developed independently for source and target languages. Lexicalized trees of the source language grammar are put in bi-directional correspondence with target language trees. The substitution and adjunction sites on the source trees are linked to substitution and adjunction sites on the target trees. By linking the source-language syntactic site for a verb's argument with the target-language site for the same argument, variations between languages in syntactic argument realization can be handled perspicuously.

In addition to the simple linkings, which pair one source-language tree with one target-language tree, *complex* linkings may also be defined. For example, entire tree families in the source grammar may be linked to target-grammar tree families, capturing generalizations about ranges of related verb frames in both languages. Another important type of complex linking pairs a *composition* of several source-language trees with a composition of target-language trees, as described below.

After a derivation tree for a source sentence has been produced, a derivation tree for the target sentence can be created by following the links and performing the indicated substitutions and adjunctions with the appropriate items from the transfer lexicon. Synchronous TAGs can readily handle the task of both recognizing an idiom and mapping from it to its natural translation in another language.

A common objection to this approach is that it is essentially transfer based, and loses the inherent efficiency of an interlingua approach. However, by enriching the lexicalized tree structures with verb class membership information, many of the advantages of an interlingua representation can be preserved.

3.1 Using verb classes to generalize over structural divergences

For example, a classic problem 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 via a prepositional phrase, as in *John will drive out of the airport.* In many cases, this is not allowed in French, where the direction becomes incorporated into the matrix verb, and the manner is adjoined on as an adverbial or a prepositional phrase, as in *Jean soit sortie l'aéroport par la route.*

A theoretical explanation of this phenomena has been provided by Talmy, [23], who differentiates between verb-framed languages (ex: English) and satellite-framed languages (ex: French): the former map the core schema into the main verb, while the latter map the core schema into the satellite [2].

This substantial difference in how the languages incorporate information is especially troublesome for transfer based systems, since there are many possible direction prepositions that can be adjoined on, and even more different manners of motion. Each possible unique combination of manner and direction must be spelled out explicitly and associated with its appropriate French equivalent.

This type of translation mismatch has been termed *structural divergence,* and is often cited as strong evidence for the advantages of an interlingual representation [7]. Dorr uses Jackendoff's LCS as her interlingua, and this allows her to tease apart the three separate concepts of motion, direction, and manner of motion [7]. Each one is represented separately in the interlingua representation, as GO, PATH and MANNER, respectively.

[[CAUSE JOHN [GO JOHN [PATH [OUT AIRPORT]]][MANNER DRIVINGLY]]]

Then they can be recombined for the French translation in a different order, this time incorporating GO and DIRECTION into *sortir* and leaving MANNER by itself.

However, the STAG approach, though fundamentally a transfer approach, can still translate these classes of sentences in an efficient manner. To be sure, it is necessary to have some shared semantic features in the grammars of both languages, in order to capture, e.g., the generalization that the *path* component of some events, which is described by *out* in English, is also expressed by the verb *sortir in* French (see Figure 3). To this degree, the STAG solution to this problem incorporates the advantages of interlingual semantic generalization. However, a full semantic decomposition of the source input is not required. Instead, the STAG transfer lexicon pairs an abstract representation of a generic directed-motion phrase from the source language with a similar abstract representation of a generic directed-motion phrase from the target language. The transfer lexicon thus abstracts away from individual pairs of lexical items to pairs of verb classes and co-indexed shared semantic features. This is possible because the lexical items in the two grammars are equipped with cross-linguistic verb-class membership features and certain other cross-linguistic semantic features such as directionality.

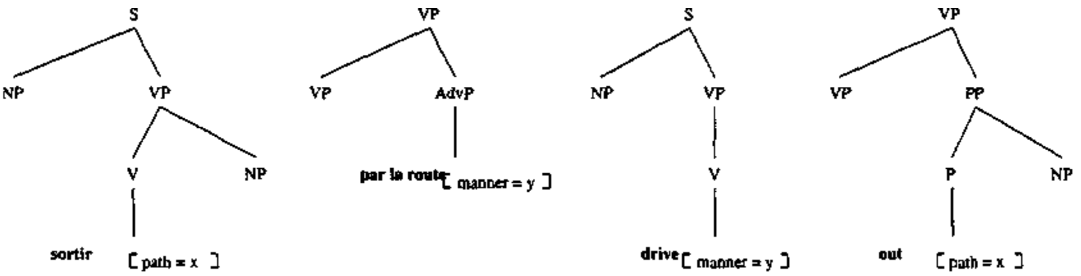

Figure 3: Lexical entries for *drive*, *out*, *sortir* and *route*

In the above example, if *drive* is categorized a *manner-of-motion* verb, and *out* is recognized as supplying *path* information, a combination of the trees anchored by these two words (using the adjunction operation) results in a directed motion event. In the transfer lexicon, a *complex* linking specifies that the composition of an English *manner-of-motion* verb tree with an English *path* prepositional tree can be mapped to a French composition of *directed-motion* verb tree and *manner* preposition phrase tree (see Figure 4). The explicit lexical items on both sides remain unspecified, so that these two partial structures can represent all the possible combinations of

[23] gives definitions for terms such as *core schema, satellite.* The *core schema* is the schematic core of the framing event. The *satellite to the verb* is the grammatical category of any constituent other than a nominal complement that is in a sister relation to the verb root.

manner of motion and path for their respective languages, as in Fig 4. However, the transfer algorithm ensures that the particular lexical items on the source side constrain the selection of correct corresponding lexical items on the target side, by enforcing co-indexation of the cross-linguistic shared semantic features, such as *path* and *manner* features. So for our example sentence only lexical items with features that unify with path: *out* and manner: *drive* can instantiate the source language tree.

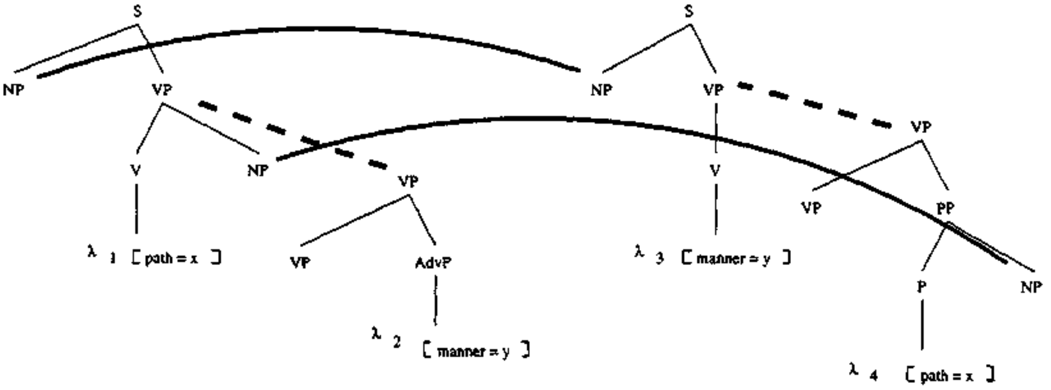

Figure 4: Transfer lexicon link between directed-motion schemas in English and French

3.2 Trees as constructions

In the previous section, we mentioned that an adjunction of the trees anchored by *drive* and *out* resulted in a directed motion event, but we were not specific about the details. By associating the directed-motion event type with the path prepositional phrase, and allowing this feature to be passed up to the S node during the adjunction onto the intransitive verb tree, we add this information to the manner-of-motion event type that is already associated with the intransitive verb. This approach effectively treats the path prepositional phrase as a *construction* in the sense of *construction grammar,* CG, [9], [10]. The claim of CG is that there are atomic, non-decomposable meaning elements at higher levels of grammatical organization — in particular, at the level of phrase and sentence schemas. For example, CG is able to directly associate meaning with a construction for ditransitive sentences, rather than having to specify how the implied meanings in such sentences are derived compositionally from the verb involved and its arguments. CG simply adds a ditransitive syntactic schema to its lexicon.

Since many FB-LTAG trees are designed as phrase and sentence schemas, they are a natural vehicle for capturing the semantic constraints specified in CG. Feature unification provides a mechanism for composing these constraints, where the violation of a constraint is signalled by the failure to find a unifier. The CG constraints on composition may be formulated in a feature-based event-algebra framework, which can be implemented in an augmented FB-LTAG which makes use of a larger set of features and event annotations than is used in the current system. [3]

In the CG approach, as discussed in [3], the verb *drive* in the English sentence *John drove out of the airport* is not analyzed as a predicate which conflates two events, as suggested by Talmy. Instead the sentence's meaning is viewed as the composition of a verb denoting a simple event, such as *drive,* with a syntactic schema which includes the path-denoting prepositional phrase,

---

[3] Some of the constraints make reference to AI notions such as preconditions, which would have to be encoded in the semantics of the events described by verbs and trees.

and which denotes a second simple event, the "directed-motion" event. The complex meaning is derived not by lexical incorporation, but by a compositional semantics of constructions.

An alternative approach to using FB-LTAGs as a vehicle for CG would be to have two separate senses for each manner-of-motion verb already in the lexicon, one that is a simple intransitive and indicates a type of motion but no change of location, and a second sense that takes the path preposition and indicates directed motion, i.e., change of location, in addition [25]. The relationship between the two intransitive senses can be captured by a productive lexical rule allowing the second sense to be generated automatically for all of the relevant verbs during a pre-processing stage. The lexical rule is intended to be compatible with Levin and Rappaport's approach to lexical representation [14]. The distinction between the two approaches can then be characterized from an implementation point of view as a choice between precompiling the pairs in the lexicon or producing them dyamically during run-time.

3.3 Accommodating lexically specific translations

The combination of generalizations based on verb class memberships with an essentially lexicalized transfer-based approach offers a powerful basis for handling the range of phenomena presented by the machine translation task. The structure of FB-LTAGs allows syntactic dependencies, such as agreement and sub categorization, to be localized in the lexical items and their associated elementary trees. Since transfer rules are stated as correspondences between nodes of the elementary trees associated with lexical entries, we can define lexical transfer rules over a larger domain of locality. This provides an elegant method for translating the language-specific idioms and constructions which can pose problems for interlingua approaches. The fixed parts of an expression are built into the tree as lexical anchors to provide for non-compositional readings. For example, verb particle constructions such as *lift up,* or *give away,* appear as trees, but here the verb and the particle both anchor the tree, and are already in the correct position. However, the flexibility of the parsing system allows noun phrases to be inserted in between the verb and its particle if necessary, so both *lift the baby up* and *lift up the baby* can be represented. In the same way, as mentioned above, idioms and light verb constructions, such as *kick the bucket* and *take advantage of,* appear in the lexicon as trees, with all of the essential lexical items anchoring the tree. Once again, these "frozen strings" are flexible enough, that where there is a need to treat them productively, where other lexical items need to be inserted, as in to *pris une vraie veste* or *vraiment pris une veste,* (to really come a cropper), these can also be accounted for [2], [1]. Since the anchor for the tree is the entire phrase, *pris une veste,* the modifier, whether it is an adjective or an adverbial, adjoins on, (albeit in different locations), and is dependent from the same node in the derivation structure. The transfer lexicon pairs the phrasal trees in one language with phrasal trees in another language, and then adjoins on modifiers appropriately. In this way, Synchronous TAGs can readily handle the task of both recognizing an idiom, and mapping from it to the natural translation in another language.

3.4 Comparison with a unification-based approach

The ACQUILEX system [6, 5] provides a typed-feature structure framework for doing MT, based on a HPSG/categorial grammar formalism for the source and target languages. Like an LTAG-based MT approach, the ACQUILEX MT framework uses a set of bi-directional transfer rules, called *tlinks,* to pair up translation equivalents in the two languages. The *tlinks* pair feature structures from the source and target languages. A translation is performed by parsing an input sentence with a unification-based source-language parser, resulting in a source-language feature structure that is mapped to a target-language feature structure, which is used to generate the output. The ACQUILEX framework supports sophisticated and powerful unification procedures, including the ability to perform unification with overrideable default information [5].

Since the linked feature structures may be partially parameterized, the ACQUILEX system, like the LTAG-based system, is able to make generalizations about the translations of semantically similar sentences, such as the directed-motion sentences discussed above.

While the two approaches share many of the same underlying assumptions about lexical semantic representation, the LTAG approach offers some unique advantages stemming from its use of tree composition operations. Its efficient handling of non-compositional phrases is particularly important in the context of MT, where it is undesirable to treat such phrases too rigidly. For example, in translating the phrase *take unfair advantage of* from English, it should be possible to recognize that the target form is systematically related to the translation of the phrase *take advantage of.* An LTAG representation of these two English phrases highlights their similarity in such a way that it is easy to make reference to the common elements of both phrases when defining a transfer lexicon [1]. This is not the case for the representation of these phrases in other grammars, such as the categorial grammars used by ACQUILEX.

4 Summary

In summary, we have presented an approach to machine translation based on Synchronous Tree Adjoining Grammars that is truly a hybrid approach - it seamlessly combines the advantages of transfer-based techniques with a capability for utilizing powerful generalizations that are usually only associated with interlingua approaches. This unique combination is possible because the lexicalized style of our grammar allows a range of syntactic and semantic features to be stored with each lexical item. These features can contain the necessary information for making interlingua style generalizations as well as allowing semi-frozen phrases to be included naturally in structural descriptions.

Much work remains to be done on finding optimal verb classification schemes and appropriate typed feature structure systems for encoding them. We will be looking closely at the ACQUILEX framework and its implementation for suggestions on sophisticated constraint satisfaction, and the most appropriate ways to organize our tree families without losing any of the linguistic desirability of the Tree-Adjoining Grammar formalism.

References

[1] Anne Abeillé and Yves Schabes. Non compositional discontinuous constituents in TAG. In *Proceedings of the International Symposium on Discontinuous Constituents,* 1990.

[2] Anne Abeillé and Yves Schabes. Parsing Idioms in Lexicalized TAGs. In *Proceedings of EACL '89,* pages 161-65, 1989.

[3] Jon Aske. Path predicates in English and Spanish: A closer look. In *Proceedings of the Berkeley Linguistics Society,* Berkeley, California, pages 1 - 15, 1989.

[4] Tilman Becker. *HyTAG: A new Type of Tree Adjoining Grammars for Hybrid Syntactic Representation of Free Word Order Languages.* PhD thesis, Universität des Saarlandes, 1993.

[5] Briscoe E. J., Copestake A., and V. de Paiva. *Inheritance, defaults and the lexicon.* Cambridge University Press, 1994.

[6] Copestake A. and Sanfilippo A. Multilingual lexical representation. In *Proceedings of the AAAI Spring Symposium: Building Lexicons for Machine Translation.* Stanford, California, 1993.

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

[8] D. Egedi, M. Palmer, H.S. Park, and A. Joshi. Korean to english translation using synchronous tags. In *First Conference of the Association for Machine Translation in the Americas,* Columbus, MD, October 1994.

[9] C. J. Fillmore, Paul Kay and Mary Catherine O'Connor. Regularity and idiomaticity in grammatical constructions: The case of *let alone.* In *Language* 64:3, pages 501 - 538, 1988.

[10] Adele E. Goldberg. *Constructions. A Construction Grammar Approach to Argument Structure.* University of Chicago Press, 1995.

[11] Chunghye Han, Fei Xia, Martha Palmer, and Joseph Rosenzweig. Capturing language specific constraints on lexical selection with feature-based lexicalized tree-adjoining grammars. In *Proceedings of the International Conference on Chinese Computing,* 1996.

[12] Aravind K. Joshi, L. Levy, and M. Takahashi. Tree Adjunct Grammars. *Journal of Computer and System Sciences,* 1975.

[13] B. Levin. *English Verb Classes and Alternations.* University of Chicago Press, 1993.

[14] B. Levin and M. Rappaport Hovav. *Unaccusativity. At the syntax-semantics interface.* MIT Press, 1995.

[15] M. Palmer, H.S. Park, and D. Egedi. The application of korean-english machine translation to a military message domain. In *Fifth Annual IEEE Dual-Use Technologies and Applications Conference,* May 1995.

[16] M. Palmer, C. Weir, R. Passonneau, and T. Finin. The kernel text understanding system. In *Artificial Intelligence* 63, pages 17- 68, October 1993. Special Issue on Text Understanding.

[17] M. Palmer, W. Zhibiao. Verb semantics for English-Chinese translation. In *Machine Translation* 10, pages 59 - 92, 1995.

[18] M. T. Rosetta. *Compositional Translation.* Kluwer, 1994.

[19] Yves Schabes. *Mathematical and Computational Aspects of Lexicalized Grammars.* PhD thesis, Computer Science Department, University of Pennsylvania, 1990.

[20] Yves Schabes, Anne Abeillé, and Aravind K. Joshi. Parsing strategies with 'lexicalized' grammars: Application to Tree Adjoining Grammars. In *Proceedings of the 12th International Conference on Computational Linguistics (COLING'88),* Budapest, Hungary, August 1988.

[21] Stuart Shieber and Yves Schabes. Synchronous Tree Adjoining Grammars. In *Proceedings of the 13th International Conference on Computational Linguistics (COLING'90),* Helsinki, Finland, 1990.

[22] Stuart Shieber and Yves Schabes. Generation and Synchronous Tree Adjoining Grammars. In *Proceedings of the Fifth International Workshop on Natural Language Generation,* Pittsburgh, Pennsylvania, 1991.

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

[24] K. Vijay-Shanker and Aravind K. Joshi. Unification Based Tree Adjoining Grammars. In J. Wedekind, editor, *Unification-based Grammars.* MIT Press, Cambridge, Massachusetts, 1991.

[25] Clare Voss. Interlingua-based Machine Translation of Spatial Expressions. PhD Dissertation, Department of CS, U of Maryland, 1996.

[26] The XTAG Group. A Lexicalized Tree Adjoining Grammar for English. Technical Report IRCS 95-03, University of Pennsylvania, 1995.