Machine Translation Divergences: A Formal Description and Proposed Solution

Bonnie J. Dorr*
University of Maryland

There are many cases in which the natural translation of one language into another results in a very different form than that of the original. The existence of translation divergences (i.e., cross-linguistic distinctions) makes the straightforward transfer from source structures into target structures impractical. Many existing translation systems have mechanisms for handling divergent structures but do not provide a general procedure that takes advantage of the systematic relation between lexical–semantic structure and syntactic structure. This paper demonstrates that a systematic solution to the divergence problem can be derived from the formalization of two types of information: (1) the linguistically grounded classes upon which lexical–semantic divergences are based; and (2) the techniques by which lexical–semantic divergences are resolved. This formalization is advantageous in that it facilitates the design and implementation of the system, allows one to make an evaluation of the status of the system, and provides a basis for proving certain important properties about the system.

1. Introduction

There are many cases in which the natural translation of one language into another results in a very different form than that of the original. The existence of translation *divergences* (i.e., cross-linguistic distinctions) makes the straightforward transfer from source structures into target structures impractical. This paper demonstrates that a systematic solution to the divergence problem can be derived from the formalization of two types of information: (1) the linguistically grounded classes upon which lexical–semantic divergences are based; and (2) the techniques by which lexical–semantic divergences are resolved. An important result of this formalization is the provision of a framework for proving that the lexical–semantic divergence classification proposed in the current approach covers all source-language/target-language distinctions based on lexical–semantic properties. Other types of divergences and mismatches are outside of the scope of this paper; these include distinctions based on purely syntactic information, idiomatic usage, aspectual knowledge, discourse knowledge, domain knowledge, or world knowledge.[1]

Although other translation approaches have attempted to account for divergences, the main innovation of the current approach is that it provides a formalization of these divergences and the techniques by which they are resolved. This is advantageous from a computational point of view in that it facilitates the design and implementation of

* Department of Computer Science, University of Maryland, A. V. Williams Building, College Park, MD 20742, USA.

1 The reader is referred to Dorr (1993a) for a discussion of how syntactic divergences are handled. Aspectual divergences are treated by Dorr (1992a). The relation of the current framework to other types of knowledge outside of lexical semantics is discussed by Dorr and Voss (1993b).

(1) **Thematic divergence:**
E: I like Mary ⇔ S: María me gusta a mí
'Mary pleases me'
(2) **Promotional divergence:**
E: John usually goes home ⇔ S: Juan suele ir a casa
'John tends to go home'
(3) **Demotional divergence:**
E: I like eating ⇔ G: Ich esse gern
'I eat likingly'
(4) **Structural divergence:**
E: John entered the house ⇔ S: Juan entró en la casa
'John entered in the house'
(5) **Conflational divergence:**
E: I stabbed John ⇔ S: Yo le di puñaladas a Juan
'I gave knife-wounds to John'
(6) **Categorial divergence:**
E: I am hungry ⇔ G: Ich habe Hunger
'I have hunger'
(7) **Lexical divergence:**
E: John broke into the room ⇔ S: Juan forzó la entrada al cuarto
'John forced (the) entry to the room'

Figure 1
Examples of translation divergences with respect to English, Spanish, and German.

the system: the problem is clearly defined in terms of a small number of divergence categories, and the solution is systematically stated in terms of a uniform translation mapping and a handful of simple lexical–semantic parameters. In addition, the formalization allows one to make an evaluation of the status of the system. For example, given the formal description of the interlingua and target-language root words, one is able to judge whether a particular target-language sentence fully covers the concept that underlies the corresponding source-language sentence. Finally, the formalization of the divergence types and the associated solution allows one to prove certain properties about the system. For example, one might want to determine whether the system is able to handle two or more simultaneous divergences that interact in some way. With the mechanism of the current approach, one is able to prove formally that such cases are handled in a uniform fashion.

This paper will focus on the problem of lexical–semantic divergences and will provide support for the view that it is possible to construct a finite cross-linguistic classification of divergences and to implement a systematic mapping between the interlingual representation and the surface syntactic structure that accommodates all of the divergences in this classification. The types of divergences under consideration are those shown in Figure 1. The first divergence type is *thematic*: in (1), the theme is realized as the verbal object (*Mary*) in English but as the subject (*María*) of the main verb in Spanish. The second divergence type, *promotional*, is one of two *head switching* divergence types: in (2), the modifier (*usually*) is realized as an adverbial phrase in English but as the main verb *soler* in Spanish. The third divergence type, *demotional*, is another type of *head switching* divergence: in (3), the word *like* is realized as a main verb in English but as an adverbial modifier (*gern*) in German.[2] The fourth

2 The distinction between promotional and demotional divergences is not intuitively obvious at first glance. In both (2) and (3), the translation mapping associates a main verb with an adverbial satellite, or vice versa (i.e., in (2), the main verb *soler* is associated with the adverbial satellite *usually*, and in (3) the main verb *like* is associated with the adverbial satellite *gern*). The distinction between these two

divergence type is *structural*: in (4), the verbal object is realized as a noun phrase (*the house*) in English and as a prepositional phrase (*en la casa*) in Spanish. The fifth divergence type is *conflational*. Conflation is the incorporation of necessary participants (or arguments) of a given action. In (5), English uses the single word *stab* for the two Spanish words *dar* (*give*) and *puñaladas* (*knife-wounds*); this is because the effect of the action (i.e., the *knife-wounds* portion of the lexical token) is *conflated* into the main verb in English. The sixth divergence type is *categorial*: in (6), the predicate is adjectival (*hungry*) in English but nominal (*Hunger*) in German. Finally, the seventh divergence type is a *lexical* divergence: in (7), the event is lexically realized as the main verb *break* in English but as a different verb *forzar* (literally *force*) in Spanish.

The next section discusses the divergence classification given above, comparing the current divergence categories with those of other researchers. Section 3 formally defines the terms used to classify divergences. Section 4 uses this terminology to formalize the divergence classification and to define the solution to the divergence problem in the context of detailed examples. Finally, Section 5 discusses certain issues of relevance to the divergence problem including the resolution of several (recursively) interacting divergence types.

2. Classification of Machine Translation Divergences

The divergence problem in machine translation has received increasingly greater attention in recent literature (see, for example, Barnett et al. 1991a, 1991b; Beaven 1992a, 1992b; Dorr 1990a, 1990b; Kameyama et al. 1991; Kinoshita, Phillips, and Tsujii 1992; Lindop and Tsujii 1991; Tsujii and Fujita 1991; Whitelock 1992; related discussion can also be found in work by Melby [1986] and Nirenburg and Nirenburg [1988]). In particular, Barnett et al. (1991a) divide distinctions between the source language and the target language into two categories: translation *divergences*, in which the same information is conveyed in the source and target texts, but the structures of the sentences are different (as in previous work by Dorr [1990a, 1990b]); and translation *mismatches*, in which the information that is conveyed is different in the source and target languages (as described by Kameyama et al. [1991]).[3] Although translation mismatches are a major problem for translation systems that must be addressed, they are outside the scope of the model presented here. (See Barnett et al. 1991a, 1991b; Carbonell and Tomita 1987; Meyer, Onyshkevych, and Carlson 1990; Nirenburg, Raskin, and Tucker 1987; Nirenburg and Goodman 1990; Nirenburg and Levin 1989; Wilks 1973; among others, for descriptions of interlingual machine translation approaches that take into account knowledge outside of the domain of lexical semantics.)

Although researchers have only recently begun to classify divergence types systematically, the notion of translation divergences is not a new one in the machine translation community. For example, a number of researchers working on the Eurotra project have sought to solve divergent source-to-target translations, although the divergences were named differently and were resolved by construction-specific transfer rules. (For cogent descriptions of the Eurotra project, see, for example, Arnold and des Tombe 1987; Copeland et al. 1991; and Johnson, King, and des Tombe 1985).

---

head switching cases will be made clearer in Section 4.3.

3 An example of the latter situation is the translation of the English word *fish* into Spanish: the translation is *pez* if the fish is still in its natural state, but it is *pescado* if the fish has been caught and is suitable for food. It is now widely accepted that, in such a situation, the machine translation system must be able to derive the required information from discourse context and a model of the domain that is being discussed.

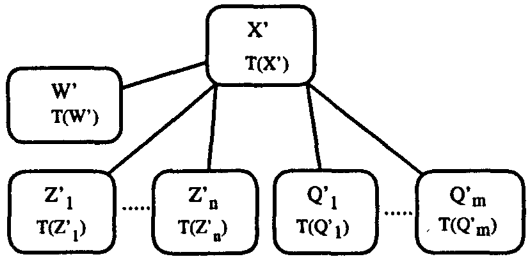

Figure 2
Formal definition of lexical conceptual structure.

A comprehensive survey of divergence examples is presented by Lindop and Tsujii (1991). The term used in this work is "complex transfer," but it describes a class of problems inherent in machine translation itself, not just in the transfer (or interlingual) approaches.

One of the claims made by Lindop and Tsujii (1991) is that the non-Eurotra literature rarely goes into great detail when discussing how divergences are handled. An additional claim is that combinations of divergences and interaction effects between divergent and nondivergent translations are not described in the literature. This paper seeks to change this perceived state of affairs by providing a detailed description of a solution to all of the (potentially interacting) divergences shown in Figure 1, not just a subset of them as would typically be found in the description of most translation systems. The framework assumed for the current approach makes use of a linguistically grounded classification of divergence types that can be formally defined and systematically resolved.

We now turn to a formal description of the terminology used to define the divergence problem.

3. Definitions

This section formally defines the lexical–semantic representation that serves as the interlingua of the system (Definitions 1–3). This representation, which is influenced primarily by Jackendoff (1983, 1990), has been described in detail elsewhere (see, for example, Dorr 1992b, 1993a) and thus will not be the focus of this paper. In addition to a formal description of the lexical–semantic representation, definitions are provided for syntactic phrases (Definition 4) and two translation mappings (Definitions 5 and 6).

Definition 1
A *lexical conceptual structure* (LCS) is a modified version of the representation proposed by Jackendoff (1983, 1990) that conforms to the following structural form:

$$[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z'_1)}\ Z'_1] \ldots [_{T(Z'_n)}\ Z'_n]\ [_{T(Q'_1)}\ Q'_1] \ldots [_{T(Q'_m)}\ Q'_m])]$$

This corresponds to the tree-like representation shown in Figure 2, in which (1) X' is the *logical head*; (2) W' is the *logical subject*; (3) $Z'_1 \ldots Z'_n$ are the *logical arguments*; and (4) $Q'_1 \ldots Q'_m$ are the *logical modifiers*. These four positions are relevant to the mapping

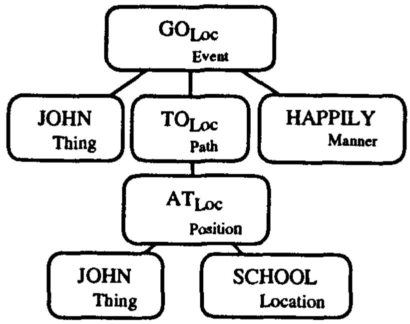

Figure 3
CLCS representation for *John went happily to school.*

between the interlingual representation and the surface syntactic representation. In addition, T(ϕ) is the logical *type* (Event, State, Path, Position, etc.) corresponding to the *primitive* ϕ (CAUSE, LET, GO, STAY, BE, etc.); Primitives are further categorized into *fields* (e.g., Possessional, Identificational, Temporal, Locational, etc.).[4]

Example 1
The LCS representation of *John went happily to school* is

$[_{Event}\ GO_{Loc}$
$([_{Thing}\ JOHN],$
$[_{Path}\ TO_{Loc}\ ([_{Position}\ AT_{Loc}\ ([_{Thing}\ JOHN],\ [_{Location}\ SCHOOL])])]$
$[_{Manner}\ HAPPILY])]$

This corresponds to the tree-like representation shown in Figure 3, in which (1) the logical head is GO_{Loc} (of type Event); (2) the logical subject is JOHN (of type Thing); (3) the logical argument is TO_{Loc} (of type Path); and (4) the logical modifier is HAPPILY (of type Manner). Note that the logical argument is itself a LCS that contains a logical argument, SCHOOL (of type Location), i.e., LCSs are recursively defined.

The LCS representation is used both in the lexicon and in the interlingual representation. The former is identified as a *root* LCS (RLCS) and the latter is identified as a *composed* LCS (CLCS):

Definition 2
A *RLCS* (i.e., a *root* LCS) is an uninstantiated LCS that is associated with a word definition in the lexicon (i.e., a LCS with unfilled variable positions).

Example 2
The RLCS associated with the word *go* (from Example 1) is

$$[_{Event}\ GO_{Loc}\ ([_{Thing}\ X],\ [_{Path}\ TO_{Loc}\ ([_{Position}\ AT_{Loc}\ ([_{Thing}\ X],\ [_{Location}\ Z])])])]$$

4 The validity of the primitives and their compositional properties is not discussed here. The LCS has been studied as the basis of a representation for multiple languages (see, for example, Hale and Keyser 1986a, 1986b, 1989; Hale and Laughren 1983; Levin and Rappaport 1986; Zubizarreta 1982, 1987) and is discussed in the context of machine translation by Dorr (1992b).

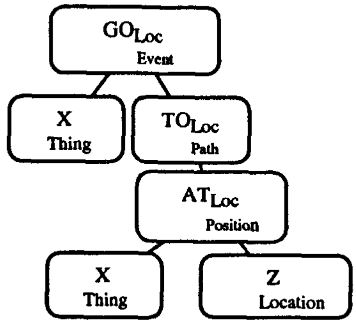

Figure 4
RLCS representation for *go*.

which corresponds to the tree-like representation shown in Figure 4.

Definition 3
A *CLCS* (i.e., a *composed* LCS) is an instantiated LCS that is the result of combining two or more RLCSs by means of unification (roughly). This is the *interlingua*, or language-independent, form that serves as the pivot between the source and target languages.

Example 3
If we compose the RLCS for *go* (in Figure 4) with the RLCSs for *John* ($[_{\text{Thing}}$JOHN]), *school* ($[_{\text{Location}}$ SCHOOL]), and *happily* ($[_{\text{Manner}}$ HAPPILY]), we get the CLCS corresponding to *John went happily to school* (shown in Figure 3).

Each (content) word in the lexicon is associated with a RLCS, whose variable positions may have certain restrictions. The CLCS is a structure that results from combining the lexical items of a source-language sentence into a single underlying pivot form by means of LCS composition.[5] The notion of *unification* (as used in Definition 3) differs from that of the standard unification frameworks (see, for example, Shieber et al. 1989, 1990; Kaplan and Bresnan 1982; Kaplan et al. 1989; Kay 1984; etc.) in that it is not directly invertible. That is, the generation process operates on the CLCS in a unification-like fashion that roughly mirrors the LCS composition process, but it is not a direct inverse of this process. The notion of unification used here also differs from others in that it is a more "relaxed" notion: those words that are mapped in a relaxed way are associated with special lexical information (i.e., the :INT, :EXT, :PROMOTE, :DEMOTE, $*$, :CAT, and :CONFLATED parameters, each of which will be formalized shortly).

A fundamental component of the mapping between the interlingual representation and the surface syntactic representation is the *syntactic phrase*.

Definition 4
A *syntactic phrase* is a maximal projection that conforms to the following structural form:

5 This process is described in detail in Dorr (1992b).

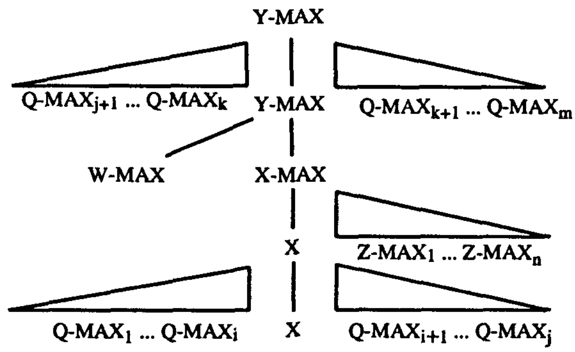

Figure 5
Formal definition of syntactic phrase.

$[_{\text{Y-MAX}}$
$\quad$ Q-MAX$_{j+1}$ ... Q-MAX$_k$
$\quad [_{\text{Y-MAX}}$
$\quad\quad$ W-MAX
$\quad\quad [_{\text{X-MAX}} [_{\text{X}}$ Q-MAX$_1$ ... Q-MAX$_i$ X Q-MAX$_{i+1}$ ... Q-MAX$_j$]
$\quad\quad$ Z-MAX$_1$ ... Z-MAX$_n$]]
$\quad$ Q-MAX$_{k+1}$ ... Q-MAX$_m$][6]

This corresponds to the tree-like representation shown in Figure 5, in which (1) X is the *syntactic head* (of category V, N, A, P, I, or C); (2) W-MAX is the *external argument;* (3) Z-MAX$_1$, ..., Z-MAX$_n$ are the *internal arguments;* and (4) Q-MAX$_1$, ..., Q-MAX$_m$ are the *syntactic adjuncts.*

Example 4
The syntactic phrase corresponding to *John went happily to school* is
$[_{\text{C-MAX}} [_{\text{I-MAX}} [_{\text{N-MAX}}$ John]
$[_{\text{V-MAX}} [_{\text{V}}$ went] $[_{\text{ADV}}$ happily] $[_{\text{P-MAX}}$ to $[_{\text{N-MAX}}$ school]]]]]

This corresponds to the tree-like representation shown in Figure 6, in which (1) the syntactic head is $[_{\text{V}}$ went]; (2) the external argument is $[_{\text{N-MAX}}$ John]; (3) the internal argument is $[_{\text{P-MAX}}$ a ...]; and (4) the syntactic adjunct is $[_{\text{ADV}}$ happily]. Note that the internal argument constituent is itself a syntactic phrase that contains an internal argument, $[_{\text{N-MAX}}$ school], i.e., syntactic phrases are recursively defined.

In addition to the representations involved in the translation mapping, it is also possible to formalize the mapping itself. The current approach is to map between

6 These syntactic structures are based on the $\overline{\text{X}}$ framework of *government-binding* theory (see Chomsky 1981, 1982, 1986a, 1986b). For ease of illustration, the word order used in all formal definitions is head-initial/spec-initial (i.e., the setting for English). The syntactic operations that determine word order are completely independent from the lexical–semantic operations that use these definitions. Thus, the formal definitions can be stated in terms of an arbitrary ordering of constituents, without loss of generality, as long as it is understood that the constituent order is independently determined.

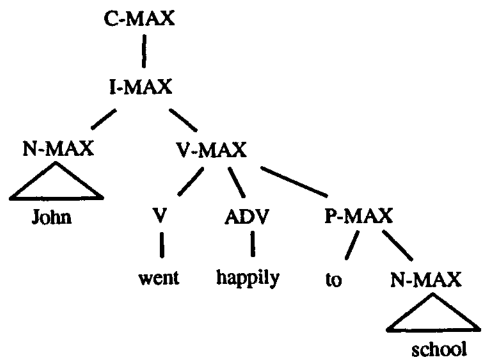

Figure 6
Syntactic phrase representation for *John went happily to school.*

the LCS representation and the surface syntactic form by means of two routines that are grounded in linguistic theory: a *generalized linking routine* ($\mathcal{GLR}$) and a *canonical syntactic realization* ($\mathcal{CSR}$). These routines are defined formally here:

Definition 5
The $\mathcal{GLR}$ systematically relates syntactic positions from Definition 1 and lexical–semantic positions from Definition 4 as follows:

1. $X' \Leftrightarrow X$
2. $W' \Leftrightarrow W$
3. $Z'_1 \ldots Z'_n \Leftrightarrow Z_1 \ldots Z_n$
4. $Q'_1 \ldots Q'_m \Leftrightarrow Q_1 \ldots Q_m$

Example 5
The correspondence between the LCS of Example 1 and the syntactic structure of Example 4 (i.e., for the sentence *John went happily to school*) is (1) $X' = \text{GO}_{\text{Loc}} \Leftrightarrow X = [_{\text{V}}$ went]; (2) $W' = \text{JOHN} \Leftrightarrow W = [_{\text{N-MAX}}$ John]; (3) $Z' = \text{TO}_{\text{Loc}} \Leftrightarrow Z = [_{\text{PP}}$ to ...]; and (4) $Q' = \text{HAPPILY} \Leftrightarrow Q = [_{\text{ADV}}$ happily].

Definition 6
The $\mathcal{CSR}$ systematically relates a lexical–semantic type $T(\phi')$ to a syntactic category $\text{CAT}(\phi)$, where ϕ' is a CLCS constituent related to the syntactic constituent ϕ by the $\mathcal{GLR}$.

Example 6
The LCS type Thing corresponds to the syntactic category N, which is ultimately projected up to a maximal level (i.e., N-MAX). The full range of realization possibilities is given in Figure 7.

LCS Type	Syntactic Category
EVENT	V
STATE	V
THING	N
PROPERTY	A
PATH	P
POSITION	P
LOCATION	ADV
TIME	ADV
MANNER	ADV
INTENSIFIER	ADV
PURPOSE	ADV

Figure 7
$\mathcal{CSR}$ mapping between LCS types and syntactic categories.

Now that we have formally defined the representations and mappings used during translation, we will turn to a classification of divergences that is based on these definitions.

4. The Divergence Problem: Formal Classification and Solution

In general, translation divergences occur when there is an exception either to the $\mathcal{GLR}$ or to the $\mathcal{CSR}$ (or to both) in one language, but not in the other.[7] This premise allows one to define formally a classification of all lexical–semantic divergences that arise during translation (i.e., divergences based on properties associated with lexical

7 Most of the examples in this paper seem to suggest that a divergence is defined in terms of a *language-to-language* phenomenon: a divergence occurs when a sentence in language L_1 translates into a sentence in L_2 in a very different form (i.e., differently shaped parse trees or similarly shaped trees with different basic categories). This definition implies that a divergence may arise between two languages L_1 and L_2, independent of the way the translation is done (i.e., direct, transfer, or interlingual). However, it is also possible to define a divergence from an *interlingual* point of view, i.e., with respect to an underlying representation (lexical conceptual structure) that has been chosen to describe the source and target language sentences. From this point of view, a divergent mapping may apply even in cases in which the source- and target-language pairs do not exhibit any distinctions on the surface (e.g., the translation of the German sentence *Hans kußt Marie gern* as the equivalent Dutch sentence *Hans kust Marie graag*, both of which literally translate to *Hans kisses Mary likingly*). In such cases, there are generally two occurrences of a *language-to-interlingua* divergence: one *from* the surface structure and one *to* the surface structure. (The terms *language-to-language* and *language-to-interlingua* are taken from Dorr and Voss 1993a.) At first glance, it might seem odd to introduce the notion of a language-to-interlingua divergence for cases that do not exhibit a language-to-language divergence. However, it is clearly the case that language-to-language divergences—a special case of language-to-interlingua divergences—do exist regardless of the translation approach adopted. Thus, we can view divergences more generally as a consequence of the internal mapping between the surface structure and the interlingual representation rather than as an external distinction that shows up on the surface. The result is that the interlingua appears to have been simplified to the extent that it accommodates constructions in one language (without any special information) more readily than it accommodates the corresponding construction in another language. However, as one reviewer points out, this is not an undesirable consequence, since the development of a suitable representation is where the interlingua builder has a choice and should choose the simplest representation format. The appropriate question to ask is whether an approach that addresses the divergence problem from a language-to-interlingua perspective is an improvement over an approach that addresses the problem strictly from a language-to-language point of view. This paper argues that the language-to-interlingua approach is the correct one given that the alternative would be to handle language-to-language divergences by constructing detailed source-to-target transfer rules for each lexical entry in the source and target language. Introducing the notion of language-to-interlingua divergence allows the translation mapping to be defined in terms of a representation that is general enough to carry over to several different language pairs.

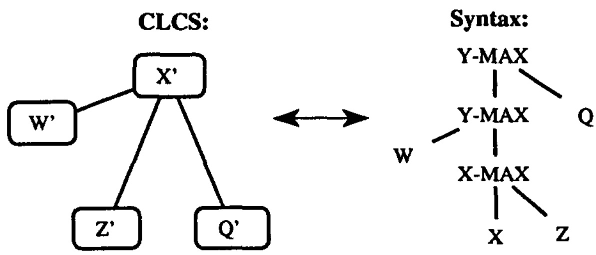

Figure 8
$\mathcal{GLR}$ mapping between the CLCS and the syntactic structure.

entries that are not based on purely syntactic information, idiomatic usage, aspectual knowledge, discourse knowledge, domain knowledge, or world knowledge).

Before we define and resolve each divergence type, we will first make some revisions to the representations used in Definitions 1 and 4 to simplify the presentation. The representation given in Definition 1 is revised so that Z′ is used to denote a logical argument from the set $\{Z'_1 \ldots Z'_n\}$ and Q′ is used to denote a logical modifier from the set $\{Q'_1 \ldots Q'_m\}$. The resulting representation is considerably simplified:

(8) $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z'],\ [_{T(Q')}\ Q'])]$

Similarly, the representation given in Definition 4 is revised so that W is used to denote the external argument, Z is used to denote an internal argument from the set $\{\text{Z-MAX}_1 \ldots \text{Z-MAX}_n\}$, and Q is used to denote a syntactic adjunct from the set $\{\text{Q-MAX}_1 \ldots \text{Q-MAX}_m\}$. The resulting representation has the following simplified form:

(9) $[_{\text{Y-MAX}}\ [_{\text{Y-MAX}}\ W\ [_{\text{X-MAX}}\ X\ Z]]\ Q]$[8]

With these simplifications, the $\mathcal{GLR}$ can be conceptualized as the following set of relations:

(10) **Simplified $\mathcal{GLR}$:**
1. $X' \Leftrightarrow X$
2. $W' \Leftrightarrow W$
3. $Z' \Leftrightarrow Z$
4. $Q' \Leftrightarrow Q$

Figure 8 shows the simplified $\mathcal{GLR}$ in terms of tree-like representations.[9]

We are now prepared to define and resolve the translation divergences of Figure 1 on the basis of the simplified formalization presented in (8)–(10) above. The

8 For the purposes of this discussion, we will retain the convention that syntactic adjuncts occur on the right at the maximal level. Note that this is not always the case: the setting of an adjunction parameter (described by Dorr [1993b]) determines the side and level at which a particular adjunct will occur.

9 For ease of illustration, this diagram omits the type specification. (There is no loss of generality, since the $\mathcal{GLR}$ mapping does not make use of this specification.) We will retain this convention throughout the rest of this paper.

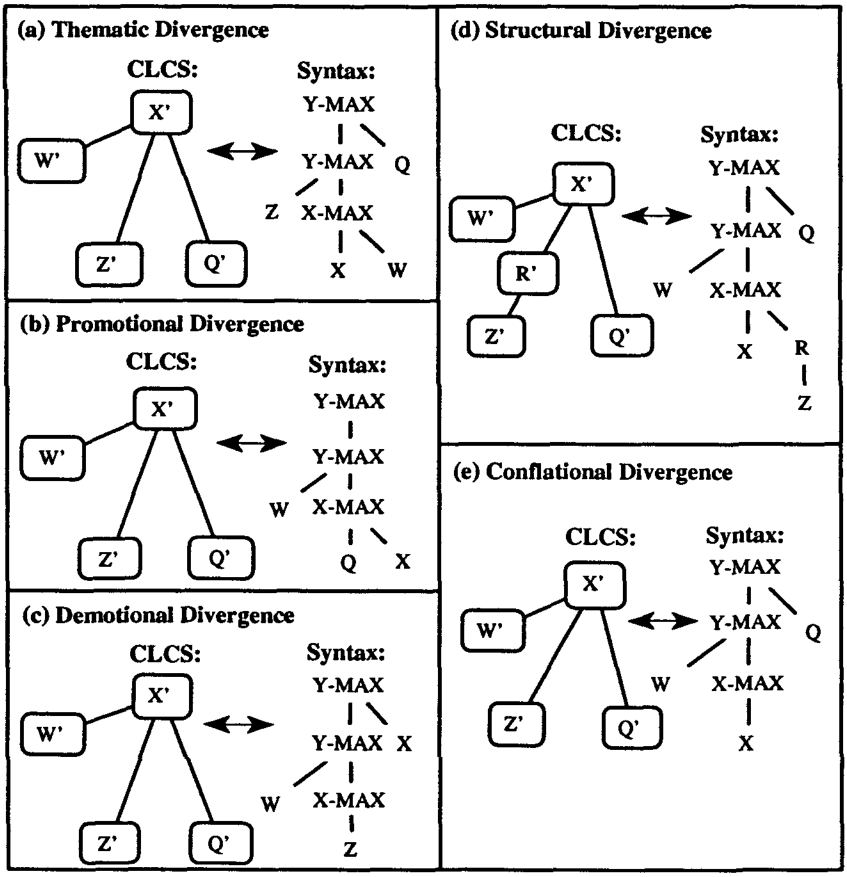

Figure 9
Translation mappings for cases in which $\mathcal{GLR}$ default positions are overridden.

solution to the divergence problem relies solely on three types of information: the $\mathcal{GLR}$; the $\mathcal{CSR}$; and a small set of parametric mechanisms. The $\mathcal{GLR}$ and $\mathcal{CSR}$ are intended to be language independent, whereas the parameters are intended to encode language-specific information about lexical items. Because the interlingual representation preserves relevant lexical–semantic relations, these three types of information are all that are required for providing a systematic solution to the divergence types shown in Figure 1. In particular, the solution given here eliminates the need for transfer rules and relies instead on parameterized mappings that are defined and applied uniformly across all languages. Seven parameters are used to invoke exceptions to the $\mathcal{GLR}$ and $\mathcal{CSR}$ functions in the context of translation divergences: :INT, :EXT, :PROMOTE, :DEMOTE, $*$, :CAT, and :CONFLATED. We will now present a formal description of each divergence type and its associated parameter.

4.1 Thematic Divergence

The first divergence type to be formalized is the one for thematic divergence, i.e., the repositioning of arguments with respect to a given head. This type of divergence arises in cases in which the $\mathcal{GLR}$ invokes the following sets of relations in place of steps 2 and 3 of (10):

(11) 2.′ W′ ⇔ Z
3.′ Z′ ⇔ W

Figure 9a shows the revised mapping.

Thematic divergence arises only in cases in which there is a logical subject. An example of thematic divergence is the reversal of the subject with an object, as in the thematic divergence example given earlier in (1). The syntactic structures and corresponding CLCS are shown here:

(12) $[_{C-MAX}\ [_{I-MAX}\ [_{N-MAX}\ \text{I}]\ [_{V-MAX}\ [_{V}\ \text{like}]\ [_{N-MAX}\ \text{Mary}]]]]$
⇕
$[_{State}\ \text{BE}_{Ident}\ ([_{Thing}\ \text{I}],$
$[_{Position}\ \text{AT}_{Ident}\ ([_{Thing}\ \text{I}],\ [_{Thing}\ \text{MARY}])],$
$[_{Manner}\ \text{LIKINGLY}])]$
⇕
$[_{C-MAX}\ [_{I-MAX}\ [_{N-MAX}\ \text{María}]\ [_{V-MAX}\ [_{V}\ \text{me gusta}]]]]$[10]

Here the object *Mary* has reversed places with the subject *I* in the Spanish translation. The result is that the object *Mary* turns into the subject *María*, and the subject *I* turns into the object *me*.

This argument reversal is resolved by means of the :INT and :EXT parameters, which force the $\mathcal{GLR}$ mapping to be overridden with respect to the positioning of the logical subject and logical argument in Spanish. The lexical entries for *like* and *gustar* illustrate the difference in the use of these parameters:

(13) (i) **Lexical entry for *like*:**
$[_{State}\ \text{BE}_{Ident}\ ([_{Thing}\ \text{W}],$
$[_{Position}\ \text{AT}_{Ident}\ ([_{Thing}\ \text{W}],\ [_{Thing}\ \text{Z}])],$
$[_{Manner}\ \text{LIKINGLY}])]$

(ii) **Lexical entry for *gustar*:**
$[_{State}\ \text{BE}_{Ident}\ ([_{Thing}\ \text{:INT W}],$
$[_{Position}\ \text{AT}_{Ident}\ ([_{Thing}\ \text{W}],\ [_{Thing}\ \text{:EXT Z}])],$
$[_{Manner}\ \text{LIKINGLY}])]$

Because the English entry does not include these parameters, the translation relies on the default argument positionings imposed by the $\mathcal{GLR}$. By contrast, the :INT/:EXT markers specified in the Spanish entry force the internal and external arguments to swap places in the syntactic structure.

10 For the purposes of this discussion, the Spanish sentence is given in its uninverted form. There are other ways of realizing this sentence. In particular, a native speaker of Spanish will frequently invert the subject to post-verbal position:

$[_{C-MAX}\ [_{I-MAX}\ e_i\ [_{V-MAX}\ [_{V-MAX}\ [_{V}\ \text{me gusta}]]\ [_{N-MAX}\ \text{María}]_i]]]$.

However, this does not affect the internal/external reversal scheme described here, since inversion is a syntactic operation that takes place independently of the process that handles thematic divergences.

The general solution to thematic divergence is diagrammed as follows:

(14) RLCS 1: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])]$
RLCS 2: $[_{T(X')}\ X'\ ([_{T(W')}\ \text{:INT}\ W'],\ [_{T(Z')}\ \text{:EXT}\ Z']\ [_{T(Q')}\ Q'])]$

Translation: $[_{Y-MAX}\ [_{Y-MAX}\ W\ [_{X-MAX}\ X\ Z]]\ Q]$

$$\Updownarrow$$

$$[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])]$$

$$\Updownarrow$$

$$[_{Y-MAX}\ [_{Y-MAX}\ Z\ [_{X-MAX}\ X\ W]]\ Q]$$

This assumes that there is only one external argument and zero or more internal arguments. If the situation arises in which more than one variable is associated with the :EXT markers, it is assumed that there is an error in the word definition.[11] Note that the :INT and :EXT markers show up only in the RLCS. The CLCS does not include any such markers, since it is intended to be a language-independent representation for the source- and target-language sentences.

Thematic divergence is one of three types of possible positioning variations that force the $\mathcal{GLR}$ to be overridden. Two additional positioning variations are promotional and demotional divergences, which will be defined in the next two sections. Whereas thematic divergence involves a repositioning of two satellites relative to a head, promotional and demotional divergences involve a repositioning of the head itself.[12] We will see in Section 5.1 that these three divergences account for the entire range of repositioning possibilities.

4.2 Promotional Divergence

Promotional divergence is characterized by the *promotion* (placement "higher up") of a logical modifier into a main verb position (or vice versa), as shown in Figure 9b. In such a situation, the logical modifier is associated with the syntactic head position, and the logical head is then associated with an internal argument position. Thus, promotional divergence overrides the $\mathcal{GLR}$, invoking the following sets of relations in place of steps 1 and 4 of (10):

(15) 1.′ $X' \Leftrightarrow Z$[13]
4.′ $Q' \Leftrightarrow X$

Figure 9b shows the revised mapping.

11 The parameters associated with the RLCS are assumed to be correctly specified for the purposes of this formal description. However, in practice, there might be errors in the lexical entries, since they are constructed by hand in the current implementation. Eventually, the intent is to automate the process of lexical entry construction so that these errors can be avoided.

12 The notions of demotion and promotion are not the same as the notions of demotion and advancement in the theory of relational grammar (see Perlmutter 1983). Dorr (1993b, pp. 269–274) argues that, although the relational representation might be a convenient tool for illustrating the promotion and demotion operations as used in the current approach, this representation is not an appropriate vehicle for interlingual translation for a number of reasons.

13 This relation does not mean that X replaces Z (if there is a Z), but that X retains the same structural relation with Z (i.e., Z remains an internal argument of X). To simplify the current description, Z is not shown in the syntactic structure of Figure 9b.

An example of promotional divergence is the case given earlier in (2). The syntactic structures and corresponding CLCS are shown here:

(16) $[_{\text{C-MAX}}$ $[_{\text{I-MAX}}$ $[_{\text{N-MAX}}$ John]
$[_{\text{V-MAX}}$ $[_{\text{V}}$ usually $[_{\text{V}}$ goes]] $[_{\text{N-MAX}}$ home]]]]
⇕
$[_{\text{Event}}$ GO_{Loc} ($[_{\text{Thing}}$ JOHN],
$[_{\text{Path}}$ TO_{Loc} ($[_{\text{Position}}$ AT_{Loc} ($[_{\text{Thing}}$ JOHN], $[_{\text{Location}}$ HOUSE])])],
$[_{\text{Manner}}$ HABITUALLY])]
⇕
$[_{\text{C-MAX}}$ $[_{\text{I-MAX}}$ $[_{\text{N-MAX}}$ Juan]
$[_{\text{V-MAX}}$ $[_{\text{V}}$ suele] $[_{\text{V-MAX}}$ $[_{\text{V}}$ ir] $[_{\text{P-MAX}}$ a casa]]]]]

Here the main verb *go* is modified by an adverbial adjunct *usually*, but in Spanish, *usually* has been placed into a higher position as the main verb *soler*, and the "going home" event has been realized as the internal argument of this verb.

Promotional divergence is resolved by the :PROMOTE parameter, which forces the $\mathcal{GLR}$ mapping to be overridden with respect to the positioning of the logical head and the logical modifier. The lexical entries for *usually* and *soler* illustrate the difference in the use of this parameter:

(17) (i) **Lexical entry for *usually*:**
$[_{\text{Manner}}$ HABITUALLY]

(ii) **Lexical entry for *soler*:**
$[_{\text{Manner}}$:PROMOTE HABITUALLY]

Because the English entry does not use this parameter, the translation relies on the default argument positionings imposed by the $\mathcal{GLR}$. By contrast, the :PROMOTE marker specified in the Spanish entry forces the head and adjunct to swap places in the syntactic structure.

The general solution to promotional divergence is diagrammed as follows:

(18) RLCS 1: $[_{T(Q')}$ Q']
RLCS 2: $[_{T(Q')}$:PROMOTE Q']

Translation: $[_{\text{Y-MAX}}$ $[_{\text{Y-MAX}}$ W $[_{\text{X-MAX}}$ X Z]] Q]
⇕
$[_{T(X')}$ X' ($[_{T(W')}$ W'], $[_{T(Z')}$ Z'] $[_{T(Q')}$ Q'])]
⇕
$[_{\text{Y-MAX}}$ $[_{\text{Y-MAX}}$ W $[_{\text{X-MAX}}$ Q [... X Z]]]]

4.3 Demotional Divergence

Demotional divergence is characterized by the *demotion* (placement "lower down") of a logical head into an internal argument position (or vice versa), as shown in Figure 9c. In such a situation, the logical head is associated with the syntactic adjunct position, and the logical argument is then associated with a syntactic head position. Thus, demotional divergence overrides the $\mathcal{GLR}$, invoking the following sets of relations in place of steps 1 and 3 of (10):

(19) 1.′ $X' \Leftrightarrow Q$[14]
3.′ $Z' \Leftrightarrow X$

Figure 9(c) shows the revised mapping.

An example of demotional divergence is the case given earlier in (3). The syntactic structures and corresponding CLCS are shown here:[15]

(20) $[_{\text{C-MAX}}\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}\ \text{I}]_i\ [_{\text{V-MAX}}\ [_{\text{V}}\ \text{like}]\ [_{\text{C-MAX}}\ \text{PRO}_i\ \text{to eat}]]]]$

$\Updownarrow$

$[_{\text{State}}\ \text{BE}_{\text{Circ}}\ ([_{\text{Thing}}\ \text{I}],$
$[_{\text{Position}}\ \text{AT}_{\text{Circ}}\ ([_{\text{Thing}}\ \text{I}],\ [_{\text{Event}}\ \text{EAT}\ ([_{\text{Thing}}\ \text{I}],\ [_{\text{Thing}}\ \text{FOOD}])])]$
$[_{\text{Manner}}\ \text{LIKINGLY}])]$

$\Updownarrow$

$[_{\text{C-MAX}}\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}\ \text{Ich}]\ [_{\text{V-MAX}}\ [_{\text{V}}\ [_{\text{ADV}}\ \text{gern}]\ [_{\text{V}}\ \text{esse}]]]]]$[16]

Here the main verb *like* takes the "to eat" event as an internal argument; but in German, *like* has been placed into a lower position as the adjunct *gern*, and the "eat" event has been realized as the main verb.

The distinction between promotional and demotional divergences may not be intuitively obvious at first glance. In both cases, the translation mapping appears to associate a main verb with an adverbial satellite, or vice versa. However, the distinction between these two head switching cases becomes more apparent when we consider the status of the participating lexical tokens more carefully. In the case of *soler–usually*, the main verb *soler* is, in some sense, the token that "triggers" the head switching operation: its presence forces the adverbial satellite *usually* to appear in English, even if we were to substitute some other event for *ir* in Spanish (e.g., *correr a la tienda, leer un libro*, etc.). By contrast, in the case of *like–gern*, the triggering element is not the main verb *like*, since we are able to use *like* in other contexts that do not require *gern* (e.g., *I like the car* ⇔ *Mir gefällt der Wagen*); instead, the triggering element is the adverbial satellite *gern*: its presence forces the verb *like* to appear in English even if we were to substitute some other event in place of *essen* in German (e.g., *zum Geschäft laufen, das Buch lesen*, etc.). We will return to this point in Section 5.2.

Demotional divergence is resolved by the :DEMOTE parameter, which forces the $\mathcal{GLR}$ mapping to be overridden with respect to the positioning of the logical head and the logical argument. The lexical entries for *like* and *gern* illustrate the difference in the use of this parameter:[17]

14 This relation does not mean that X replaces Q (if there is a Q), but that X retains the same structural relation with Q (i.e., Q remains a syntactic adjunct of X). To simplify the current description, Q is not shown in the syntactic structure of Figure 9c.

15 The default object being eaten is FOOD, although this argument does not appear on the surface for the current example.

16 The German syntactic structure is shown here in the uninverted base form. In the German surface structure, the verb is moved up into verb-second position and the subject is topicalized:

$[_{\text{C-MAX}}\ [_{\text{N-MAX}}\ \text{Ich}]_i\ [_{\text{V}}\ \text{esse}]_j\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}\ \text{t}]_i\ [_{\text{V-MAX}}\ [_{\text{V}}\ [_{\text{ADV}}\ \text{gern}]\ [_{\text{V}}\ \text{t}]_j]]]]$.

17 Both definitions of *like* in (21) use the circumstantial field, which means that the Y argument must be an Event (e.g., *like to eat*) rather than a Thing (e.g., *like Mary*). Thus, the definitions for *like* and *gern* are slightly different from the definitions of *like* given earlier in (13) (i.e., these are additional lexical entries for *like*).

(21) (i) **Lexical entry for *like*:**

$$[_{\text{State}}\ \text{BE}_{\text{Circ}}\ ([_{\text{Thing}}\ \text{W}],\ [_{\text{Position}}\ \text{AT}_{\text{Circ}}\ ([_{\text{Thing}}\ \text{W}],\ [_{\text{Event}}\ \text{Z}])],\ [_{\text{Manner}}\ \text{LIKINGLY}])]$$

(ii) **Lexical entry for *gern*:**

$$[_{\text{State}}\ \text{BE}_{\text{Circ}}\ ([_{\text{Thing}}\ \text{W}],\ [_{\text{Position}}\ \text{AT}_{\text{Circ}}\ ([_{\text{Thing}}\ \text{W}],\ [_{\text{Event}}\ \text{:DEMOTE Z}])],\ [_{\text{Manner}}\ \text{LIKINGLY}])]$$

Because the English entry does not use this parameter, the translation relies on the default argument positionings imposed by the $\mathcal{GLR}$. By contrast, the :DEMOTE marker specified in the German entry forces the head and internal argument to swap places in the syntactic structure.

The general solution to demotional divergence is diagrammed as follows:

(22) RLCS 1: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])]$
RLCS 2: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ \text{:DEMOTE}\ Z']\ [_{T(Q')}\ Q'])]$

Translation: $[_{Y-MAX}\ [_{Y-MAX}\ W\ [_{X-MAX}\ X\ Z]]\ Q]$

$$\Updownarrow$$

$$[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])]$$

$$\Updownarrow$$

$$[_{Y-MAX}\ [_{Y-MAX}\ W\ [_{X-MAX}\ Z]\ [\ \ldots\ X\ Q]]]$$

4.4 Structural Divergence

Structural divergence differs from the last three divergence types in that it does not alter the positions used in the $\mathcal{GLR}$ mapping, but it changes the nature of the relation between the different positions (i.e., the "$\Leftrightarrow$" correspondence). Figure 9d characterizes the alteration that takes place. Note that the mapping of Z' to the corresponding internal argument position is altered so that it is positioned under the constituent that corresponds to R'.

An example of structural divergence is the case given earlier in (4). The syntactic structures and corresponding CLCS are shown here:

(23) $[_{C-MAX}\ [_{I-MAX}\ [_{N-MAX}\ \text{John}]\ [_{V-MAX}\ [_{V}\ \text{entered}]\ [_{N-MAX}\ \text{the house}]]]]$

$$\Updownarrow$$

$$[_{\text{Event}}\ \text{GO}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{JOHN}],\ [_{\text{Path}}\ \text{TO}_{\text{Loc}}\ ([_{\text{Position}}\ \text{IN}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{JOHN}],\ [_{\text{Location}}\ \text{HOUSE}])])])]$$

$$\Updownarrow$$

$$[_{C-MAX}\ [_{I-MAX}\ [_{N-MAX}\ \text{Juan}]\ [_{V-MAX}\ [_{V}\ \text{entró}]\ [_{P-MAX}\ \text{en}\ [_{N-MAX}\ \text{la casa}]]]]]$$

Here the verbal object is realized as a noun phrase (*the house*) in English and as a prepositional phrase (*en la casa*) in Spanish.

Structural divergence is resolved by means of the $*$ marker, which forces logical constituents to be realized compositionally at different levels. In particular, the $*$ serves as a pointer to a RLCS position that must be combined with another RLCS in order to arrive at a (portion of a) CLCS. The lexical entries for *enter* and *entrar* illustrate the difference in the use of this parameter:

(24) (i) **Lexical entry for *enter*:**
$[_{Event}\ GO_{Loc}\ ([_{Thing}\ W],$
$[_{Path}\ TO_{Loc}\ ([_{Position}\ IN_{Loc}\ ([_{Thing}\ W],\ [_{Location}\ *\ Z])])])]$

(ii) **Lexical entry for *entrar*:**
$[_{Event}\ GO_{Loc}\ ([_{Thing}\ W],$
$[_{Path}\ *\ TO_{Loc}\ ([_{Position}\ IN_{Loc}\ ([_{Thing}\ W],\ [_{Location}\ Z])])])]$

Because the English entry contains a $*$ marker in the $[_{Location}\ Z]$ position, this constituent is realized on the surface as the object (i.e., *the house*) of the main verb. By contrast, the $*$ marker is associated with a "higher" position $[_{Path}\ TO_{Loc}\ \ldots]$ in the Spanish entry, thus forcing this constituent to have a more complex realization (i.e., *en la casa*) in the syntactic structure.

The general solution to structural divergence is diagrammed as follows:

(25) RLCS 1: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(R')}\ R'\ ([_{T(Z')}\ *\ Z'])]\ [_{T(Q')}\ Q'])]$
RLCS 2: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(R')}\ *\ R'\ ([_{T(Z')}\ Z'])]\ [_{T(Q')}\ Q'])]$

Translation: $[_{Y-MAX}\ [_{Y-MAX}\ W\ [_{X-MAX}\ X\ Z]]\ Q]$
$\Updownarrow$
$[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])]$
$\Updownarrow$
$[_{Y-MAX}\ [_{Y-MAX}\ W\ [_{X-MAX}\ X\ [\ \ldots\ R\ Z]]]\ Q]$

Note that the logical argument R' is associated with a $*$ marker in the RLCS of the target language, but not in the RLCS of the source language. This forces the target language syntactic structure to realize a phrase R that dominates Z; in contrast, no such dominating phrase occurs in the source-language structure.

4.5 Conflational Divergence

Conflational divergence is another case in which the "$\Leftrightarrow$" correspondence is changed. In particular, conflational divergence is characterized by the suppression of a CLCS constituent (or the inverse of this process). The constituent generally occurs in logical argument or logical modifier position; thus, the "$\Leftrightarrow$" correspondence of either step 3 or step 4 of the $\mathcal{GLR}$ is changed, depending on which position is conflated. Figure 9e characterizes the alteration that takes place. Note that the Z' position in the CLCS does not have a corresponding realization in the syntax.

An example of conflational divergence is the case given earlier in (5). The syntactic structures and corresponding LCS are shown here:

(26) $[_{C-MAX}\ [_{I-MAX}\ [_{N-MAX}\ I]$
$[_{V-MAX}\ [_{V}\ stabbed]\ [_{N-MAX}\ John]]]]$
$\Updownarrow$
$[_{Event}\ CAUSE$
$([_{Thing}\ I],$
$[_{Event}\ GO_{Poss}$
$([_{Thing}\ KNIFE\text{-}WOUND],$
$[_{Path}\ TOWARD_{Poss}$
$([_{Position}\ AT_{Poss}\ ([_{Thing}\ KNIFE\text{-}WOUND],\ [_{Thing}\ JOHN])])])])]$
$\Updownarrow$
$[_{C-MAX}\ [_{I-MAX}\ [_{N-MAX}\ Yo]$
$[_{V-MAX}\ [_{V}\ le\ di]\ [_{N-MAX}\ puñaladas]\ [_{P-MAX}\ a\ Juan]]]]$

Here, English uses the single word *stab* for the two Spanish words *dar* (*give*) and *puñaladas* (*knife-wounds*); this is because the effect of the action (i.e., the *knife-wound* portion of the lexical token) is incorporated into the main verb in English.

Conflational divergence is resolved by means of the :CONFLATED marker, which suppresses the realization of the filler of a particular position. The lexical entries for *stab* and *dar* illustrate the difference in the use of this parameter:[18]

(27) (i) **Lexical entry for *stab*:**

$[_{\text{Event}}$ CAUSE
 $([_{\text{Thing}}$ W$]$,
  $[_{\text{Event}}$ GO_{Poss}
   $([_{\text{Thing}}$:CONFLATED KNIFE-WOUND$]$,
    $[_{\text{Path}}$ $\text{TOWARD}_{\text{Poss}}$
     $([_{\text{Position}}$ AT_{Poss} $([_{\text{Thing}}$ KNIFE-WOUND$]$, $[_{\text{Thing}}$ Z$])])])])$]$

(ii) **Lexical entry for *dar*:**

$[_{\text{Event}}$ CAUSE
 $([_{\text{Thing}}$ W$]$,
  $[_{\text{Event}}$ GO_{Poss}
   $([_{\text{Thing}}$ $*$ Y$]$,
    $[_{\text{Path}}$ $\text{TOWARD}_{\text{Poss}}$
     $([_{\text{Position}}$ AT_{Poss} $([_{\text{Thing}}$ Y$]$, $[_{\text{Thing}}$ Z$])])])])$]$

Because the English entry contains a :CONFLATED marker in the logical position corresponding to KNIFE-WOUND, this constituent is suppressed in the syntactic structure. By contrast, this marker does not appear in the corresponding position in Spanish, thus forcing this constituent to be overtly realized (i.e., *puñaladas*) in the syntactic structure.

The general solution to conflational divergence is diagrammed as follows:

(28) RLCS 1: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ *\ Z']\ [_{T(Q')}\ Q'])]$
RLCS 2: $[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ \text{:CONFLATED}\ Z']\ [_{T(Q')}\ Q'])]$

Translation: $[_{\text{Y-MAX}}\ [_{\text{Y-MAX}}\ W\ [_{\text{X-MAX}}\ X\ Z]]\ Q]$

$$\Updownarrow$$

$$[_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])]$$

$$\Updownarrow$$

$$[_{\text{Y-MAX}}\ [_{\text{Y-MAX}}\ W\ [_{\text{X-MAX}}\ X]]\ Q]$$

Note that the logical argument Z′ is associated with a :CONFLATED marker in the RLCS of the target language, but not in the RLCS of the source language. This forces the target language syntactic structure to suppress the realization of this constituent.

18 Note that the :CONFLATED marker appears to be in complementary distribution with the $*$ marker. In fact, one might consider the use of the :CONFLATED marker to be unnecessary, since its presence could be implied by the absence of the $*$ marker. However, the :CONFLATED marker plays an important role in the lexical-semantic representation: it specifies that the "constant" term (e.g., the KNIFE-WOUND of the *stab* RLCS) must obligatorily fill the position and, moreover, that this constant must be a legal LCS primitive of the system. In addition, there is an inherent asymmetry between the :CONFLATED marker and the $*$ marker: whereas the former always occurs in a leaf node position, the latter may occur in *any* position in the RLCS. Because the notion of *conflation* is not meaningful in non-leaf node positions, it would be unreasonable to make the assumption that every non-leaf position without the $*$ marker is conflated. The :CONFLATED marker is used to identify truly conflated positions, not just those positions without the $*$ marker.

4.6 Categorial Divergence

Unlike the previous five divergence types, categorial divergence affects the operation of the $\mathcal{CSR}$, not the $\mathcal{GLR}$. It is characterized by a situation in which CAT(ϕ) is forced to have a different value than would normally be assigned to T(ϕ') by means of the mapping specified in Figure 7. Thus, categorial divergence is formally described as follows: a lexical–semantic type T(ϕ') is related to a syntactic category CAT(ϕ), where CAT(ϕ) $\neq$ $\mathcal{CSR}$(T(ϕ')). In such a case, CAT(ϕ) must be specified through lexical parameterization.

An example of categorial divergence is the case given earlier in (6). The syntactic structures and corresponding CLCS are shown here:

(29) $[_{\text{C-MAX}}\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}$ I] $[_{\text{V-MAX}}\ [_{\text{V}}$ am] $[_{\text{A-MAX}}$ hungry]]]]

$\Updownarrow$

$[_{\text{State}}$ BE_{Ident} ($[_{\text{Thing}}$ I],
$[_{\text{Position}}$ AT_{Ident} ($[_{\text{Thing}}$ I], $[_{\text{Property}}$ HUNGRY])])]

$\Updownarrow$

$[_{\text{C-MAX}}\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}$ Ich] $[_{\text{V-MAX}}\ [_{\text{N-MAX}}$ Hunger] $[_{\text{V}}$ habe]]]][19]

Here, the predicate is adjectival (*hungry*) in English but nominal (*Hunger*) in German.

Categorial divergence is resolved by means of the :CAT parameter. The lexical entries for *be* and *haben* illustrate the difference in the use of this parameter:[20]

(30) (i) **Lexical entry for *be*:**
$[_{\text{State}}$ BE_{Ident} ($[_{\text{Thing}}$ W],
$[_{\text{Position}}$ AT_{Ident} ($[_{\text{Thing}}$ W], $[_{\text{Property}}$ Y])])]

(ii) **Lexical entry for *haben*:**
$[_{\text{State}}$ BE_{Ident} ($[_{\text{Thing}}$ W],
$[_{\text{Position}}$ AT_{Ident} ($[_{\text{Thing}}$ W], $[_{\text{Property}}$:CAT(N) Y])])]

Because the English entry does not contain a :CAT marker in the position corresponding to $[_{\text{Property}}$ Y], this constituent is realized in the syntactic structure as $\mathcal{CSR}$(Property) $\rightarrow$ Adjective (i.e., *hungry*). By contrast, the :CAT(N) marker specified in the German entry forces the default $\mathcal{CSR}$ category to be overridden and the Property argument is realized as a noun form (i.e., *Hunger*).

19 The German structure shown here is the base form of the following surface syntactic tree:

$[_{\text{C-MAX}}\ [_{\text{N-MAX}}$ Ich$]_i$
$[_{\text{C}}\ [_{\text{V}}$ habe$]_j]$
$[_{\text{I-MAX}}\ [_{\text{N-MAX}}$ t$]_i$ $[_{\text{V-MAX}}\ [_{\text{N-MAX}}$ Hunger] $[_{\text{V}}\ t]_j]$ $[_{\text{I}}\ e]]]$

This form is derived by the syntactic processor *after* lexical-semantic processing is complete (see Dorr 1993a, for details).

20 As might be expected, there are two lexical entries for the verb *haben*, only one of which is listed here. The one not shown here corresponds to the possessional sense of *haben*, i.e., the meaning corresponding to the word *have* in English. This entry does not contain a :CAT marker.

The general solution to categorial divergence is diagrammed as follows:

(31)

$$
\begin{array}{l}
\text{RLCS 1: } [_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ {:}Z']\ [_{T(Q')}\ Q'])] \\
\text{RLCS 2: } [_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ ({:}\text{CAT}\ \delta)\ Z']\ [_{T(Q')}\ Q'])] \\
\\
\text{Translation: } [_{\text{Y-MAX}}\ [_{\text{Y-MAX}}\ W\ [_{\text{X-MAX}}\ X\ Z]]\ Q] \\
\qquad\qquad\qquad\Updownarrow \\
\qquad [_{T(X')}\ X'\ ([_{T(W')}\ W'],\ [_{T(Z')}\ Z']\ [_{T(Q')}\ Q'])] \\
\qquad\qquad\qquad\Updownarrow \\
\qquad [_{\text{Y-MAX}}\ [_{\text{Y-MAX}}\ W\ [_{\text{X-MAX}}\ X\ Z]]\ Q] \\
\text{where CAT(Z)} = \delta\text{-MAX.}
\end{array}
$$

4.7 Lexical Divergence

Lexical divergence arises only in the context of other divergence types.[21] This is because the choice of lexical items in any language relies crucially on the realization and composition properties of those lexical items. Because the six preceding divergences potentially alter these properties, lexical divergence is viewed as a side effect of other divergences. Thus, the formalization thereof is considered to be some combination of those given above.

Unlike the first six divergence types, lexical divergence is solved during the process of lexical selection.[22] Thus, there is no specific override marker that is used for this type of divergence. For example, in the lexical divergence (7), a conflational divergence forces the occurrence of a lexical divergence. The syntactic structures and corresponding CLCS for this example are shown here:

(32)

$$
\begin{array}{l}
[_{\text{C-MAX}}\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}\ \text{John}] \\
\qquad\qquad [_{\text{V-MAX}}\ [_{\text{V}}\ \text{broke}]\ [_{\text{P-MAX}}\ \text{into}\ [_{\text{N-MAX}}\ \text{the room}]]]]] \\
\qquad\qquad\qquad\Updownarrow \\
[_{\text{Event}}\ \text{CAUSE} \\
\quad ([_{\text{Thing}}\ \text{JOHN}], \\
\quad [_{\text{Event}}\ \text{GO}_{\text{Loc}} \\
\qquad ([_{\text{Thing}}\ \text{JOHN}], \\
\qquad [_{\text{Path}}\ \text{TO}_{\text{Loc}} \\
\qquad\quad ([_{\text{Position}}\ \text{IN}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{JOHN}],\ [_{\text{Location}}\ \text{ROOM}])])])] \\
\quad [_{\text{Manner}}\ \text{FORCEFULLY}])] \\
\qquad\qquad\qquad\Updownarrow \\
[_{\text{C-MAX}}\ [_{\text{I-MAX}}\ [_{\text{N-MAX}}\ \text{Juan}] \\
\qquad\qquad [_{\text{V-MAX}}\ [_{\text{V}}\ \text{forzó}]\ [_{\text{N-MAX}}\ \text{la entrada}]\ [_{\text{P-MAX}}\ \text{al cuarto}]]]]
\end{array}
$$

Because the word–particle pair *break into* subsumes two concepts (forceful spatial motion and entry to a location), it is crucial that the word *forzar* (literally, *force*) be selected in conjunction with *entrada* (literally, *entry*) for the underlying break-into concept.

---

21 As noted by a reviewer, this is not strictly true, since there are many cases in which a source-language word maps to more than one target-language word without the simultaneous occurrence of another divergence type. An example of such a case is the English word *eat*, which maps to *essen* (for humans) or *fressen* (for animals). Such cases are considered to be outside of the classes of lexical–semantic divergences considered here (see Footnote 3). However, a simple approach to resolving such cases would be to use featural restrictions during syntactic processing.

22 The solution to lexical divergence is trivial for transfer machine translation systems, since transfer entries map source-language words directly to their target-language equivalents. In general, lexical selection is not seen as a problem in these systems.

There is also a structural divergence in this example, since the prepositional phrase *into the room* must be translated into a noun phrase *entrada al cuarto*. This divergence compounds the lexical divergence problem, since it is necessary to choose the target-language word *a* in the absence of a source-language counterpart.

Lexical divergence also shows up in three previously presented examples, (12), (29), and (26), owing to the presence of thematic, categorial, and conflational divergences, respectively: in (12) the word *like* is chosen for the word *gustar* (literally, *to please*); in (29) the word *haben* (literally, *to have*) is chosen for the word *be*; and in (26) the word *dar* (literally, *to give*) is chosen for the word *stab*.

5. Discussion

This section discusses certain issues of relevance to the formal classification and resolution of translation divergences. In particular, we will discuss (1) the limits imposed on the range of repositioning possibilities; (2) the justification for distinguishing between promotional and demotional divergences; (3) the notion of full coverage in the context of lexical selection; and (4) the resolution of interacting divergence types.

5.1 Limits on Repositioning Divergences

In Section 4.1 we made the claim that the thematic, promotional, and demotional divergences account for the entire range of repositioning possibilities. We will now explore the validity of this claim.

There are two potential types of syntactic relations that exist between a head and a satellite: the first is complementation (i.e., involving the internal argument), and the second is adjunction.[23] Given these two types of relations, there are only a small number of ways syntactic entities may be repositioned. The three CLCS positions that are involved in these relations are X', Z', and Q'. If we compute the repositionings combinatorically, there are $3^3 = 27$ configurations (i.e., X', Z', and Q' would map into any of three positions). However, we can eliminate 15 of these (since a CLCS must contain exactly one head), thus leaving only 12 possible configurations. One of these corresponds to the default $\mathcal{GLR}$ mapping (i.e., the logical head, logical argument, and logical modifier map into canonical positions). The remaining 11 configurations can be factored into three cases as follows:

1. $X' \Leftrightarrow X$

    1.1 $Q' \Leftrightarrow Z; Z' \Leftrightarrow Z.$
    1.2 $Z' \Leftrightarrow Q; Q' \Leftrightarrow Q.$
    1.3 $Q' \Leftrightarrow Z; Z' \Leftrightarrow Q.$

2. $Q' \Leftrightarrow X$

    2.1 $X' \Leftrightarrow Z; Z' \Leftrightarrow Z.$
    2.2 $X' \Leftrightarrow Z; Z' \Leftrightarrow Q.$

23 We have left out the possibility of an external argument as a participant in the head-satellite relation. Of course, the external argument *is* a satellite with respect to the head, but it turns out that the external argument, which corresponds to the logical subject in the CLCS, has a special status and does not have the same repositioning potential that internal arguments and syntactic adjuncts have. In particular, the external argument has the unique property that it never participates as the incorporated argument of a conflational verb. Hale and Keyser (1989) provide evidence that this property holds across all languages. Thus, we take the external argument to have a special status (universally) that exempts it from participating in divergences other than thematic divergence.

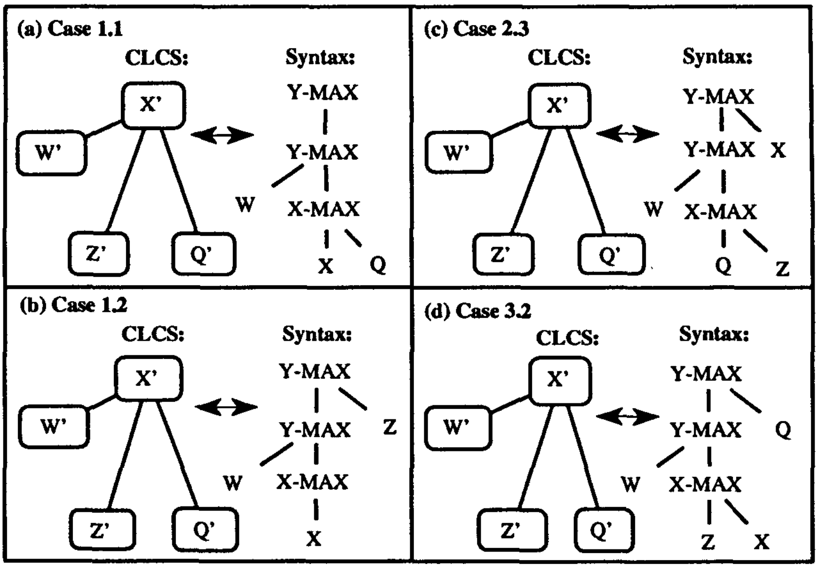

Figure 10
Illegal translation mappings for natural language.

2.3 $X' \Leftrightarrow Q$; $Z' \Leftrightarrow Z$.
2.4 $X' \Leftrightarrow Q$; $Z' \Leftrightarrow Q$.

3. $Z' \Leftrightarrow X$

3.1 $X' \Leftrightarrow Z$; $Q' \Leftrightarrow Z$.
3.2 $X' \Leftrightarrow Z$; $Q' \Leftrightarrow Q$.
3.3 $X' \Leftrightarrow Q$; $Q' \Leftrightarrow Z$.
3.4 $X' \Leftrightarrow Q$; $Q' \Leftrightarrow Q$.

We will discuss in detail how four of these cases, i.e., the ones characterized in Figure 10, are ruled out. Of the remaining, 2.1 and 3.4 correspond to the definitions of promotional and demotional divergences illustrated in Figures 9b and 9c, respectively; cases 1.3, 2.2, 2.4, 3.1, and 3.3 are ruled out for the same reasons that cases 1.1 and 1.2 are ruled out (as we will see shortly), namely, that an internal argument Z can never be associated with a logical modifier Q′ and that a syntactic adjunct (Q) can never be associated with a logical argument Z′.

It cannot be the case that a logical modifier maps to an internal argument position (case 1.1) because a logical modifier is an *optional* participant of a particular action, i.e., it need not be "governed" by the lexical item that it modifies. Internal argument positions are reserved for cases in which a government relation must hold; thus, logical modifiers must necessarily be mapped into syntactic adjunct positions.

Similarly, it also cannot be the case that a logical argument maps to a syntactic adjunct position (case 1.2). A logical argument is a *necessary* participant of a particular action, and as such, it must be "governed" by the lexical item that selects it. By contrast, adjunct positions are reserved for *optional* modifying participants that do not need to be governed by the lexical item that they are modifying; thus, logical arguments must necessarily be mapped into internal argument positions.

Another case that is eliminated is the renaming of a logical head as a syntactic adjunct whose head corresponds to a logical modifier (case 2.3). The idea is simply that modification is a one-way relation. If a logical head has a modifier, the head cannot become an adjunct of that modifier because the modifying relation would be reversed (i.e., the logical head would modify the syntactic head rather than the other way around). In contrast, a logical head of a CLCS *can* be mapped to an internal argument position in cases in which a logical modifier is mapped to a syntactic head (i.e., the case of promotional divergence presented earlier), since there is no violation of the one-way relation.

A similar argument is used to eliminate the case in which a logical head is mapped to an internal argument whose head corresponds to a logical argument (case 3.2). The idea is that heads and arguments participate in a one-way relation. If a logical head has an argument, the head cannot become an internal argument of that argument because the head-argument relation would be reversed (i.e., the logical head would be an argument of the syntactic head rather than the other way around). In contrast, the logical head of a CLCS *can* be mapped to an adjunct (i.e., modifier) position in cases in which a logical argument is mapped to a syntactic head (i.e., the case of demotional divergence presented earlier), since there is no violation of the one-way relation.

The argument for the elimination of the last two cases could be viewed as an appeal to a constraint that is analogous to the θ-criterion in syntax. Essentially, this constraint states that all arguments and modifiers must be *licensed* (see Chomsky 1986a; Abney 1989) in order to appear either in the syntactic structure or in the conceptual structure. In the context of conceptual structure, a logical modifier may license the realization of a logical head in an internal argument position, but not in an adjunct position, since the modifier relation is already satisfied by virtue of the relation between the head and the modifier. Similarly, a logical argument may license the realization of a logical head in a syntactic adjunct position, but not in an internal argument position, since the head–argument relation is already satisfied by virtue of the relation between the head and the argument. Having eliminated the meaningless possibilities, we are left with the promotional and demotional cases presented above.

5.2 Promotional versus Demotional Divergences

We will now provide justification for the earlier claim that promotional and demotional divergence should be classified differently, even though they exhibit some of the same properties. It might be argued that these divergences are essentially the same, since both cases involve an association of a main verb with an adverbial satellite or vice versa. In the examples given earlier, the *promotional* divergence referred to a mapping between the adverbial *usually* and the main verb *soler* and *demotional* divergence referred to a mapping between the adverbial *gern* and the main verb *like*. However, as mentioned in Section 4.3, these are taken to be in distinct classes: the difference between these two cases is determined by the "triggering" element (i.e., promotion is triggered by a main verb such as *soler*, whereas demotion is triggered by an adverb such as *gern*).

Another factor that distinguishes between promotional and demotional divergences is the fact that verbs such as *like* and verbs such as *soler* do not have parallel syntactic distributions, nor do they have analogous logical interpretations. The verb *like* may take a sentential complement that has its own event structure (as in *I like to eat*), or it may take a nominal complement without an event structure (as in *I like the car*). In either case, the verb *like* generally means the same thing (i.e., it describes a state in which an event or a thing is somehow desirable to that person). By contrast, the verb *soler* is a modal verb that contributes an aspectual component of meaning that crucially relies on a verbal complement with an event structure; in a sense, *soler* is analogous to the modal *must* in English in that it cannot be used in isolation, but requires the presence of a verbal complement in order for it to be interpretable. In such a configuration, the modal *soler* allows the event to be interpreted as being habitual in nature.

Given these distinctions, it would not be appropriate to consider the head switching mapping to be the same for the *soler–usually* and *like–gern* cases. Not only do they have different triggering elements (i.e., the main verb in the former and the adverb in the latter), but they do not have identical syntactic distributions and their logical interpretations are not analogous. Thus, they are taken to be two independent mappings with entirely different syntactic and lexical-semantic ramifications.

The handling of promotional and demotional divergences is a topic that has received recent attention, although it has been labeled differently, depending on how it is manifested. An example of such a case is the *way* construction. This phenomenon has been studied by Jackendoff (1990) in his extended version of the original LCS framework:

(33) Bill belched his way out of the restaurant

In such cases, Jackendoff claims that *belching* is subordinated to a higher predicate like GO, "in effect demoting the meaning of the lexical verb to a subordinate accompaniment or means modifier" (Jackendoff 1990, p. 214). This characterization is essentially equivalent to that of the *soler–usually* example (i.e., promotional divergence) given earlier.[24]

5.3 Lexical Selection: Full Coverage Constraint

Because of the compositional nature of the LCS representation, the current framework automatically imposes a full coverage constraint during the lexical selection process. Formally, this constraint is defined as follows:

(34) **Full coverage constraint:**
A RLCS R matches a CLCS C if and only if R **fully covers** C.

where R *fully covers* C under the following conditions:

(35) A RLCS R **fully covers** a CLCS C if and only if:

(a) there is no portion of R that does not match C;

24 Jackendoff's approach to handling the *way* construction has been criticized by Marantz (1992) for its use of arbitrary exceptions to the "usual mappings." Marantz takes issue with the characterization of such cases as an idiosyncratic relation between syntax and semantics and proposes, instead, that the conceptual structure looks different from what Jackendoff envisions. Whichever of these proposals is correct, neither Jackendoff nor Marantz considers their proposals in the context of interlingual machine translation. If the exceptional mappings are indeed arbitrary, then one needs to explain how this affects the handling of different languages. Moreover, neither Jackendoff nor Marantz mentions the possibility that the number of exceptional mappings might not be arbitrarily large, but that there might be a fixed number of exceptions, delineated in such a way that only a handful need to be considered at any time for any given language. This is why the formalization described in this paper is a valuable resource: it provides a means for proving that only certain types of exceptions are allowed and that the number of such exceptions is actually quite small. Finally, neither Jackendoff nor Marantz considers keeping Jackendoff's version of the LCS intact and using a single parameterized mapping along the lines proposed in the current framework.

(b) either R completely matches C (i.e., there is no portion of C that does not match R) or R matches all of C except some portion C′ (i.e., a subcomponent of C) that is fully covered by some other RLCS R′.

In cases in which more than one lexical entry matches a current concept, this constraint is used to determine which possibilities should be ruled out (if any).

One of the main advantages to the formalization defined in this paper is that it allows one to judge whether a target-language concept *fully covers* the concept underlying the source-language sentence and, thus, to make an evaluation of the status of the system. As an illustration of this point, consider the *stab–dar* example of Section 4.5. The notion of *full coverage* is manifested through the lexical-selection process. The basic idea is that RLCSs are chosen such that they entirely cover the underlying concept. In terms of the mapping from the source-language sentence to the CLCS, this implies that the RLCSs must compose in such a way as to provide a full cover (i.e., there must be a path from the root to the leaf nodes that includes all the [content] words of the sentence). In terms of the mapping from the CLCS to the target-language sentence, this implies that the CLCS must be decomposed into (potentially overlapping) RLCSs whose "union" covers the entire CLCS.

From the English sentence in (26), the RLCSs must be chosen for each word in the sentence such that they provide a coherent CLCS. The RLCS for *stab* is[25]

(36) $[_{\text{Event}}$ CAUSE
 $([_{\text{Thing}} * \text{W}],$
 $[_{\text{Event}}$ GO_{Poss}
 $([_{\text{Thing}}$ Y KNIFE-WOUND :CONFLATED],
 $[_{\text{Path}}$ $\text{TOWARD}_{\text{Poss}}$
 $([_{\text{Position}}$ AT_{Poss} $([_{\text{Thing}}$ Y KNIFE-WOUND], $[_{\text{Thing}} * \text{Z}])])])])]$

Because Y does not have a $*$ specification, only W and Z need to be filled in. Once these positions are filled in, the resulting CLCS fulfills the full coverage requirement, since there is a path from the root to the leaves that covers all of the words of the source-language sentence. The resulting CLCS is

(37) $[_{\text{Event}}$ CAUSE
 $([_{\text{Thing}}$ I],
 $[_{\text{Event}}$ GO_{Poss}
 $([_{\text{Thing}}$ KNIFE-WOUND],
 $[_{\text{Path}}$ $\text{TOWARD}_{\text{Poss}}$
 $([_{\text{Position}}$ AT_{Poss} $([_{\text{Thing}}$ KNIFE-WOUND], $[_{\text{Thing}}$ JOHN])])])])]

To complete the translation, this CLCS must be decomposed into RLCSs that satisfy the full coverage requirement. The RLCS that is selected as a match for this top-level CLCS is that of the word *dar*:

25 In (27) we abbreviated the lexical entries for *stab* and *dar*, showing the $*$ marker only in the position that was relevant to example (26). In 36 and 38, we show the complete form of the lexical entries to illustrate the notion of full coverage.

(38) $[_{\text{Event}}$ CAUSE
$([_{\text{Thing}} * \text{W}],$
$[_{\text{Event}}$ GO$_{\text{Poss}}$
$([_{\text{Thing}} * \text{Y}],$
$[_{\text{Path}} *$ TOWARD$_{\text{Poss}}$
$([_{\text{Position}}$ AT$_{\text{Poss}}$ $([_{\text{Thing}}$ Y$], [_{\text{Thing}}$ Z$])])])])]$

Unlike the RLCS for *stab*, the Y position is associated with a * marker; thus, it is necessary to find RLCSs for all three positions W, Y, and Z. Positions W and Y are filled at the leaf level and Z is filled at the TOWARD$_{\text{Poss}}$ level. Once these positions are filled, the combination of the RLCSs for *yo, dar, puñaladas, a,* and *Juan* covers the entire concept. Thus, the full coverage requirement is satisfied.

It should be noted that a number of other systems have attempted to tackle divergences similar to those discussed in this paper without appealing to the notion of *full coverage*. Three examples of such systems are (1) GETA/ARIANE (Vauquois and Boitet 1985; Boitet 1987); (2) LMT (McCord 1989); and (3) METAL (Alonso 1990; Thurmair 1990). In particular, these approaches address the problem of thematic divergence by means of transfer rules of the following form, respectively:

(39) like(SUBJ(ARG2:GN),OBJ1(ARG1:GN))
plaire(SUBJ(ARG1:GN),OBJ1(ARG2:PREP,GN)) ⇔

(40) `gverb(like(dat:*,nom:X),ge+fall,*:X)`

(41) like V ⇒ gustar V
NP ([ROLE SUBJ]) ⇒ NP ([ROLE IOBJ])
NP ([ROLE DOBJ]) ⇒ NP ([ROLE SUBJ])

One problem with these approaches is that surface syntactic decisions are, in a sense, performed off-line by means of lexical entries and transfer rules that specifically encode language-specific syntactic information. Such a scheme is limited in that it has no potential for relating thematic divergence to the rest of the space of divergence possibilities. Moreover, although transfer rules might be deemed suitable for local divergences such as simple subject–object reversal, it is well known that simple transfer rules of this type do not readily accommodate more complicated divergences.

Consider a more complicated case such as the following promotional divergence:[26]

(42) **Promotional divergence:**
E: The baby just fell ⇒ F: Le bébé vient de tomber
'The baby just (verb-past) of fall'

Here, the English adverbial *just* is translated as the French main verb *venir*, which takes the falling event as its complement *de tomber*.

At first glance, it might seem difficult to construct transfer rules that handle such cases. However, the LFG-MT system by Kaplan et al. (1989) does, in fact, handle such cases by means of mappings between source and target *functional structures* (f-structures). The f-structures that correspond, respectively, to the English and French sentences in this example are the following:

26 This example was taken from Kaplan et al. (1989).

(43) (i)

$$\begin{bmatrix} \text{PRED} & \text{'JUST}\langle(\uparrow \text{ARG})\rangle\text{'} \\ \text{ARG} & \begin{bmatrix} \text{PRED} & \text{'FALL}\langle(\uparrow \text{SUBJ})\rangle\text{'} \\ \text{TENSE} & \text{PAST} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'BABY'} \\ \text{NUM} & \text{SG} \\ \text{SPEC} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{'THE'} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

(ii)

$$\begin{bmatrix} \text{PRED} & \text{'VENIR}\langle(\uparrow \text{SUBJ})(\uparrow \text{XCOMP})\rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'BÉBÉ'} \\ \text{GENDER} & \text{MASC} \\ \text{NUMB} & \text{SG} \\ \text{SPEC} & \begin{bmatrix} \text{DEF} & + \\ \text{PRED} & \text{'LE'} \end{bmatrix} \end{bmatrix} \\ \text{XCOMP} & \begin{bmatrix} \text{PRED} & \text{'TOMBER}\langle(\uparrow \text{SUBJ})\rangle\text{'} \\ \text{COMPL} & \text{DE} \\ \text{TENSE} & \text{INF} \\ \text{SUBJ} & \begin{bmatrix}\text{BÉBÉ}\end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The translation mapping is performed by a transfer equation that relates the source- and target-language f-structures:

(44)
$$\begin{array}{l} (\tau \uparrow \text{PRED 'JUST}\langle(\uparrow \text{ARG})\rangle\text{'}) = \text{VENIR} \\ (\tau \uparrow \text{XCOMP}) = \tau\ (\uparrow \text{ARG}) \end{array}$$

This equation identifies *venir* as the corresponding French predicate, and it maps the argument of *just* to a complement that is headed by the prepositional complementizer *de*.

Although such a case is handled in the LFG-MT system, there are a number of problems with this approach. A serious flaw concerns the handling of divergences in the context of embedded clauses. (For additional discussion, see Sadler and Thompson 1991.) In particular, if the English sentence in (42) were realized as an embedded complement such as *I think that the baby just fell,* it would not be possible to generate the French output. The reason for this is that the LFG-MT system is not designed to handle an interaction between a (divergent) matrix clause and a (nondivergent) embedded clause. This sentence is broken down into predicate–argument relations that conform (roughly) to the following logical specification:

(45) think(I,fall(baby))
just(fall(baby))

Because the logical constituent fall(baby) is viewed as an argument of two logical heads, "think" and "just," the LFG-MT generator cannot determine how to compose these concepts and produce an output string.

The divergence solution proposed in the LCS framework overcomes this difficulty by imposing the full coverage constraint. In particular, specific relations are set up between logical heads and their associated arguments and modifiers so that there could never be any question of how two concepts are composed, even for embedded cases. For the current example, the LCS approach would reduce the logical relations to a single specification for both French and English:

(46) think(I,fall(baby,just))

That is, the "just" component of meaning is a modifier of the "falling" action, regardless of how this constituent is realized on the surface.

The full benefit of this approach is further demonstrated when one considers the extent to which the full coverage constraint carries over to other divergence categories. Consider the following (frequently cited) conflational divergence:

(47) **Conflational divergence:**
E: John swam across the river ⇒
F: John a traversé la rivière à la nage
'John crossed the river by swimming'

In this example, the English path component (*across*) is translated as the French main verb (*traverser*), and the English main verb (*swim*) is translated as the French manner component (*à la nage*). In a system like TAUM (Isabelle 1987), transfer rules impose lexical transformations in order to establish correspondences between source and target structures, i.e., rules of the following form:

(48) (X swim (across Y)) ⇔ (X traverser Y (à la nage))

This rule maps the path component (*across*) to the French main verb (*traverser*), and the English main verb (*swim*) to the French adverbial (*à la nage*).[27]

The disadvantage to this approach is that it requires a new transfer rule for every adverbial that might potentially participate in the traverser construction (e.g., *à pied, en courant, en marchant*, etc.). The LCS approach, on the other hand, resolves this type of divergence compositionally by relying on the full coverage constraint and the :CONFLATED marker, analogous to the handling of example (26). The underlying LCS for the two sentences in (47) is the following:

(49) $[_{\text{Event}}\ \text{GO}_{\text{Loc}}$
$([_{\text{Thing}}\ \text{JOHN}],$
$[_{\text{Path}}\ \text{ACROSS}_{\text{Loc}}\ ([_{\text{Position}}\ \text{AT}_{\text{Loc}}\ ([_{\text{Thing}}\ \text{JOHN}],\ [_{\text{Location}}\ \text{RIVER}])])],$
$[_{\text{Manner}}\ \text{SWIMMINGLY}])]$

The solution to this example relies on the assumption that the distinction between the English and French exists by virtue of the fact that the word *swim* includes the manner component *swimmingly*, whereas the word *traverser* does not. Because of this conflational distinction, the manner component is suppressed in English, but is overtly realized (as *à la nage*) in French.[28] The important point is that this entire concept is fully

27 A related, but more general, strategy would be to handle such cases in bilingual lexical entries (see, for example, Beaven 1992a, 1992b; Whitelock 1992; Trujillo 1992).

28 The use of the word *traverser* (i.e., *cross*) instead of *nager* (i.e., *swim*) is independently determined by the fact that the path component ACROSS is present in the conceptual representation (i.e., there would be no way to realize the path component in conjunction with the word *nager*).

covered (in the sense of [34]) by both the source- and target-language sentences, even though these two sentences do not have the same structural representation. That is, as long as all conceptual components of (49) are somehow retrievable, the suppression/realization of the individual components in the surface structure may be forced by the presence/absence of the :CONFLATED marker in the relevant lexical entries.

5.4 Interacting Divergence Types

We now turn to another important issue that has only recently received the attention it deserves, namely, that of handling interacting divergence types. In particular, there have been criticisms (see, for example, Lindop and Tsujii 1991) of systems that perform transfer on relatively shallow analyses (e.g., early METAL [Alonso 1990, Thurmair 1990] and LTAG [Abeillé, Schabes, and Joshi 1990]) owing to the fact that such systems are not likely to be able to handle divergence interactions (although they may be able to handle each divergence type in isolation). The solution adopted in the current approach does not appeal to a shallow analysis (i.e., it does not use a set of already-coded canned "frames" with predetermined argument structure). Rather, the syntactic structures are derived *compositionally* on the basis of two pieces of information: the structure of the CLCS (i.e., the language-independent predicate-argument information) and the lexical entries (i.e., the RLCSs and their associated language-dependent information). It would not be possible to handle interacting divergence types in an approach that maps directly from a set of hard-wired source-language frames to a set of hard-wired target-language frames. This is because an argument that occurs in a divergent phrasal construction might itself be a divergent phrasal construction.

Consider the following example:

(50) **Promotional and thematic divergence:**
S: Leer libros le suele gustar a Juan
'Reading books (him) tends to please (to) John'
E: John usually likes reading books

This example exhibits a simultaneous occurrence of two types of divergences: the verb *soler* exhibits a promotional divergence with respect to its internal argument *gustar a Juan*, which itself exhibits a thematic divergence. The recursive nature of the $\mathcal{GLR}$ is crucial for handling such cases.

The CLCS for (50) is the following:

(51) $[_{\text{State}}$ BE_{Circ}
 $([_{\text{Thing}}$ JOHN],
 $[_{\text{Position}}$ AT_{Circ}
 $([_{\text{Thing}}$ JOHN], $[_{\text{Event}}$ READ $([_{\text{Thing}}$ JOHN], $[_{\text{Thing}}$ BOOK])])],
 $[_{\text{Manner}}$ LIKINGLY],
 $[_{\text{Manner}}$ HABITUALLY])]

Note that there are two modifiers, LIKINGLY and HABITUALLY. It is the job of the $\mathcal{GLR}$ function to determine the appropriate decomposition of the event on the basis of the language-specific requirements of the RLCSs involved in the mapping. In the current example, the LIKINGLY component is, in a sense, an "inherent" modifier, since it appears in the RLCS of both *like* and *gustar*. In contrast, the HABITUALLY modifier is an independent constituent that corresponds to independent RLCSs for *usually* and *soler*.

We will now formally analyze how this example is handled. The two relevant override mappings are specified in (15) and (11), repeated here for convenience:

(52) **Promotional override:**
1.′ $X' \Leftrightarrow Z$
4.′ $Q' \Leftrightarrow X$

(53) **Thematic override:**
1.′ $W' \Leftrightarrow Z$
4.′ $Z' \Leftrightarrow W$

In terms of the logical constituents that participate in the divergence mapping, X′ corresponds to [$_{State}$ BE$_{Circ}$ ...], W′ corresponds to [$_{Thing}$ JOHN], Z′ corresponds to [$_{Event}$ READ ...], and Q′ corresponds to [$_{Manner}$ HABITUALLY].

Formally, the structure of the English sentence in (50) has the default syntactic representation

(54) $[_{Y-MAX}\ [_{Y-MAX}\ W\ [_{X-MAX}\ X\ Z]]\ Q]$

In contrast, the equivalent Spanish sentence has an entirely different syntactic representation:

(55) $[_{Y-MAX}\ [_{Y-MAX}\ Z\ [_{X-MAX}\ Q\ [\ \ldots\ X\ W]]]]$

Note that this structure differs from the default representation in that the external argument is Z, not W, and that the syntactic head is Q, not X. Also, because the promotional mapping forces X into an internal argument position, additional structure is created so that X retains its status as a head. This head selects an internal argument that, in this case, is W rather than Z because of the interaction with the thematic divergence.

Suppose we were to generate the Spanish sentence for this example. The RLCSs for the Spanish case conform to the following formal specifications:

(56) *gustar*: $[_{T(X')}\ X'\ ([_{T(W')}\ \text{:INT}\ W'],\ [_{T(Z')}\ \text{:EXT}\ Z'])]$
soler: $[_{T(Q')}\ \text{:PROMOTE}\ Q']$

When the $\mathcal{GLR}$ is applied to the CLCS of (51), the promotional override (52) is immediately triggered by the :PROMOTE marker in the RLCS for *soler*; this invocation crucially precedes the invocation of the thematic override (53).[29] The promotional override forces Q′ (i.e., [$_{Manner}$ HABITUALLY]) to be realized as the syntactic head *soler*. This head takes an internal argument corresponding to X′ ([$_{State}$ BE$_{Circ}$ ...]) that is

29 This seems to indicate that there is some notion of prioritization during the resolution of interacting divergence types. In particular, the head swapping (promotional and demotional) divergences appear to take priority over the "argument swapping" cases (thematic). The decision to impose this prioritization is not entirely unprincipled. It would not be possible to apply these two overrides in the opposite order, since a head *must* be properly positioned (potentially via a promotional override) in order to identify the relative positions for its satellites (potentially via a thematic override). Although the formal ramifications of this ordering have not yet been established, it should be noted that the prioritization fits in naturally in the current framework, given that the syntactic realization process starts by realizing "outer" phrases, but then recursively realizes "inner" phrases before any attachments are made.

realized as the verb *gustar*. Recall that the structural relation between the subordinated verb and its internal argument is not changed by the promotional mapping (see Footnote 13). However, the promotional operation does change the relation between the subordinated verb and its external argument by realizing the argument in an external position relative to the main verb. Normally, this would mean that the CLCS constituent [$_{\text{Thing}}$ JOHN] would become an external argument of *soler* such as in the noninteracting case, *John suele leer libros*. In the current example, however, the attachment of the external argument is delayed until the recursive application of the $\mathcal{GLR}$ on [$_{\text{State}}$ BE$_{\text{Circ}}$ ...].

At this point, the :INT and :EXT markers trigger the thematic interchange, and the logical subject [$_{\text{Thing}}$ JOHN] is realized as the internal argument *a Juan*. The [$_{\text{Event}}$ READ ...] constituent is then taken to be the external argument of *gustar*, except that this constituent cannot be attached inside of the subordinate phrase owing to the promotional divergence. Instead, the current phrase is completed and the external argument is "passed up" to the higher phrase, which then attaches it in an external position relative to the verb *soler*. The final structure is then generated:

(57) [$_{\text{C-MAX}}$
  [$_{\text{I-MAX}}$
    [$_{\text{C-MAX}}$ leer libros]
    [$_{\text{V-MAX}}$ [$_{\text{V}}$ le suele]
      [$_{\text{V-MAX}}$ [$_{\text{V}}$ gustar] [$_{\text{P-MAX}}$ a Juan]]]]]

Thus, we have shown how the current framework provides a formal means for demonstrating how interacting divergence types are handled.

6. Limitations and Conclusions

The current approach has been implemented in a system called UNITRAN (Dorr 1990a, 1990b, 1993b).[30] Many of the problems associated with the direct replacement and transfer approaches of previous systems have been eliminated in this new design. In particular, UNITRAN does not make use of analysis/synthesis rules that are meticulously tailored to each of the source and target languages, nor does it require detailed source-to-target transfer rules. On the other hand, because the system is designed (deliberately) to operate on one sentence at a time, it has a number of inherent limitations. In particular, the lack of a theory of multisententiality makes high quality translation difficult, since it is often the case that a single sentence in one language should be translated as two or more sentences in the other language; currently, UNITRAN does not allow for a mismatch between the number of source- and target-language sentences.[31] Since generation has not been the primary focus of the current research, this and other generation problems have not yet been addressed in this framework. Such problems include cohesion (Granville 1983), selecting propositional and rhetorical goals (McKeown 1985), selecting open-class items from "deep knowledge" (Goldman

30 The name UNITRAN stands for UNIversal TRANslator, that is, the system serves as the basis for translation across a variety of languages, not just two languages or a family of languages.

31 In general, systems should be designed so that sentences are realized differently in different languages, depending on the speaker's intended effect. In the words of a reviewer, "What's a long-winded and boring sentence to an American may be precisely a very fine formal sentence to a German." Even for languages as close as French and English, sentence boundaries differ about 10% of the time (see Brown et al. 1991).

1975; Jacobs 1985; Kittredge, Iordanskaja, and Polguère 1988; Nirenburg and Nirenburg 1988; Nirenburg et al. 1992; among others), ordering propositions for producing coherent text (Hovy 1988), resolving anaphora (Derr and McKeown 1984; Sondheimer, Cumming, and Albano 1990; Werner and Nirenburg 1988), and many others. In fact, not all of these issues (e.g., selecting propositional and rhetorical goals) are directly relevant to the task of machine translation, which already has the advantage (at least from the generation point of view) that the source-language sentence and, in the current model, the conceptual analysis underlying this sentence are available at the onset of the generation process.

Instead, the current research focuses on demonstrating the utility of the LCS-based interlingua and the associated parameters for resolving translation divergences while maintaining the systematic relation between the interlingua and the syntax. The tasks involved in achieving this objective have been reduced to what might be considered the standard "what" and "how" questions of generation: (1) *lexical selection*, i.e., the task of deciding what target-language words accurately reflect the meaning of the corresponding source-language words; and (2) *syntactic realization*, i.e., the task of determining how target-language words are mapped to their appropriate syntactic structures. In the context of the current model, the first task consists of matching the LCS-based interlingua (the CLCS) against the LCS-based entries (the RLCS) in the dictionary in order to select the appropriate word, and the second task consists of realizing the positions marked by $*$ (and other parametric markers) into the appropriate syntactic structure.

The question of lexical choice is one that deserves further discussion. The details of the matching process that achieves lexical selection of the target-language RLCSs (e.g., the selection of the RLCSs for *like* [or *gustar*] from the underlying CLCS shown in Definition 1) have not been presented here, but see Dorr (1993a) for a discussion with examples. Roughly, lexical selection in UNITRAN is a "reverse" unification-like process that matches the CLCS to the RLCS templates in the lexicon and chooses the associated lexical words accordingly. One of the important problems that must be considered with respect to lexical selection is that of overgeneration.[32] In particular, one might ask how the matcher knows whether it should try to choose phrases that restate the source language more succinctly in the target language, or whether it should be allowed to restate the source-language phrases more verbosely in the target language. This comes up in cases such as the *stab* example given earlier. It turns out that the word *stab* can be translated into Spanish as the succinct form *apuñalar*, or as the more verbose form *dar puñaladas*. Similarly, in the reverse direction, the translation of *dar puñaladas* is the more succinct form *stab*; however, one could conceive of a more verbose translation (e.g., *inflict knife wounds* or even *give knife wounds*). Currently, there is no preference assignment during lexical selection (as in Nirenburg and Nirenburg 1988; Wilks 1973); instead, the system requires an exact match of the CLCS to the target-language RLCS (or some combination of RLCSs). If there is more than one way of matching the CLCS, multiple forms will be generated (although we have discussed only one target-language form, *dar puñaladas*, for the *stab* example). A first-pass approach to resolving such cases of overgeneration (based on aspectual features) is discussed in Dorr (1992a) and in more detail in Dorr (1993b). In addition, a model

32 The complexity of the lexical selection process is a well-studied problem. See, for example, the work by Reiter (1990), which shows that the selection of the optimal set of adjectives for a noun phrase under some very strong conditions is NP-complete. Presumably, the general lexical selection problem is considerably harder.

of generation based on theories of tense by Allen (1983, 1984), Hornstein (1990), and Reichenbach (1947) is discussed in Dorr and Gaasterland (1992) and Dorr (1993b).

The difficulty of the lexical choice problem will become more important as the system grows beyond its current prototypical state. Research is currently underway to extend the system by developing automatic lexical acquisition procedures that make use of a small set of conceptual structures, as a starting point, and then acquire syntactic and semantic information on the basis of these initial representations plus machine-readable definitions from the *Longman's Dictionary of Contemporary English* (Proctor 1978). (LDOCE is useful because it includes collocations and sense frequency, thus making it possible to determine the argument structures for different words.) This investigation will benefit from the work of several researchers in the field of automatic lexicon construction, most notably, Brent (1993), Boguraev and Briscoe (1989), Boguraev and Pustejovsky (1990), Briscoe and Copestake (1990), Byrd et al. (1987), Farwell, Guthrie, and Wilks (1992), Montemagni and Vanderwende (1992), Pustejovsky (1987), Pustejovsky and Bergler (1987), and Pustejovsky, Bergier, and Anick (1993), among others. In particular, it has been argued convincingly by Farwell, Guthrie, and Wilks (1992) that resources such as the LDOCE are useful for constructing dictionary representations for languages other than English, thus paving the way for scaling up interlingual machine translations so that they have broader coverage. Once this extension is complete, we intend to scale up the UNITRAN system and test the LCS approach to lexical choice on a broader set of phenomena using a larger lexicon.

The current framework provides a systematic classification of machine translation divergences. We have shown how this classification can be formally defined and systematically resolved through the use of general mapping relations and a small set of cross-linguistic parameters. Because the parameters are used to *factor out* "transfer" information, the current approach obviates the need for transfer rules. We have provided evidence that supports the view that the lexical–semantic divergence classification proposed in the current framework covers all lexical–semantic divergences that arise during translation (i.e., divergences based on properties associated with lexical entries that are not based on purely syntactic information, idiomatic usage, aspectual knowledge, discourse knowledge, domain knowledge, or world knowledge). Since the characterization of the range of potential divergences is manageably small, the task of accommodating divergences is immensely simplified. We have also demonstrated the usefulness of a full coverage requirement as a tool that allows one to judge whether a particular target-language sentence fully covers the concept that underlies the corresponding source-language sentence. Finally, we have shown, formally, that the current model accommodates interacting divergence types.

Acknowledgments
This paper describes research done at the University of Maryland Institute for Advanced Computer Studies. Support for this research has been provided in part by the National Science Foundation under a Young Investigator Award IRI-9357731 and grant IRI-9120788, by the DARPA Basic Research Program under grant N00014-92-J-1929, by the Army Research Office under contract DAAL03-91-C-0034 through Batelle Corporation, and by the Army Research Institute under contract MDA-903-92-R-0035 through Microelectronics and Design, Inc. Useful guidance and commentary during the research and preparation of this document were provided by Bob Berwick, Bruce Dawson, Ken Hale, Clare Voss, and Amy Weinberg. The author would also like to thank three anonymous reviewers for their helpful comments during the preparation of this document.

References

Abeillé, Anne; Schabes, Yves; and Joshi, Aravind K. (1990). "Using lexicalized tags for machine translation." In *Proceedings, 13th International Conference on*

Computational Linguistics, Helsinki, Finland, 1–6.

Abney, S. (1989). "A computational model of human parsing." *Journal of Psycholinguistic Research* 18:129–144.

Allen, James F. (1983). "Maintaining knowledge about temporal intervals." *Communications of the ACM* 26(11):832–843.

Allen, James F. (1984). "Towards a general theory of action and time." *Artificial Intelligence* 23(2):123–160.

Alonso, Juan Alberto. (1990). "Transfer interstructure: designing an 'interlingua' for transfer-based MT systems." In *Proceedings, Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Linguistics Research Center, The University of Texas, Austin, Texas, 189–201.

Arnold, Doug, and des Tombe, Louis. (1987). "Basic theory and methodology in Eurotra." In *Machine Translation: Theoretical and Methodological Issues*, edited by Sergei Nirenburg, 114–135. Cambridge: Cambridge University Press.

Barnett, Jim; Mani, Inderjeet; Martin, Paul; and Rich, Elaine (1991a). "Reversible machine translation: What to do when the languages don't line up." In *Proceedings, Workshop on Reversible Grammars in Natural Language Processing, ACL-91*, University of California, Berkeley, California, 61–70.

Barnett, Jim; Mani, Inderjeet; Rich, Elaine; Aone, Chinatsu; Knight, Kevin; and Martinez, Juan C. (1991b). "Capturing language-specific semantic distinctions in interlingua-based MT." In *Proceedings, Machine Translation Summit*, Washington, DC, 25–32.

Beaven, John (1992a). "Lexicalist unification-based machine translation." Doctoral dissertation, University of Edinburgh, Edinburgh, UK.

Beaven, John (1992b). "Shake and bake machine translation." In *Proceedings, 14th International Conference on Computational Linguistics*, Nantes, France, 603–609.

Boguraev, Branimir, and Briscoe, Ted (1989). *Computational Lexicography for Natural Language Processing*. London: Longman.

Boguraev, Branimir, and Pustejovsky, James (1990). "Lexical ambiguity and the role of knowledge representation in lexical design." In *Proceedings, 13th International Conference on Computational Linguistics*, Helsinki, Finland, 36–41.

Boitet, Christian (1987). "Research and development on MT and related techniques at Grenoble University (GETA)." In *Machine Translation: The State of the Art*, edited by Margaret King, 133–153. Edinburgh University Press, Edinburgh.

Brent, Michael (1993). "From grammar to lexicon: Unsupervised learning of lexical syntax." *Computational Linguistics* 19(2):243–262.

Briscoe, E. J., and Copestake, A. A. (1990). "Enjoy the paper: Lexical semantics via lexicology." In *Proceedings, 13th International Conference on Computational Linguistics*, Helsinki, Finland, 42–47.

Byrd, Roy J., Calzolari, Nicoletta; Chodorow, Martin S.; Klavans, Judith L.; Neff, Mary S.; and Rizk, Omneya A. (1987). "Tools and methods for computational linguistics." *Computational Linguistics* 13(3–4):219–240.

Carbonell, Jaime G., and Tomita, Masaru (1987). "Knowledge-based machine translation, the CMU approach." In *Machine Translation: Theoretical and Methodological Issues*, edited by Sergei Nirenburg, 68–89. Cambridge: Cambridge University Press.

Chomsky, Noam A. (1981). *Lectures on Government and Binding*. Dordrecht, Holland: Foris Publications.

Chomsky, Noam A. (1982). *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge: MIT Press.

Chomsky, Noam A. (1986a). *Barriers*. Cambridge: MIT Press.

Chomsky, Noam A. (1986b). *Knowledge of Language: Its Nature, Origin and Use*. Cambridge: MIT Press.

Copeland, C.; Durand, J.; Krauwer, S.; and Maegaard, B. (1991). "The Eurotra linguistic specifications." In *Studies in Machine Translation and Natural Language Processing, Volume 1*, edited by Erwin Valentini. Brussels: Commission of the European Communities.

Derr, M., and McKeown, K. (1984). "Using focus to generate complex and simple sentences." In *Proceedings, Tenth International Conference on Computational Linguistics*, Stanford, California, 319–326.

Dorr, Bonnie J. (1990a). "A cross-linguistic approach to machine translation." In *Proceedings, Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Linguistics Research Center, The University of Texas, Austin, Texas, 13–32.

Dorr, Bonnie J. (1990b). "Solving thematic divergences in machine translation." In *Proceedings, 28th Annual Conference of the Association for Computational Linguistics*, University of Pittsburgh, Pittsburgh, Pennsylvania, 127–134.

Dorr, Bonnie J. (1992a). "Lexical semantics for interlingual machine translation." *Machine Translation* 7(3).

Dorr, Bonnie J. (1992b). "A parameterized approach to integrating aspect with lexical-semantics for machine translation." In *Proceedings, 30th Annual Conference of the Association of Computational Linguistics*, University of Delaware, Newark Delaware, 257–264.

Dorr, Bonnie J. (1993a). "Interlingual machine translation: A parameterized approach." *Artificial Intelligence* 63(1,2).

Dorr, Bonnie J. (1993b). *Machine Translation: A View from the Lexicon*. Cambridge: MIT Press.

Dorr, Bonnie J., and Gaasterland, Terry (1992). "Reflecting time in generated text: Tense, aspect and temporal connecting words." Technical report UMIACS TR 92-92, CS TR 2950, Department of Computer Science, University of Maryland, College Park, Maryland.

Dorr, Bonnie J., and Voss, Clare R. (1993a). "Constraints on the space of MT divergences." In *Building Lexicons for Machine Translation*, Papers from the 1993 Spring Symposium, Technical report SS-93-02, Stanford University, Stanford, California.

Dorr, Bonnie J., and Voss, Clare R. (1993b). Machine translation of spatial expressions: Defining the relation between an interlingua and a knowledge representation system." In *Proceedings, Twelfth Conference of the American Association for Artificial Intelligence*, Washington, DC.

Farwell, David; Guthrie, Louise; and Wilks, Yorick (1992). "The automatic creation of lexical entries for a multilingual MT system." In *Proceedings, 14th International Conference on Computational Linguistics*, Nantes, France, 532–538.

Goldman, Neil M. (1975). "Conceptual memory and inference." In *Conceptual Information Processing*, edited by Roger C. Schank, 289–371. Amsterdam, Holland: Elsevier Science Publishers.

Granville, Robert (1983). "Cohesion in computer text generation: Lexical substitution." Technical Report, LCS Technical Report 310, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hale, Kenneth, and Keyser, S. Jay (1986a). "Some transitivity alternations in English." Technical report, Lexicon Project Working Paper 7, Center for Cognitive Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hale, Kenneth, and Keyser, S. Jay (1986b). "A view from the middle." Technical report, Lexicon Project Working Paper 10, Center for Cognitive Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hale, Kenneth, and Keyser, S. Jay (1989). "On some syntactic rules in the lexicon." Technical Report, Center for Cognitive Science, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hale, Kenneth, and Laughren, Mary (1983). "Warlpiri lexicon project: Warlpiri dictionary entries." Technical report, Warlpiri Lexicon Project, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Hornstein, Norbert (1990). *As Time Goes By*. Cambridge: MIT Press.

Hovy, Eduard (1988). *Generating Natural Language Under Pragmatic Constraints*. New Jersey: Lawrence Erlbaum Associates.

Isabelle, Pierre (1987). "Machine translation at the TAUM group." In *Machine Translation: The State of the Art*, edited by Margaret King, 247–277. Edinburgh: Edinburgh University Press.

Jackendoff, Ray S. (1983). *Semantics and Cognition*. Cambridge: MIT Press.

Jackendoff, Ray S. (1990). *Semantic Structures*. Cambridge: MIT Press.

Jacobs, Paul S. (1985). "PHRED: A generator for natural language interfaces." *Computational Linguistics* 11(4):219–242.

Johnson, Rod; King, Maghi; and des Tombe, Louis (1985). "Eurotra: A multilingual system under development." *Computational Linguistics* 11(2,3):155–169.

Kameyama, Megumi; Ochitani, Ryo; Peters, Stanley; and Sirai, Hidetoshi (1991). "Resolving translation mismatches with information flow." In *Proceedings, 29th Annual Meeting of the Association for Computational Linguistics*, University of California, Berkeley, California, 193–200.

Kaplan, Ronald M., and Bresnan, Joan (1982). "Lexical-functional grammar: A formal system for grammatical representation." In *The Mental Representation of Grammatical Relations*, edited by Joan Bresnan, 173–281. Cambridge: MIT Press.

Kaplan, Ronald M.; Netter, Klaus; Wedekind, Jurgen; and Zaenen, Annie (1989). "Translation by structural correspondences." In *Proceedings, Fourth Conference of the European Chapter of the Association for Computational Linguistics*, Manchester, UK, 272–281.

Kay, Martin (1984). "Functional unification grammar: A formalism for machine translation." In *Proceedings, 10th International Conference on Computational Linguistics*, Stanford University, Stanford, California, 75–78.

Kinoshita, Satoshi; Phillips, John; and Tsujii, Jun-ichi (1992). Interaction between structural changes in machine translation." In *Proceedings, 14th International Conference on Computational Linguistics*, Nantes, France, 679–685.

Kittredge, Richard I.; Iordanskaja, Lidija; and Polguère, Alain (1988). "Multi-lingual text generation and the meaning-text theory." In *Proceedings, Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Levin, Beth, and Rappaport, Malka (1986). "The formation of adjectival passives." *Linguistic Inquiry* 17:623–662.

Lindop, Jeremy, and Tsujii, Jun-ichi (1991). "Complex transfer in MT: A survey of examples." Technical report, CCL/UMIST Report 91/5, Center for Computational Linguistics, UMIST, Manchester, UK.

Marantz, Alec (1992). "The way-constructions and the semantics of direct arguments in English: A reply to Jackendoff." Technical report, Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, Massachusetts.

McCord, Michael C. (1989). "Design of LMT: A prolog-based machine translation system." *Computational Linguistics* 15(1):33–52.

McKeown, Kathleen (1985). *Text generation: Using discourse strategies and focus constraints to generate natural language text.* Cambridge: Cambridge University Press.

Melby, A. K. (1986). "Lexical transfer: Missing element in linguistic theories." In *Proceedings, 11th International Conference on Computational Linguistics*, Bonn, Germany.

Meyer, Ingrid; Onyshkevych, Boyan; and Carlson, Lynn (1990). "Lexicographic principles and design for knowledge-based machine translation." Technical report, CMU CMT Technical Report 90-118, Carnegie Mellon University, Pittsburgh, Pennsylvania.

Montemagni, Simonetta; and Vanderwende, Lucy (1992). "Structural patterns vs. string patterns for extracting semantic information from dictionaries." In *Proceedings, 14th International Conference on Computational Linguistics*, Nantes, France, 546–552.

Nirenburg, Sergei; Carbonell, Jaime; Tomita, Masaru; and Goodman, Kenneth (1992). *Machine Translation: A Knowledge-Based Approach.* San Mateo, California: Morgan Kaufmann.

Nirenburg, Sergei, and Goodman, Kenneth (1990). "Treatment of meaning in MT systems." In *Proceedings, Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Linguistics Research Center, The University of Texas, Austin, Texas, 171–187.

Nirenburg, Sergei, and Levin, Lori (1989). "Knowledge representation support." *Machine Translation* 4(1):25–52.

Nirenburg, Sergei, and Nirenburg, Irene (1988). "A framework for lexical selection in natural language generation." In *Proceedings, 12th International Conference on Computational Linguistics*, Budapest, Hungary, 471–475.

Nirenburg, Sergei; Raskin, Victor; and Tucker, Allen B. (1987). "The structure of interlingua in translator." In *Machine Translation: Theoretical and Methodological Issues*, edited by Sergei Nirenburg, 90–113. Cambridge: Cambridge University Press.

Perlmutter, David M. (1983). *Studies in Relational Grammar 1.* Chicago: The University of Chicago Press.

Proctor, P. (1978). *Longman Dictionary of Contemporary English.* London: Longman.

Pustejovsky, James (1987). "On the acquisition of lexical entries: The perceptual origin of thematic relations." In *Proceedings, 25th Annual Conference of the Association for Computational Linguistics*, Stanford University, Stanford, California, 172–178.

Pustejovsky, James, and Bergler, Sabine (1987). "The acquisition of conceptual structure for the lexicon." In *Proceedings, Sixth Conference of the American Association of Artificial Intelligence*, Seattle, Washington, 566–570.

Pustejovsky, James; Bergler, Sabine; and Anick, Peter (1993). "Lexical semantic techniques for corpus analysis." *Computational Linguistics* 19(2):331–358.

Reichenbach, H. (1947). *Elements of Symbolic Logic.* London: Macmillan.

Reiter, Ehud (1990). "Generating appropriate natural language object descriptions." Doctoral dissertation, Harvard University, Cambridge, MA.

Sadler, Louisa, and Thompson, Henry S. (1991). "Structural non-correspondence in translation." In *Proceedings, Fifth Conference of the European Chapter of the Association for Computational Linguistics*, Berlin, Germany, 293–298.

Shieber, Stuart M.; van Noord, Gertjan;

Moore, Robert C.; and Pereira, Fernando C. N. (1989). "A semantic-head-driven generation algorithm for unification-based formalisms." In *Proceedings, 27th Annual Conference of the Association for Computational Linguistics*, University of British Columbia, Vancouver, British Columbia, Canada, 7–17.

Shieber, Stuart M.; van Noord, Gertjan; Moore, Robert C.; and Pereira, Fernando C. N. (1990). "Semantic-head-driven generation." *Computational Linguistics* 16(1).

Sondheimer, N.; Cumming, S.; and Albano, R. (1990). "How to realize a concept: Lexical selection and the conceptual network in text generation." *Machine Translation* 5(1):57–78.

Thurmair, Gregor (1990). "Complex lexical transfer in metal." In *Proceedings, Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Linguistics Research Center, The University of Texas, Austin, Texas, 91–107.

Trujillo, Arturo (1992). "Locations in the machine translation of prepositional phrases." In *Proceedings, Fourth International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages*, Montreal, Canada, 13–20.

Tsujii, Jun-ichi, and Fujita, Kimikazu (1991). "Lexical transfer based on bilingual signs: Towards interaction during transfer." In *Proceedings, European Chapter of the Association for Computational Linguistics*, Berlin, Germany, 275–280.

Vauquois, Bernard, and Boitet, Christian (1985). "Automated translation at Grenoble University." *Computational Linguistics* 11(1):28–36.

Werner, P., and Nirenburg, S. (1988). "A specification language that supports the realization of intersentential anaphora." In *Proceedings, AAAI Workshop on Natural Language Generation*, Saint Paul, Minnesota.

Whitelock, Pete (1992). "Shake-and-bake translation." In *Proceedings, 14th International Conference on Computational Linguistics*, Nantes, France, 784–791.

Wilks, Yorick (1973). "An artificial intelligence approach to machine translation." In *Computer Models of Thought and Language*, edited by Roger C. Schank and K. M. Colby, 114–151. San Francisco: Freeman.

Zubizarreta, Maria Luisa (1982). "On the relationship of the lexicon to syntax." Doctoral dissertation, Department of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Zubizarreta, Maria Luisa (1987). *Levels of Representation in the Lexicon and in the Syntax*. Dordrecht, Holland: Foris Publications.