

Lexical and Idiomatic Transfer in Machine Translation: An LFG Approach

ONE-SOON HER, DAN HIGINBOTHAM, and
JOSEPH PENTHEROUDAKIS

1. Introduction

This paper first describes the transfer approach of machine translation (MT) in the ECS system, which is based on the linguistic theory of LFG (lexical-functional grammar) (Kaplan and Bresnan 1982, Kaplan and Maxwell 1988), and implements a unification formalism. Linguistic information is stored in rules and a lexicon composed of three types of entries: source, transfer, and target. All source lexical entries have corresponding transfer, or bilingual, entries, through which lexical replacement and necessary structure transformation are specified and performed. We will then describe the treatment of idioms within this LFG system. For purposes of translation, we identify three types of idioms and treat each type differently according to its syntactic and semantic nature.

2. The Transfer Approach in the ECS System

The ECS system employs an indirect-translation transfer approach, where analysis and generation are independently motivated and therefore a transfer component is needed for each language direction, e.g. English-to-Chinese. Within this approach, the transfer component takes the analysis of the source sentence and transforms it into that of the desired target sentence (e.g. Slocum 1985, Tucker 1987).

LFG, a theory of grammar that has been developed to serve as a computationally precise model of natural language (Bresnan 1982, Sells 1985), assigns two levels of syntactic representation to a sentence: a constituent structure (c-structure), represented as a tree, and a functional structure (f-structure). While the c-structure reflects the phrasal hierarchy and linear ordering in a sentence, the f-structure is an abstraction of the grammatical and functional information away from both phrasal constituency and ordering. Fig. 1 illustrates with a simple example the LFG co-description of c-structure, represented as a tree, and f-structure, shown

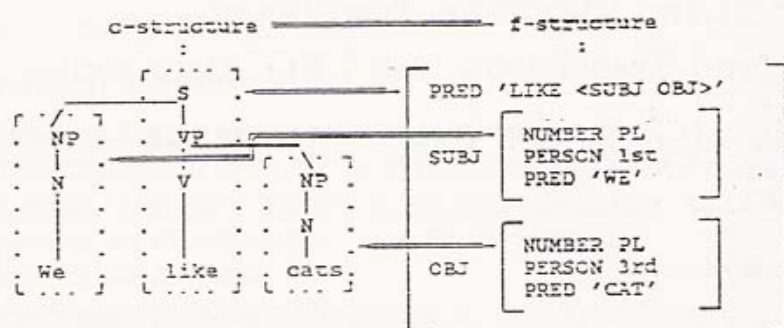


Fig. 1. LFG Co-description of c- and f-structures

in a bracketed dag (directed acyclic graph) representation of feature-value pairs.

One of the advantages of using LFG to implement a machine translation system is that its f-structure, which, unlike constituent structure, is order-free and relatively language-independent, provides a very suitable medium for transfer among different languages; thus, in recent years LFG has been gaining more popularity than other contemporary linguistic theories in the implementation of MT systems (e.g. Kudo and Nomura 1986, Her *et al.* 1989, Kaplan *et al.* 1989, Levin 1990, Sadler *et al.* 1990, Pentheroudakis 1990a). In accordance with LFG, the ECS system is stratificational in its organization. For a source sentence, analysis rules construct a c-structure which is mapped to an f-structure via functional equations annotated to the lexical entries and phrase structure rules. During transfer, the source words in the source f-structure are replaced with the appropriate target words; also, necessary structure changes are performed by transfer rules, which are instantiated through lexical transfer entries. The output of the transfer operations is the so-called 'transfer dag', which serves as the basis for generation. Generation proceeds in three stages: first, target word look-up; second, inflection of morphological elements and insertion of function words, if any; and finally, the linearization of the lexical items and grammatical functions in the target f-structure. Fig. II depicts the schema of the translation process in the ECS system.

To date, four operational language pairs have been developed at ECS: English to Chinese/Japanese/Korean, and Korean to English. Prototypes include English to Arabic/French and Arabic/Chinese/French/Japanese to English. Also, bi-directional translation between English and Norwegian has been under development at the MANTRA project of the University of Bergen in Norway. All examples in this paper will be given in the context of English to Chinese translation. Since the transfer approach

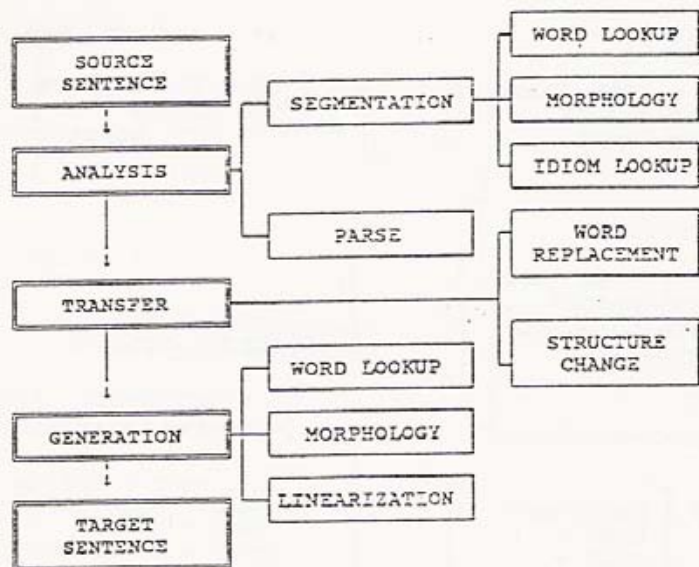


Fig. II. Translation Process in the ECS System

allows independently motivated analysis and generation, the same English grammar and lexicon have been used in the analysis portion in all language pairs mentioned above that have English as the source. Higinbotham (1990b) and Pentheroudakis (1990b) address the design philosophy behind the ECS MT system. See Her *et al.* (1989) and Pentheroudakis (1990a) for a more detailed account of the transfer design of the system. Her (1990) provides a general description of the LFG-based linguistic formalism employed in the ECS system and its application in the analysis of Chinese verbs.

3. Lexical Transfer

A lexical entry consists of feature-value pairs, representing grammatical specifications such as functional structure, subcategorization, syntactic category, lexically driven rules, etc. A lexical entry, being a structure associated with terminal nodes, is formally a dag, as are all structures associated with individual nodes in the tree. Since certain grammatical specifications are shared by an entire word-class, to capture linguistic generalizations as well as to preserve space and efficiency, a lexical entry may mention explicitly only feature-value pairs that are idiosyncratic, or unique, to that lexical item, as shown in 1a-1e. The shared or predictable specifications can be expressed as standard inheritance entries. Feature-

- 1a. `eW_book ::` \eW = English word
- | | |
|--|--|
| $\left[\begin{array}{l} \text{FS [FORM 'book']} \\ \backslash \text{STD-N-BOOK} \\ \backslash \text{STD-CAT-N} \end{array} \right]$ | $\begin{array}{l} \backslash \text{FS} = \text{f-structure} \\ \backslash = \text{inheritance} \\ \text{operator} \end{array}$ |
|--|--|
-
- 1b. `eW_car ::`
- | | |
|---|---|
| $\left[\begin{array}{l} \text{FS [FORM 'car']} \\ \backslash \text{STD-N-BOOK} \\ \backslash \text{STD-USE-VHCL} \\ \backslash \text{STD-CAT-N} \end{array} \right]$ | $\begin{array}{l} \backslash \text{VHCL} = \text{vehicle} \\ \backslash \text{CAT} = \text{category} \end{array}$ |
|---|---|
-
- 1c. `eX_STD-N-BOOK ::` \eX = Eng inheritance
- | | |
|---|---|
| $\left[\begin{array}{l} \text{FS} \left[\begin{array}{l} \text{ANIM -} \\ \text{HUMN -} \\ \text{CNRT +} \end{array} \right] \end{array} \right]$ | $\begin{array}{l} \backslash \text{ANIM} = \text{animate} \\ \backslash \text{HUMN} = \text{human} \\ \backslash \text{CNRT} = \text{concrete} \end{array}$ |
|---|---|
-
- 1d. `eX_STD-USE-VHCL ::`

$$\left[\text{FS [USE VHCL]} \right]$$
-
- 1e. `eX_STD-CAT-N ::`
- | | |
|---|--|
| $\left[\begin{array}{l} \text{CAT N} \\ \text{FS ANY} \end{array} \right]$ | |
|---|--|

value pairs contained in inheritance entries will be unified by extension into the lexical entry at the time of word lookup during analysis and generation or word replacement during transfer.

The monolingual lexical entries and respective language-specific analysis and generation rules are independently defined for the source language and the target language. The transfer component of a specific direction of a language pair, e.g. English to Chinese, consists of bilingual lexical entries and transfer rules. Since all source lexical entries have a counterpart transfer lexical entry, during transfer source words in the source f-structure are replaced with target words specified in the bilingual transfer entries, as shown in 2a–2c. Note that for typographical ease we will use the Chinese *pinyin* romanization in this paper, although in the actual system Chinese characters are fully supported.

As we shall see later in the discussion of idioms, a transfer entry may specify the proper translation of a source lexical item as a WORD or as

2a. ec_book ::

[	WORD	[	FORM 'shu1'	]
			TECH GENERAL	
			\STD-T-TRAN	

'ec = English-Chinese

'TECH = technical domain

2b. ec_car ::

[	WORD	[	FORM 'qi4che1'	]
			TECH GENERAL	
			\STD-T-TRAN	

2c. eX_STD-T-TRAN ::

[ACTION [TRAN (<↑ FORM> , <↓ FORM>)]]

the head of an IDIOM. While ↑ refers to the current level of the source word in the source f-structure, ↓ points to the dag under WORD (or IDIOM) in the transfer entry. The internally defined action TRAN requires two arguments and replaces the source form with the specified target form in the transfer entry. Thus, in the above examples, *book* and *car* will be replaced with *shu1* and *qi4che1* respectively during transfer. The translation here is straightforward in that a source lexical element is replaced with a single target lexical element.

3.1 Structure Change

Between the source f-structure and the target f-structure, often structure changes and/or multiple target words may be needed for the proper translation of a source word in certain contexts. Such transformations are performed by transfer rules, which are lexically driven and specified within transfer lexical entries. For example, many English causative verbs appear in the construction of [V + OBJ] and have a straightforward translation in Chinese, e.g. *enlarge* (*zeng1da4*) and *enrage* (*ji1nu4*); however, some such verbs, e.g. *endanger* and *sadden*, require a causative construction of [V + OBJ + XCOMP] in Chinese. Therefore, *A saddens B* translates as *A shi3 B bei1shnag1*, 'A causes B to be sad'. The necessary f-structure change is specified and instantiated through the transfer entry of *sadden* (3a-3b).

The action instantiated through the transfer entry of *sadden* specifies that (1) the OBJ of the current level unifies (indicated by =) with the

3a. *ec_sadden* ::

-sadden ..		
WORD	FORM 'beilshang1' TECH GENERAL \STD-T-CAUS-V	'sad
\STD-T-TRAN		

3b. *eX_STD-T-CAUS-V* ::

X_STD-T-CAUS-V::			
ACTION	[	<↑ XCOMP SUBJ> = <↑ OBJ>	'1
		TRAN (<↑ FORM> , <↑ XCOMP FORM>)	'2
		REPVAL (↑ , FORM , 'shi3')	'3

XCOMP's SUBJ; (2) *beilshang1* 'sad' is to be the value of FORM in the XCOMP of the current level of the f-structure; and (3) the internally defined REPVAL (replace value) replaces the value of FORM of the current level with *shi3* 'cause'. We will show the f-structures before and after the transfer action in 4a and 4b.

4a. 'before transfer of *eX_STD-T-CAUS-V*

[	FORM 'sadden'	]
	SUBJ [FORM 'A']	
	OBJ [FORM 'B']	

4b. 'after

after

FORM	'shi3'	
SUBJ	[FORM 'A']	
OBJ	[FORM 'B']	-----
XCOMP	[FORM 'beilshang1'	
	SUBJ [---]	-----

3.2 Lexical Selection in Transfer

Frequently a word may have several translations in the target language; the translation selection may depend upon the specific domain of the text, the syntactic context of the word, or the semantic context. For instance, when the technical domain of the input text is specifically related to computers, the translation *mao2bing4* is to be selected for *bug*; otherwise *chong2zi* will be selected, according to the transfer entry of *bug* in 5a.

5a. *ec_bug* ::

WORD {	[	FORM 'mao2bing4'	]
		TECH COMPUTER	
	[	FORM 'chong2zi'	]
		TECH GENERAL	
}			
\STD-T-TRAN			

On the other hand, the syntactic environment of *tell* affects its proper translation: it translates as *shuo1* when *tell* takes an object and on other arguments, as in *tell lies* or *tell a story*; in other contexts *gao4su4* is appropriate. In other words, *gao4su4* is the more common, the default, translation. The transfer entry of *tell* is shown in 5b.

5b. *ec_tell* ::

WORD	[	FORM 'shuo1'	]
		TECH GENERAL	
		\STD-T-PRED-OBJ	
	[		]
	[	FORM 'gao4su4'	]
		TECH GENERAL	
	]		]
		</	

5c. *ex_STD-T-PRED-OBJ* ::

IF	[	<↑ PRED> =c <SUBJ OBJ>	]
ACTION	[	SUCCEED	]

According to the transfer entry of *tell*, *shuo1* will be selected as the translation only when PRED of *tell* has exactly the value <SUBJ OBJ>; =c indicates a constraint. Otherwise, the next translation will be chosen. *Cut*, however, as specified in the transfer entry of 5d, is a transitive verb that translates as *kail* only when the object is *check*; it translates as *jiang4di1* when the object is an abstract noun. Only when the object is not *check* or abstract will the default translation *gie1* be selected.

As shown in 5d-5f, the dag structures of the lexical entries and f-structures allow the translation selection to have access to information in all three areas, i.e. technical domain, syntactic context, and semantic

5d. ec_cut ::

WORD	[	FORM 'kail'	]	'5e
		TECH GENERAL		
		\STD-T-OBJ-CHECK		
	[	FORM 'jiang4dil'	]	'5f
		TECH GENERAL		
		\STD-T-OBJ-ABST		
	[	FORM 'qiel'	]	
		TECH GENERAL		
	)			
	\STD-T-TRAN			

5e. ex_STD-T-PRED-OBJ ::

IF	[	< OBJ FORM > = 'check'	]
		ACTION [SUCCEED	

5f. ex_STD-T-OBJ-ABST ::

IF	[	< OBJ ABST > =c +	]	'ABST = ABSTRACT
		ACTION [SUCCEED		

characteristics, used in evaluating selection conditions associated with each lexical matching. This approach of lexically driven transfer facilitates the modular development of linguistic description of languages and is more flexible than others that integrate source lexical entries with translations, e.g. Li and Erickson (1987). Pentheroudakis (1990a), where examples from the bi-directional English-Korean translation can be found, constitutes a more technical and detailed account of the design of the transfer module and demonstrates an alternative to the projection transfer approach advocated in Kaplan *et al.* (1990). Moreover, Higinbotham (1991, 1990a, 1990c) provides extensive discussions on the use of the semantic co-occurrence networks in resolving lexical selection in the ECS machine translation system.

4. Idiomatic Transfer

Idioms constitute a problem for parsing as well as for translation. There are two approaches for the analysis of idioms: one considers idioms

similar to words as basic lexical units of syntactic structure, and the other treats idioms as phrases whose internal syntactic structures participate in the structure of the sentence (Stock 1987). In the ECS system, three types of idioms in English are recognized: lexical idioms, transfer idioms, and phrasal verbs. While the first approach is adapted for lexical idioms, transfer idioms and phrasal verbs are treated as phrases. The basic criterion used to determine whether a string of words, continuous or not, is an idiom is whether the meaning of this string is the sum of its parts; in other words, an idiom is a multi-word expression whose meaning is a property of the expression as a whole (van der Linden and Kraaij 1990). In the context of a translation system, therefore, this criterion can simply be whether the proper translation of the entire expression can be achieved by the sum of the translation of its individual parts. If not, then such an expression is considered an idiom.

4.1 Lexical Idiom

Two criteria are used to distinguish a string of words as a lexical idiom: (1) the string must be a continuous unit; and (2) its internal syntactic structure, if any, must be of no relevance to the structure of the sentence. In essence, syntactically a lexical idiom behaves just like other basic, non-phrasal lexical elements of its category; thus, a lexical idiom consists of a distinct entry in the English lexicon. Like all other lexical items of the source language, a lexical idiom also has a unique corresponding transfer entry specifying the approach translation of the idiom as a whole. Good examples are multi-word proper names such as *American Airlines*, *the United States of America*, *One-Soon Her*, etc.; conjunctions like *as long as* and *as far as*; prepositions like *in regard to* and *with respect to*; and adverbs like *of course* and *per se*. In short, a lexical idiom is treated just like a word (see 6a-6d).

6a. eI_ "One-Soon Her" ::

'eI = English
idiom

```

[ FS [ FORM 'One-Soon Her' ]
  \STD-N-PERSON
  \STD-PROPER-PLUS
  \STD-CAT-N ]

```

6b. ec_ "One-Soon Her" ::

```

[ WORD [ FORM 'He2wan4shun4'
  \TECH GENERAL ]
  \STD-T-TRAN ]

```

6c. eI "American Airlines" ::

FS [FORM 'American Airlines'] \STD-PROPER-PLUS \STD-USE-GROUP \STD-USE-VHCL \STD-CAT-N	
--	--

6d. ec "American Airlines" ::

WORD [FORM 'Mei3guo2hang2kong1gong1si1' TECH GENERAL] \STD-T-TRAN	
---	--

Note that, in cases where the inner structure of the lexical idiom allows a compositional analysis, e.g. *American Airlines* meaning also *airlines of America*, in order to prevent multiple analyses and resolve ambiguity, alternate analyses may be barred if they use any sub-part of a recognized idiom. In doing so, the conventional and often non-compositional meaning of idioms is preferred.

4.2 Transfer Idiom

Transfer idioms are distinguished according to the following three criteria: (1) the internal syntactic structure of the expression must be of relevance to that of the sentence; (2) its syntactic behaviour must be largely predictable from the normal behaviour of its parts without extra syntactic statement; and (3) the phrase must have both a literal (compositional) and an idiomatic (non-compositional) reading, and therefore must have two possible translations in the target language. As far as the syntax, or parsing, of the source language is concerned, there is no need to treat transfer idioms as in any way special. However, in order to obtain the idiomatic, non-compositional translation, these entries undergo special treatment during transfer, as their name suggests.

Good examples of transfer idioms are verb-noun collocations; for instance, the phrase *to kick the bucket* has a regular reading that is compositional and also an idiomatic reading, *to die*, which cannot be derived from the compositional meaning of *to kick* plus *the bucket*, and therefore has to be recognized as a property of the phrase as a whole. However, that syntactically *the bucket* is the object of *kick* can be accounted for by existing information on the transitive characteristic of *kick*; no extra statement is needed. What does need to be accounted for is that (1) the phrase *to kick the bucket* may have two entirely different

translations, one literal, the other idiomatic; and (2) once the phrase is altered syntactically in certain manner, the idiomatic meaning no longer exists, and only the literal reading remains.

- 7a. He kicked the red bucket.
- 7b. He kicked the buckets.
- 7c. He kicked a bucket.
- 7d. He kicked buckets.
- 7e. He kicked three buckets.
- 7f. The bucket was kicked by him.

None of the sentences in 7 has the idiomatic reading *to die*. We therefore observe that only under certain syntactic circumstances can *to kick the bucket* be idiomatic. However, these required circumstances may vary greatly among transfer idioms; for example, it is possible to passivize *tabs* in the idiom *to keep tabs on* and still maintain its idiomatic reading. It is therefore necessary to specify the minimally required syntactic information in the transfer entry of each transfer idiom's head verb. Since in the translation system only the source f-structure enters the transfer stage, what needs to be accounted for syntactically, to state it more formally in LFG terms, is the fact that, for a sentence to have the idiomatic translation, its f-structure must contain at least all the information specified within the f-structure of the idiomatic phrase. Any conflict will cause the loss of the idiomatic reading and thus will result in the literal translation. The minimally required f-structure for *to kick the bucket* to have the idiomatic translation is specified in 8a.

8a.

[	FORM 'kick'	]
	PRED <SUBJ OBJ>	
[	VOICE ACTIVE	]
	OBJ [	
	SPFORM 'the'	
	FORM 'bucket'	
	NUMBER SG	
	ADJS NONE	
	]	

According to this minimally required f-structure, we can now easily account for the fact that sentences 7a–7f cannot have the idiomatic reading. In 7a the ADJUNCTS of *bucket* is *red* while the transfer entry specifies that its ADJS must be NONE; 7d fails because of a conflict of the value for NUMBER; 7c fails because of the conflict of SPFORM; 7d fails because of NUMBER and SPFORM; 7e fails for the same reason as 7a; and 7f fails because of the conflict of VOICE. Only sentences whose f-structures are subsumed by 8a may have the preferred idiomatic translation, e.g. *Johnny kicked the bucket yesterday*, as shown in 8b.

8b. 'Johnny kicked the bucket yesterday.'

FORM 'kick'	
PRED <SUBJ OBJ>	
VOICE ACTIVE	
TENSE PAST	
SUBJ [FORM 'Johnny']	
OBJ [SPFORM 'the'	
FORM 'bucket'	
NUMBER SG	
ADJS [FORM 'yesterday']	

8c.

FORM 'kick'	
PRED <SUBJ OBJ>	
VOICE ACTIVE	
TENSE PAST	
SUBJ [FORM 'Johnny']	
OBJ [SPFORM 'the'	
FORM 'bucket'	
NUMBER SG	
ADJS [FORM 'yesterday']	

8d. ec_kick ::

WORD [FORM 'til'	
TECH GENERAL	
IDIOM [E1 'kick'	
E2 'the'	
E3 'bucket'	
C1 'si3'	
TECH GENERAL	
\STD-T-IDM-VO2V	
\STD-T-TRAN	

Once it is determined that a sentence can have the idiomatic translation, the necessary f-structure change will then follow. In the case of *to kick the bucket*, the appropriate translation in Chinese is *si3* 'to die', an intransitive verb; hence, the label OBJ in the source f-structure must be deleted, resulting in 8c.

Since there are other similar idioms, e.g. *hit the road*, which have identical f-structure requirements for their idiomatic reading and also need the same transformations on their f-structures, rather than making

8e. *ec_hit* ::

WORD	FORM 'da3' TECH GENERAL	
IDIOM	E1 'hit' E2 'the' E3 'road' C1 'chulfa1' TECH GENERAL \STD-T-IDM-VO2V	
\STD-T-TRAN		

8f. *eX_STD-T-IDM-VO2V* ::

IF	† :- FORM †E1 PRED <SUBJ OBJ> OBJ SPFORM †E2 FORM †E3 NUMBER SG ADJS NONE	
ACTION	TRAN (<† C1> , <† FORM>) DELLAB († , OBJ)	

the requirements unique to the transfer entry of *kick*, we opt to define the manner of translation and structure change in a separate template (shown in 8f). The idiosyncratic source and target lexical forms relevant to the individual idioms would still have to be supplied by the transfer entries, as shown in 8d and 8e.

Note that 'A :- B' indicates that 'A' is subsumed by 'B'. Since, whenever the f-structure of a sentence is subsumed by the minimally specified f-structure of an idiom in a transfer entry, translation of the sentence will proceed according to the conventional meaning of the idiom, ambiguity is resolved again in favour of the idiomatic meaning.

4.3 Phrasal Verbs

Phrasal verbs are verbs with certain prepositions or adverbial particles that they govern. Although the translation of a phrasal verb may not be

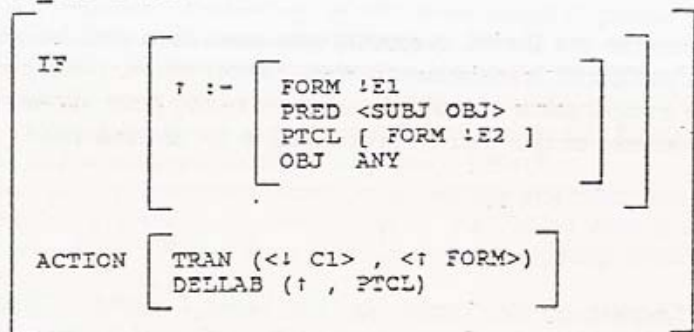
the sum of the translation of its parts, it is different from a transfer idiom in that the internal syntactic structure of a phrasal verb, though also relevant to that of the sentence, cannot be predicted from the general behaviour of the verb and thus needs extra information in the source lexical entry for the correct analysis. Consider the following examples:

- 9a. I will look up her phone number.
- 9b. I will look her phone number up.
- 9c.* I will look down her phone number.
- 9d.* I will look her phone number down.

The verb *look* ordinarily does not behave like a transitive verb; however, *her phone number* in 9a and 9b does behave like an object. Also, although *look* may co-occur with the particle *up* and behave transitively, what particles it may or may not co-occur with in this syntactic construction cannot be predicted, as shown in 9c and 9d. Therefore, for sentences 9a and 9b to parse correctly, it is necessary to have information in the lexical entry of *look* about the particular adverbial particle(s) that it can govern, and only when the specified particle is present can the verb take an object. Furthermore, for the phrase to translate appropriately, it is also necessary to have a lexical transfer entry similar to that of a transfer idiom that specifies the minimal f-structure of the phrasal verb and how it should be translated. And for a sentence to have the idiomatic reading of *look... up*, its functional structure must be subsumed by the f-structure minimally specified in the transfer entry of the head-verb of the phrasal verb, i.e. *look*. Again, in order to maintain generality in accommodating other similar idioms that need the same structure change, e.g. *look...*

10a. ec_look ::

WORD	[	FORM 'kan4'	]
		TECH GENERAL	
IDIOM {	[	E1 'look'	]
		E2 'up'	]
		C1 'cha2'	]
		TECH GENERAL	]
		\STD-T-IDM-VPTCL2V	]
	[	E1 'look'	]
		E2 'over'	]
		C1 'jian3cha2'	]
		TECH GENERAL	]
		\STD-T-IDM-VPTCL2V	]
	}		
	\STD-T-TRAN		

10b. `ex_std-t-idm-vptcl2v ::`

over, we abstract this minimally specified f-structure of the idiom and the necessary operations on the source f-structure from the transfer entries and put them into a separate standard entry (shown in 10b).

The operations specified in ACTION first replace the FORM of the verbs in *look... up* and *look... over* with the appropriate Chinese translation *Cha2* and *jian3cha2* respectively; then, since both Chinese verbs are transitive verbs, the label PTCL (particle) will be deleted by the internally defined function DELLAB (delete label).

5. Conclusion

In this paper we have demonstrated that the LFG-based ECS machine translation system has the capability to handle lexical items as well as idioms efficiently. The implementation of standard inheritance entries and the dag representation of lexical entries and f-structures facilitates the flexible and stratificational development of the lexicon. Accordingly, structure changes and proper translation selection are sufficiently accomplished during the transfer stage.

The treatment of idioms in the ECS machine translation system is based on the following three principles: (1) If the internal syntactic structure of the idiom is relevant to parsing, then the idiom will not constitute a separate entry in the source language. (2) However, if additional information is needed for correct analysis, then such information is specified in the source lexical entry of the head word in that idiom. (3) The account of appropriate idiomatic translation is left to the transfer entries.

This approach of treating idioms clearly observes the 'conventionality principle' that the non-compositional, idiomatic, conventional meaning of phrases should be preferred (van der Linden and Kraaij 1990). By recognizing certain phrases as lexical elements and also shifting the

resolution between the literal, compositional translation and the idiomatic, non-compositional translation to the transfer stage, this approach bypasses the complication of multiple analyses arising from idioms; thus, there is no increase of the load on processing in the analysis stage.

References

- Arnold, D., Krauwer, S., des Tombe, L., and Sadler, L. (1988), '“Relaxed” Compositionality in Machine Translation', in *Proceedings of the Second International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language* (Pittsburg, Pa).
- Bresnan, J. (ed.) (1982), *The Mental Representation of Grammatical Relations* (Cambridge).
- Her, O. (1990), 'Grammatical Functions and Verb Subcategorization in Mandarin Chinese', Ph.D. thesis, University of Hawaii.
- Higinbotham, D., and Pentheroudakis, J. (1989), 'An LFG-Based Machine Translation System', *Computer Processing of Chinese and Oriental Languages* (Taipei).
- Higinbotham, D. (1990a), 'Semantic Co-occurrence Networks and Automatic Resolution of Lexical Ambiguity in Machine Translation', doctoral dissertation, University of Texas at Austin.
- (1990b), 'Digital Thought Mapping: Automatic Translation of Natural Language', in *Proceedings of the Topical Meeting on Advances in Human Factors Research on Man/Computer Interactions: Nuclear and Beyond* (Nashville, Tenn.), 365–9.
- (1990c), 'Semantic Co-occurrence Networks', in *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages* (Austin, Texas), 75–84.
- (1991), 'The Resolution of Lexical Ambiguity in Machine Translation', in *Proceedings of the Eleventh Annual Symposium of the Deseret Language and Linguistics Society* (Provo, Utah).
- Kaplan, R. and Bresnan, J. (1982), 'Lexical-Functional Grammar: A Formal System for Grammatical Representation', in Bresnan (1982), 173–281.
- and Maxwell, J. (1988), 'An Algorithm for Functional Uncertainty', in *Proceedings of COLING 1988* (Budapest), 297–302.
- Netter, K., Wedekine, J., and Zaenen, A. (1989), 'Translation by Structural Correspondence', in *Proceedings of the 4th Conference of EACL* (Manchester), 272–81.
- Kudo, I. and Nomura, H. (1986), 'Lexical-Functional Transfer: A Transfer Framework in a Machine Translation System Based on LFG', in *Proceedings of COLING 1986* (Bonn), 112–14.
- Levin, L. (1990), 'Syntactic Theory and Grammar Design for Machine Translation', in *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language* (Austin, Texas), 109–20.
- Li, M. and Erickson, G. (1987), 'Modular Dictionary Design for Chinese-to-

- English Machine Translation', *Proceedings of 1987 International Conference on Chinese and Oriental Languages Computing* (Chicago), 249-52.
- Nirenburg, S. (ed.) (1987), *Machine Translation: Theoretical and Methodological Issues* (New York).
- Pentheroudakis, J. (1990a), 'Complex Bidirectional Transfer in the ECS English-Korean System', in *Proceedings SICONLP '90* (Seoul), 108-19.
- (1990b), 'You Can Get There from Here: Design and Implementation Issues in Machine Translation Systems', in *Proceedings of the Topical Meeting on Advances in Human Factors Research on Man/Computer Interactions: Nuclear and Beyond* (Nashville, Tenn.), 370-5.
- Sadler, L., Crookston, I., and Arnold, D. (1990), 'LFG and Translation', in *Proceedings of the Third International Conference on Theoretical and Methodological Issues in Machine Translation of Natural Language* (Austin, Tex.).
- Sells, P. (1985), *Lectures on Contemporary Syntactic Theories: An Introduction to Government-Binding Theory, Generalized Phrase Structure, Grammar, and Lexical-Functional Grammar* (Stanford, Calif.).
- Shieber, S. (1986), *Introduction to Unification-Based Approaches to Grammar* (Stanford, Calif.).
- Slocum, J. (1985), 'A Survey of Machine Translation: Its History, Current Status, and Future Prospects', *Computational Linguistics* 11: 1-17.
- Stock, O. (1987), 'Getting Idioms into a Lexicon-Based Parser's Head', *1987 ACL Proceedings* (Stanford, Calif.), 52-8.
- Tucker, A. (1987), 'Current Strategies in Machine Translation Research and Development', in Nirenburg (1987), 22-41.
- van der Linden, E. and Kraaij, W. (1990), 'Ambiguity Resolution and the Retrieval of Idioms: Two Approaches', *COLING-90* (Helsinki), 245-50.