

Semantically Governed Machine Translation of BE-clauses with Adverbs and Prepositional Phrases.

Demonstration of a Ring-Model for German/English, English/German including Analysis-Synthesis, Transformation, Transition, and Generation.

Klaus Günther Schweisthal, Forschungsgruppe LIMAS, Bonn

Research in methods of lexicalization, classification, analysis-synthesis, transformation, transition, and generation, which has been carried on independently in the German and English sections at LIMAS, has recently been consolidated for the purpose of constructing a Ring-Translation-Model for German/English, English/German.

This model has to meet the following requirements:

1. It must produce a translation which is identical in meaning to the original and stylistically acceptable.
2. Some of the difficulties which have heretofore appeared in machine translation must be resolved.

The starting point (Figure 1) is a German adverbial phrase, which is to be analyzed and then converted into a factor formula by means of a morpho-nomo-transformation. An explicative nomo-nomo-transformation together with prepositional governing factors then expand the factor formula for the adverbial expression into a formula for a prepositional phrase, which in turn becomes expressible through a nomo-morpho-transformation. The factor formula for the German prepositional phrase is converted into an English factor formula according to the coordinator's system of maximum agreement. This takes place through a nomo-nomo-transition if the identical factor formula is present in English, otherwise through a nomo-nomo-transformation. A nomo-morpho-transformation then creates an expressible English prepositional phrase which is semantically equivalent to the German. The factor formula for an English adverbial phrase is made possible by a reductive nomo-nomo-transformation plus a nomo-morpho-transformation. The last step of the cycle again takes place in the coordinator. The factor formula for the English adverbial phrase is converted into the corresponding formula for the German adverbial phrase by means of transition or a nomo-nomo-transformation. The morphological realization of this factor formula is identical to that of the initial German adverbial phrase.

A sample output of the ring would be:

Vater ist drinnen ---- Vater ist innerhalb des Hauses ----
Father is inside (of) the house ---- Father is inside ----
Vater ist drinnen.

The automatic reversible process of the ring model is determined by several association lists and function matrices. At first a

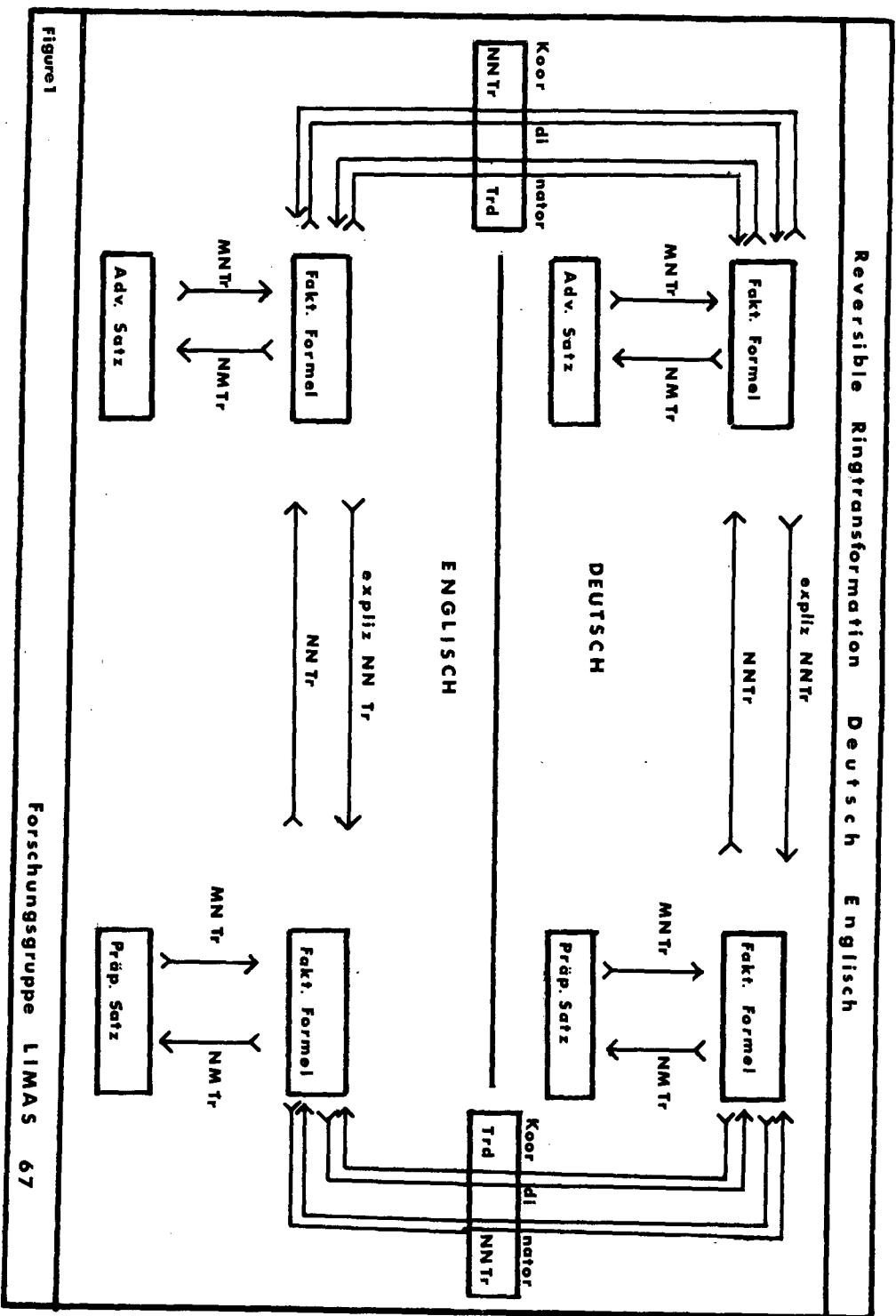


Figure 1

lexicon of German and English semantic factors is necessary.

This list forms the upper row for the necessary nomo-association lists of the German-English words (Figure 2). In this figure are illustrated three different types of semantic association.

List A/D illustrates the association of the factors to the German adverbs, A/P to the German prepositions, and N/D to the various German noun classes. In a strict sense the noun class association is to be specified once again according to the two syntactic functions of the noun in the sentence, i.e. BE-subject* and BE-complement*.

For English the association has been organized in a similar manner in the lists A/E, P/E, N/E, but in such a way that they are independent of German.

An example may make this a little more clear. Assume that A/D₁ is the German adverb "drinnen" and the factors 1, 3, 5, 8, →¹ are coded symbols for the factors Ort, Lage, Einschluß, Betehung der einschließenden Grenzen, Hinweis, then according to the principle of maximum agreement we are able to determine the same factor combination Place, Location, Interior, Accent on boundaries, Reference, in A/E under 1, which in English is the word "inside".

With this same set of factors we find that in the lists P/D and P/E according to maximum agreement the results are P/D₁ and P/E₁, which are expressed as "innerhalb" and "inside of".

In the list N/D all five preposition factors appear in the noun class B, so that if we were to choose as BE-complements the words Garten or Haus the resultant phrase would be: innerhalb des Gartens (Hauses). Because this set of factors also appears under B in N/E, it is possible through transition to produce the English phrase "inside of the garden (house)".

In view of this schematic arrangement one might think that it is possible to associate all German prepositions word for word with English prepositions. (This method was used frequently in the early days of machine translation resulting in untold difficulties.) Such word for word association is almost 100 % possible for the prepositions of place.

But if, for example, we are dealing with the German word nach, which can either be P/D₂ plus factor 7 (Zeit) or P/D₄ plus factor 2 (Ort), we are not able to make a direct association, but must decide according to the method of maximum agreement whether English "after" or "to" is meant.

The corresponding words in A/D and A/E would be "danach"-"dort-hin", and "after that"-"there".

§ See accompanying sheet for LIMAS terminology

NOMO - ZUORDNUNGEN VON WOERNERN DEUTSCH - ENGLISCH

Sprache	Wortarten	Inhalt Faktoren Semantic Factors										
		1	2	3	4	5	6	7	8	9	→ n	
D	Adv. (Liste A/D)	1	■		■		■		■		■	
		2						×				
		3	■		■		■				■	
		4		×								■
		↓	■		■			■				■
		n										
E	Adv. (List A/E)	1	■		■		■		■		■	
		2						×				
		3	■		■		■				■	
		4		×								■
		↓	■		■			■				■
		n										
D	Präp. (Liste P/D)	1	■		■		■		■		■	
		2						×				
		3	■		■		■				■	
		4		×								■
		↓	■		■			■				■
		n										
E	Prep. (List P/E)	1	■		■		■		■		■	
		2						×				
		3	■		■		■				■	
		4		×								■
		↓	■		■			■				■
		n										
D	Nomen Inhalt Klassen (Liste N/D)	A	■		■		■		■		■	
		B										
		C	■		■		■				■	
		↓		×								■
		Q	■		■		■				■	
E	Semantic Noun Classes (List N/E)	A	■		■		■		■		■	
		B										
		C	■		■		■				■	
		↓		×								■
		Q	■		■		■				■	

Figure 2

The selection of the factors Place and Time for the respective noun classes in the noun class lists N/D and N/E would allow the generation of the phrases: "nach dem Mittagessen", "nach Köln", and "after lunch", "to Cologne".

Already it is evident how our system differs from those of earlier machine translation projects.

In the nomo-association list we have established the prerequisites for generation, transformation and transition. In the middle function matrix P/D - P/E (Figure 3) of this diagram are contained the restrictions as to which prepositions may combine with which semantic noun classes to produce a meaningful BE-complement. A quick glance shows that the preposition-noun relationship is differentiated in various ways. This is one proof for the necessity of classifying nouns according to function or semantic classes.

This can be illustrated in the following example. The codes in the preposition column are replaced by prepositions, and the row of noun classes by a representative of each class.

	A-Vater	B-Garten	C-Köln	→Mittag- essen	Q-Gedanke
1. innerhalb		x	x		
2. rechts	x	x	x		
↓ n. nach (Zeit)				x	
n. vor (Abstr. Ber.)					x

Using the lexicon the permitted BE-complements can be programmed from this matrix.

The various possibilities for generating meaningful adverbial BE-clauses can be seen in the matrix N/D - A/D. Here again the necessity of using noun classes is evident, even though some noun classes in the matrix (A, B, C,) appear to be similar.

Example: Vater ist drinnen
 draußen
 rechts
but not danach

But with Mittagessen all four adverbs are possible:

drinnen
 draußen
 rechts
 danach

A further prerequisite for the generation of meaningful BE-clauses containing prepositional phrases is the coordination of the BE-subject noun class with the BE-complement noun class.

The matrix S_{nr}/N/D - S_{ne}/N/D supplies the necessary information.

NOMO - ZUORDNUNG VON SATZGLIEDERN DEUTSCH- ENGLISCH

DEUTSCH	MATRIX	ENGLISCH																																																				
<p>ERLAUBTE ADVERBIAL-SATZE</p> <p>Matrix:</p> <p>N/D — A/D</p>	<table border="1"> <tr> <td colspan="2" rowspan="2"></td> <td colspan="2">Adv.</td> <td colspan="5">Sn E</td> </tr> <tr> <td colspan="2">Nomen</td> <td>1</td> <td>2</td> <td>3</td> <td>→</td> <td>n</td> </tr> <tr> <td rowspan="5">Sn tr</td> <td>A</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> <td></td> </tr> <tr> <td>B</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> <td></td> </tr> <tr> <td>C</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> <td></td> </tr> <tr> <td>↓</td> <td>X</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> </tr> <tr> <td>Q</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>X</td> </tr> </table>			Adv.		Sn E					Nomen		1	2	3	→	n	Sn tr	A	X	X	X				B	X	X	X				C	X	X	X				↓	X	X	X	X			Q						X	<p>PERMITTED ADVERBIAL CLAUSES</p> <p>Matrix:</p> <p>N/E — A/E</p>
				Adv.		Sn E																																																
		Nomen		1	2	3	→	n																																														
Sn tr	A	X	X	X																																																		
	B	X	X	X																																																		
	C	X	X	X																																																		
	↓	X	X	X	X																																																	
	Q						X																																															
<p>ERLAUBTE PRAEPOS. PHRASEN</p> <p>Matrix:</p> <p>P/D — N/D</p>	<table border="1"> <tr> <td colspan="2" rowspan="2"></td> <td colspan="2">P N</td> <td colspan="4">A B C → Q</td> </tr> <tr> <td colspan="2">Seins-ergän-zurig</td> <td>1</td> <td>2</td> <td>3</td> <td>→</td> <td>n</td> </tr> <tr> <td rowspan="4">Be-Com-plement</td> <td>1</td> <td></td> <td>X</td> <td>X</td> <td></td> <td></td> </tr> <tr> <td>2</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> </tr> <tr> <td>↓</td> <td></td> <td></td> <td></td> <td>X</td> <td></td> </tr> <tr> <td>n</td> <td></td> <td></td> <td></td> <td></td> <td>X</td> </tr> </table>			P N		A B C → Q				Seins-ergän-zurig		1	2	3	→	n	Be-Com-plement	1		X	X			2	X	X	X			↓				X		n					X	<p>PERMITTED PREPOS. PHRASES</p> <p>Matrix:</p> <p>P/E — N/E</p>												
				P N		A B C → Q																																																
		Seins-ergän-zurig		1	2	3	→	n																																														
Be-Com-plement	1		X	X																																																		
	2	X	X	X																																																		
	↓				X																																																	
	n					X																																																
<p>ERLAUBTE FUNKTIONEN</p> <p>Sntr /N Inh. Klasse —</p> <p>Sn E /N Inh. Klassen</p> <p>Matrix:</p> <p>Sntr/N/D — SnE/N/D</p>	<table border="1"> <tr> <td colspan="2" rowspan="2"></td> <td colspan="2">Sn E tr</td> <td colspan="4">A B C → Q</td> </tr> <tr> <td colspan="2">Sntr</td> <td>A</td> <td>B</td> <td>C</td> <td>→</td> <td>Q</td> </tr> <tr> <td rowspan="5">Sntr</td> <td>A</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> </tr> <tr> <td>B</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> </tr> <tr> <td>C</td> <td>X</td> <td>X</td> <td>X</td> <td></td> <td></td> </tr> <tr> <td>↓</td> <td>X</td> <td>X</td> <td>X</td> <td>X</td> <td></td> </tr> <tr> <td>Q</td> <td></td> <td></td> <td></td> <td></td> <td>X</td> </tr> </table>			Sn E tr		A B C → Q				Sntr		A	B	C	→	Q	Sntr	A	X	X	X			B	X	X	X			C	X	X	X			↓	X	X	X	X		Q					X	<p>PERMITTED FUNCTIONS</p> <p>BeS/Sem. Noun Class —</p> <p>BeC/Sem. Noun Classes</p> <p>Matrix:</p> <p>BeS/N/E — BeC/N/E</p>						
				Sn E tr		A B C → Q																																																
		Sntr		A	B	C	→	Q																																														
Sntr	A	X	X	X																																																		
	B	X	X	X																																																		
	C	X	X	X																																																		
	↓	X	X	X	X																																																	
	Q					X																																																

Figure 3

For example, a BE-subject from the noun class "abstract" is compatible only with a BE-complement from the same class.

	Die Idee ist hinter dem Gedanken
<u>but not</u>	Der Gedanke ist hinter dem Garten
	Der Garten ist hinter dem Haus
<u>but not</u>	Der Garten ist hinter dem Mittagessen

Not until all of these factors have been formalized and all conditions which were illustrated in the matrices are taken into consideration it is possible to set up a final function matrix. This matrix then aids the translation process in the generation of meaningful BE-clauses with prepositional phrases. (Figure 4)

For example:

A	is to be replaced by	Vater	-	(father)
B		Haus	-	(house)
C		Köln	-	(Cologne)
→		Mittagessen (action)	-	(lunch)
Q		Gedanke	-	(thought)
9		vor (place)	-	(in front of)

then

NPN/D	-->	A 9 A	-->	Vater ist vor dem Vater
		A 9 B	-->	vor dem Haus
		A 9 C	-->	vor Köln
		A 9 →		<u>(not permitted!)</u> Vater ist vor dem Mittagessen
		A 9 Q		<u>(not permitted!)</u> Vater ist vor dem Gedanken

NPN/E	-->	A 9 A	-->	Father is in front of the father
		A 9 B	-->	in front of the house
		A 9 C	-->	in front of Cologne
		A 9 →		<u>(not permitted!)</u> Father is in front of lunch
		A 9 Q		<u>(not permitted!)</u> Father is in front of the thought

In order to show that this model exists not only on paper but also in practise, we have compiled a German and English lexicon consisting of 16 prepositions, 16 adverbs, and one noun from each of the five noun classes. The upper row of the nomo-association chart, which was seen in figure 2, is filled with over 30 semantic factors.

The outputs show that the ring model works satisfactorily.

With the aid of these matrices (Figures 1 - 4), a semantically equivalent German/English, English/German translation can be executed for all the cases in which the function matrices in both languages are identical in structure, e.g. (NPN/D = NPN/E) (transition).

For those cases in which the function matrices are not identi-

S_{nE}
S_{nPr}

Erlaubte ist-Sätze mit
Pröp.-Phrasen

KOMPLEXE MATRIX

NPN/D

	A	B	C	→	Q
A	9	14 2	1 9	5	
	14 2	3 7	14 2	10	
3 7		4	3 7	15	
4		12	4		
	9	1 9	1 9		
	14	2	14 2		
3 7		3 7	3 7		
4		4	4		
	9	1 9	1 9		
	14 2	3 7	2		
3 7		3 7	3 7		
4		4	4		
	9	1 9	1 9	5	
	14 2	14 2	14 2	10	
3 7		3 7	3 7	15	
4		4	4		
					13
					11
					8
					16

Permitted is-clauses
containing prep-phrases

KOMPLEXE MATRIX

NPN/E

○ nur für "geht"

Figure 4

cal (NPN/D \neq NPN/E), the proper translation is insured by means of an auxiliary program. This takes place through verbalization of the semantic factors of the source language in the target language.

For example: The English sentence The candle is inside (of) the father cannot be translated into German as it now is because there is no corresponding matrix NPN/D. The verbalization program, however, makes possible the translation Die Kerze ist im Körper des Vaters eingeschlossen.

If a German adverb has no English counterpart, an auxiliary program will create a semantically equivalent prepositional phrase, e.g. dahinter ---- behind it. All German adverbs must be provided with the possibility of such a transformation, so that in the reverse process - in case it should be necessary - the proper German adverb can be selected for the English combination preposition + pronoun.

Some Terms of LIMAS-Terminology

1. Transformation

- a) Morpho-transformation: A change within the morpho-structure involving formal constituents.
- b) Morpho-semantic-transformation: A change within the morpho-level involving those semantic factors which are referred to as "Denkweisen". These are: sentence constituent, framed, stressed, cyclopic, nominal, actional, and designative.
- c) Morpho-nomo, and nomo-morpho-transformation: A transfer from the morpho-level to the nomo-level or vice versa. (This is similar to a change from the deep structure to the surface structure and vice versa in Transformational-Generative Grammar.)
- d) Nomo-nomo-transformation:
 - 1) Within the same language: A change of the factor formula by means of exchange, addition, or elimination of certain factors.
 - 2) Between the source language and the target language: The factor formula of the source language is not acceptable in the target language, and must therefore be reshaped.

2. Transition

The simple transfer of a factor formula from one metalanguage into another.

3. Explication

To make explicit a semantic factor in a sentence which contains no morpheme to represent that factor. The morpheme can be lacking for three reasons:

- a) ellipsis
- b) non-expressed context relations
- c) a combination of semantic factors which are understood but not expressed.

It is therefore important to differentiate between a) ellipsis explication, b) context-relation explication, and c) factor explication. These three types of explication can, however, be combined.

4. Verbalization

The process of verbally expressing a semantic factor within the morpho-level which has been determined through explication.

5. Reduction

The opposite of explication, i.e. a factor formula of the meta-language is not permitted to be verbalized. The factor is not really eliminated, but is implied within the morpho-level.

6. Explicating-transformation

- a) Explicating nomo-nomo-transformation: A semantic factor which has been determined through explication changes the factor formula.
- b) Explicating nomo-morpho-transformation: The semantic factor becomes verbalized.

Faktorenformel	-	factor formula
Funktionsmatrix	-	function matrix
Inhaltfaktor	-	semantic factor
ist-Satz	-	Be-clause
Nomeninhaltklasse	-	semantic noun class
Seinsträger (Sntr)	-	Be-subject (BeS)
Seinergänzung (SnE)	-	Be-complement (BeC)
Traduktion (Trd)	-	transition
Transformation (Tr)	-	transformation
Zuordnungsliste	-	association list

Criteria for "meaningful" BE-clauses with adverbs and prepositional phrases:

- a) Approximately 40,000 examples of modern German taken from various sources.
- b) A smaller collection of English examples.
- c) Interviews with several native speakers of English from England and America.