

1965 International Conference

on

Computational Linguistics

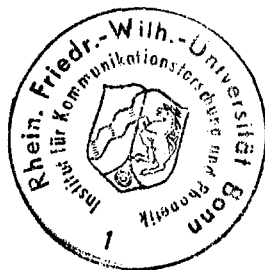
DATA PREPARATION FOR SYNTACTIC TRANSLATION

L. W. Tosh

The University of Texas

P. O. Box 8611

Austin, Texas 78712



TOSH

ABSTRACT

The following paper discusses the preparation of syntactic data for use in a generalized language translation system, developed by the Linguistics Research Center at The University of Texas. Capabilities and limitations of translation by syntactic model are outlined and compared with the word-for-word model.

BACKGROUND

In January of this year the Linguistics Research Center held its first demonstration of an operational system for experimental translation of languages. We prepared a limited set of test data and used a pre-selected input text to demonstrate the operational status of computer programs in the system. I shall discuss briefly the model on which the translation system is based and the preparation of linguistic data used in the demonstration.¹

LINGUISTICS RESEARCH CENTER

Two principal objectives at the Linguistics Research Center have been the development of a generalized automatic translation system and the development of a linguistics computer system consisting of programs designed to facilitate the collection and maintenance of data for the translation system [7]. In addition to these objectives, we have undertaken related studies in information retrieval and automatic classification [1, 2]. The philosophy behind our research effort may be characterized as one of seeking general solutions to language description and translation as opposed to one of designing specialized

algorithms. The general principles underlying our research have been discussed elsewhere, and I shall not dwell on them here [4; 5; 6; 8, pp. 3-14; 9].

Three organizational subdivisions of the Center are the Theoretical Linguistics Group, the Descriptive Linguistics Group and the Systems Group. Activities at the Center are distributed over these and other specialized areas in order to facilitate research. Results reported in this paper are presented from the point of view of activities in the Descriptive Linguistics Group.²

The Descriptive Linguistics Group is currently engaged in maintaining research data in six languages: Chinese, English, German, Hebrew, Russian and Spanish. We are also maintaining data for independent, non-supported research in Hindi and Old Saxon. We have just begun maintaining data for Japanese. Plans are being made to add French to the data in the Linguistics Research System in the near future.

LINGUISTICS RESEARCH SYSTEM

The Linguistics Research System is a hierarchical system of computer programs, which, in addition to programs in the experimental translation system, includes programs designed to support a stratified description of language data

(see fold-out entitled LINGUISTICS RESEARCH SYSTEM). In the illustration the large boxes marked MAINTAINANCE at the upper and lower part of the page represent the system of programs in which we collect and maintain language and descriptive linguistic data. The system of large boxes running across the middle of the page represents the translation system. Details of these programs will be found in [8, pp. 83-103]. I outline the functions of programs in the translation system below.

TRANSLATION MODELS

Various models have been proposed for automatic translation of languages. The models have been characterized into at least three levels of increasing complexity and sophistication: 1. Word-for-word, 2. Rule-for-rule or syntactic, 3. Transformational-semantic. The inadequacies of type 1. are known. Most of current investigation is concentrated in some form or other on type 2., while type 3. models remain largely speculative. Translation programs have been completed which will simulate models 1. and 2.

In model 1. we may perform word-for-word translation by presenting an input corpus (see fold-out) to

the LEXICAL ANALYSIS program. Analysis results in recognition of whatever forms have been defined in the lexical grammar. The results are transferred from the analysis program in MONOLINGUAL RECOGNITION to the LEXICAL ANALYSIS program in INTERLINGUAL RECOGNITION. Intermediate display programs are ordinarily by-passed in the translation mode. The data then pass to an INPUT TRANSFER tape before entering the TRANSFER program. This program processed INPUT TRANSFER data against data from the INTERLINGUA tape to produce an OUTPUT TRANSFER tape. OUTPUT TRANSFER data pass into the LEXICAL SYNTHESIS program in INTERLINGUAL PRODUCTION to be converted to an acceptable form for input to LEXICAL SYNTHESIS in MONOLINGUAL PRODUCTION. The resulting data pass on to the OUTPUT CORPUS tape which serves as input to the CORPUS DISPLAY program.

Output from this lowest level of translation would be word-for-word, morph-for-morph, etc. matching the order of input forms. There would be no control over output morphology or syntax. We have not considered it worthwhile to attempt to use model 1. translation independently of model 2.

Model 2. translation in the Linguistics Research System performs in a fashion operationally similar to model 1. Instead of operating (horizontally on the fold-out) directly through the lexical level, however, we initiate

the translation input in LEXICAL ANALYSIS and pass the resulting data (vertically) into SYNTACTIC ANALYSIS.

Model 2. translation now continues horizontally on the syntactic level analogously to the manner described for the lexical level.

Output resulting from the syntactic translation model observes the requirements for well-formedness in output language morphology and syntax. Examples from the January demonstration are given below. With large volumes of grammar data, this model is not expected to provide all the semantic collocational controls which we as linguists will want to maintain. Nor will it properly account for problems such as pronominal reference. These and other transformational problems will be dealt with in a still higher order of description and programming. The semantic order of programming has only recently been undertaken.

The translation model used in the January demonstration is essentially a type 2. model, although it contains some features proposed for type 3. models. Analysis is performed on the input language with a context-free phrase structure grammar. The structures which are thus identified are transformed into equivalent output language structures by the so-called transfer grammar. Translation output is then generated through a context-free phrase structure grammar of the output language [13].

Rules for use in a similar model are given by Ilse Langerhans [3]. The essential difference, however, between our model and that proposed by Langerhans is that in the latter the input language is analyzed into kernels, the kernels matched with equivalent output language kernels, and the output language kernels transformed into finished expressions.

PREPARATION OF DATA

For the demonstration, we selected a text in psychology to use as a test corpus in German, the input language (Appendix A). The corpus consisted of the first six paragraphs of an essay appearing in UNIVERSITAS [10]. Members of our staff then prepared an English translation to be used as a test corpus in the output language (Appendix B). We use test corpora for verifying the morpho-syntactic description in each language before attempting to use the grammars in the translation system. To illustrate the details of data preparation, I have chosen the second sentence from the third paragraph of text (Fig. 1). This sentence was chosen for reasons of simplicity and economy of description. It is typical, however, of transformational problems in syntactic translation. We pro-

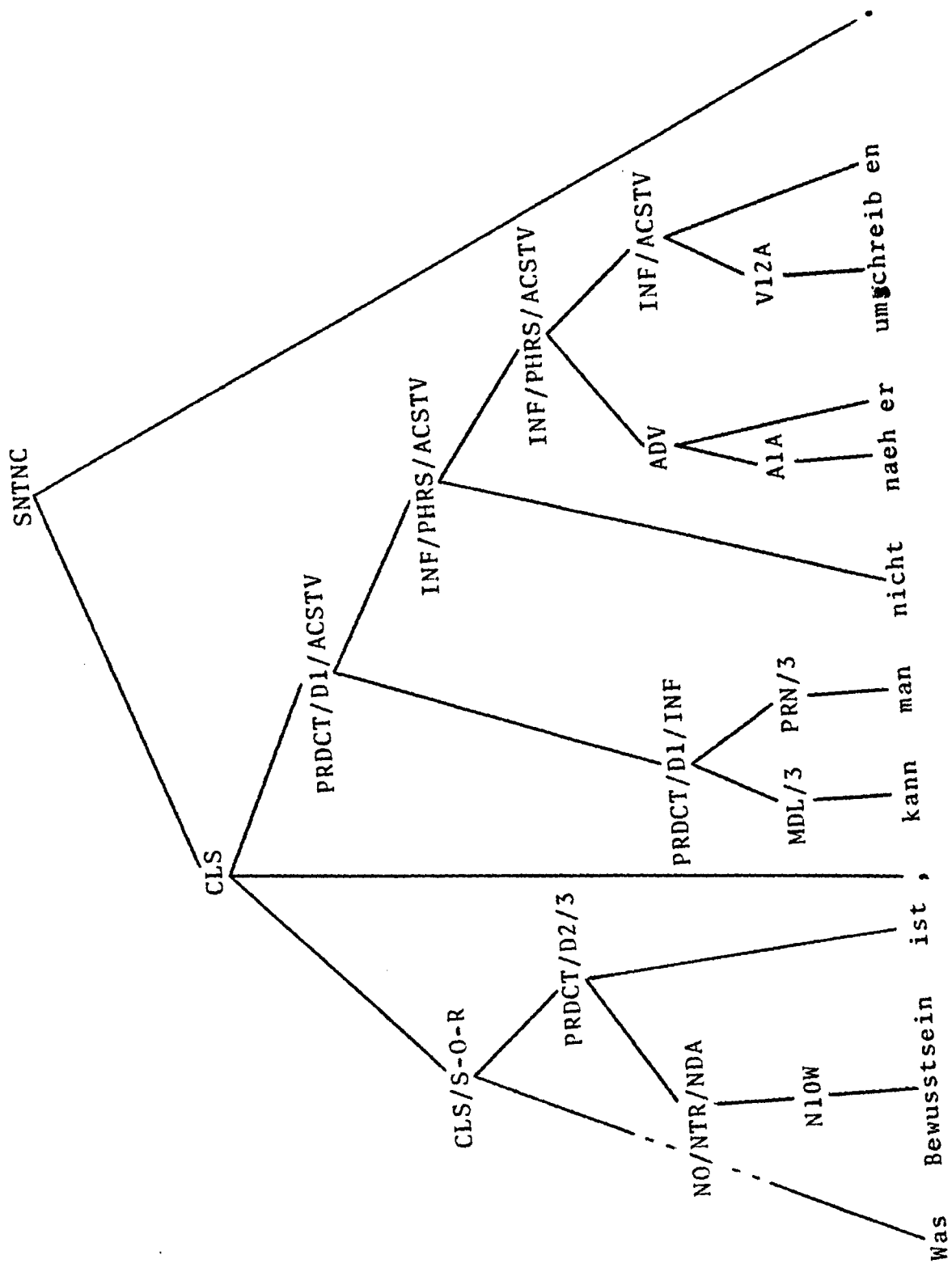


Figure 1

vided a phrase structure description for the sentence, labelling those features of construction which would be necessary for morpho-syntactic (as opposed semantic) grammaticality in German. The description contains, therefore, more information than is necessary for recognition. But we are designing our grammars, in general, for bi-directional use. A similar description was provided for the English translation (Fig. 2).

After diagramming each sentence, we encoded the information contained in the diagrams in an equivalent phrase structure notation [14]. The data were then compiled in the computer system. As rules are compiled for each language, each rule is randomly assigned a permanent identification number. After the respective grammars are compiled and displayed, we refer to them for the identification of each rule and record the appropriate number by each occurrence of a rule in the diagram. The diagrams then appear as in Figures 3 and 4.

VERIFICATION OF DATA

To insure that a description for any given sentence is complete, we perform analysis on the sentence in the computer, using the grammar data accumulated up to that point. If automatic analysis is successful, we ex-

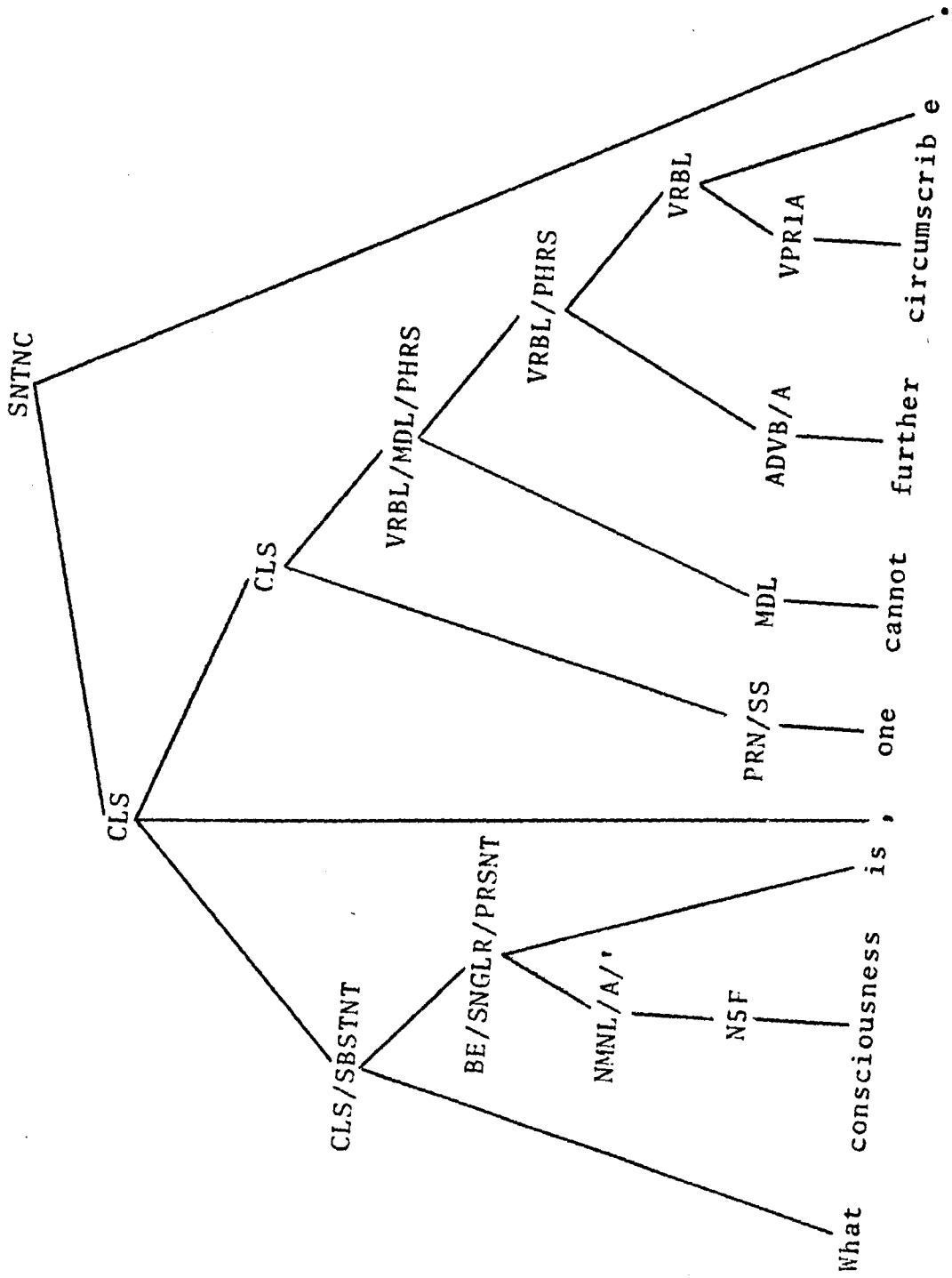


Figure 2

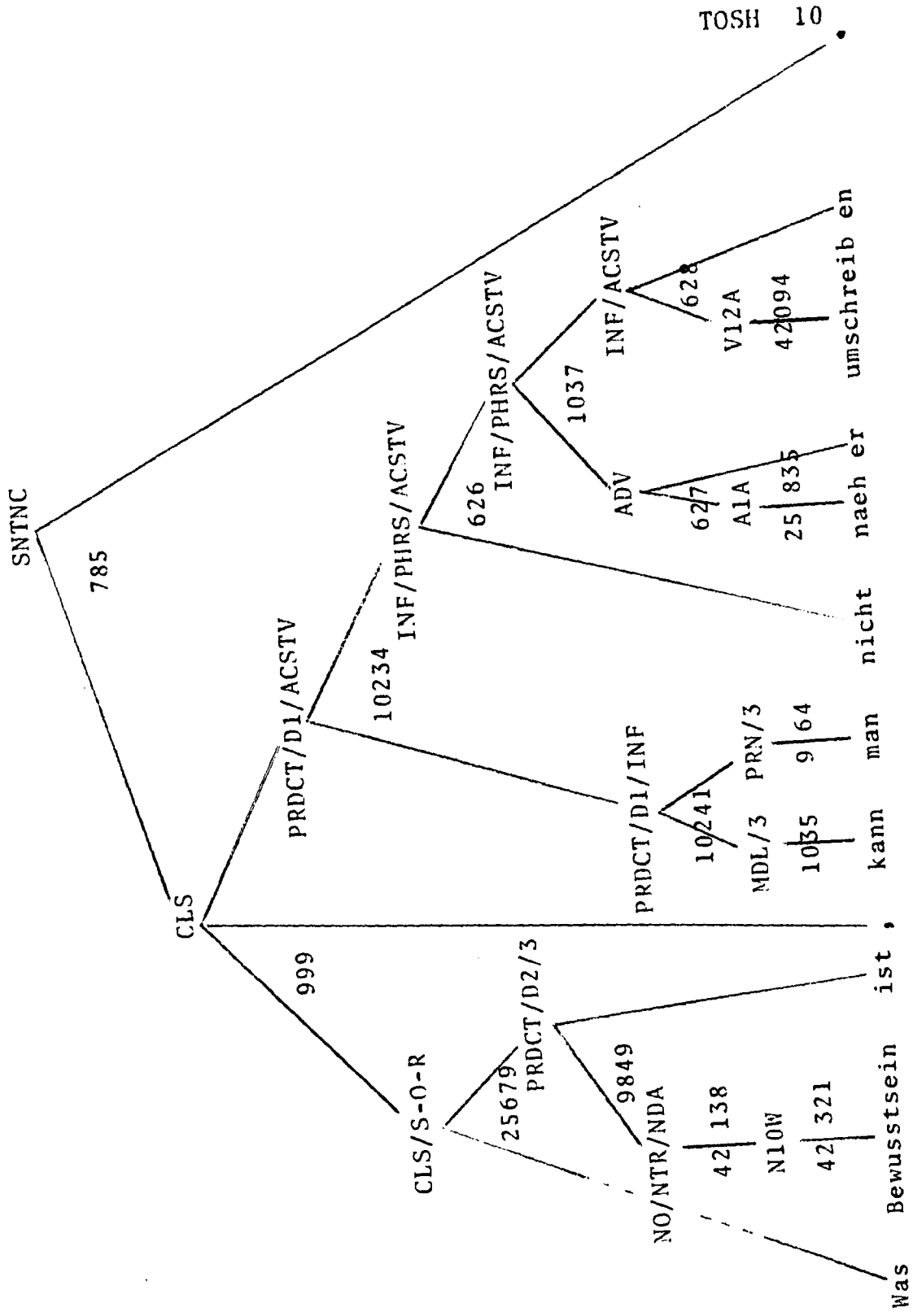


Figure 3

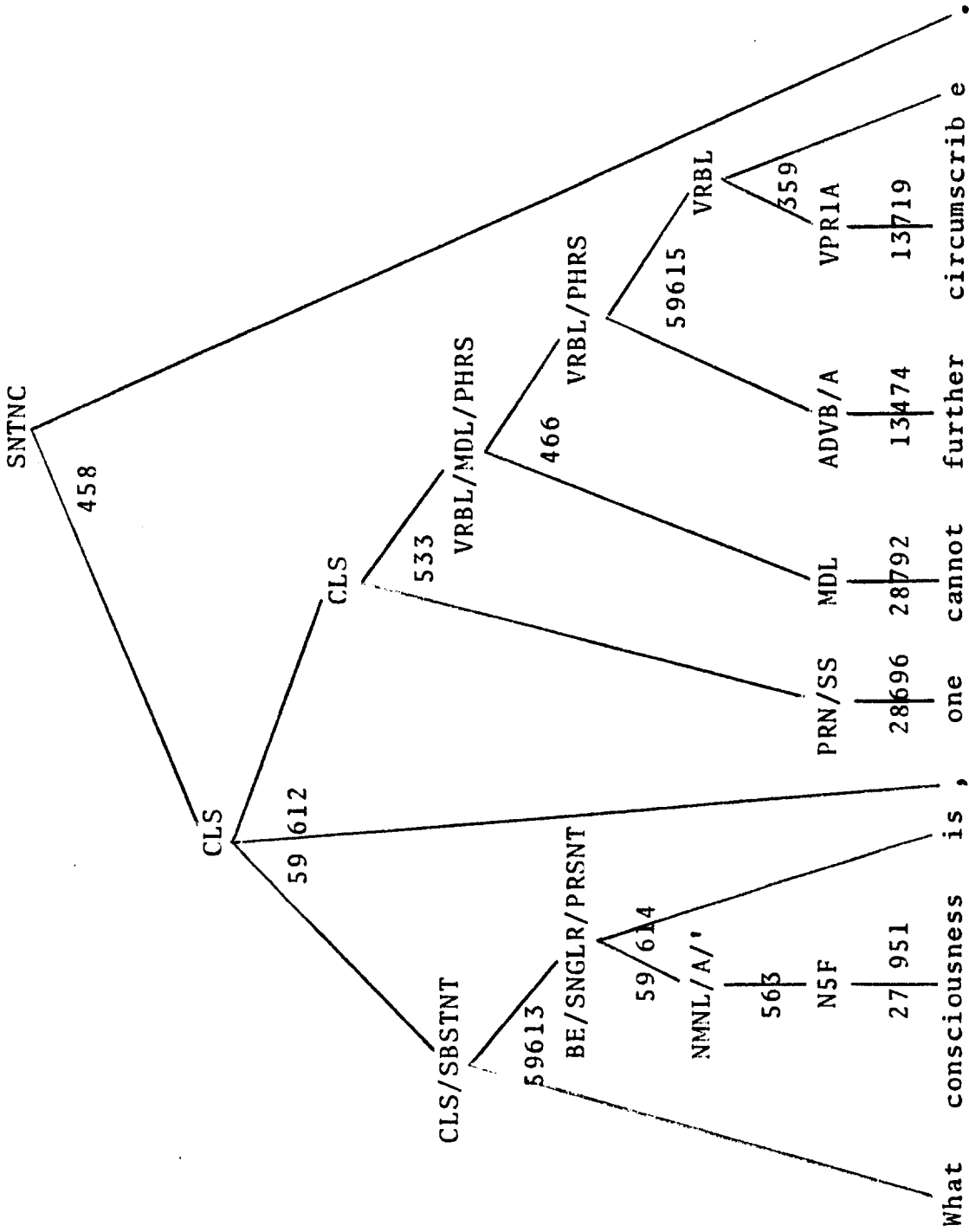


Figure 4

pect to see at least the analysis output corresponding to the information recorded in the diagram for the sentence. Often there are alternative analyses. If automatic analysis is incomplete, we reconstruct the rules needed and (re)compile them in the grammar. I shall not go into the details of analysis here, as they have been presented elsewhere [8, 12, 13].

TRANSLATIONAL TRANSFORMATIONS

After we verified the descriptions in each language, we went on to define the basis of interlingual transformation relationships. We selected a pair of sentences, one from each of the two languages. They are defined as equivalent in meaning by bi-lingual informants. Given the pair of sentences, we mapped corresponding sub-structures from one sentence on to the other. This information was recorded on the diagrams by circumscribing the sub-structures (Fig. 5). Normally these lines are added directly to the diagrams. For the sake of simplicity, I have omitted branching diagrams and class names from the illustrations. After we established the correspondences between each pair of sub-structures, we inspected each sub-structure to see of what it was

composed. I have represented this information in Figures 6 and 7 by the rule number(s) contained in each sub-structure.

Suppose now we want to "transform," i.e., translate the expression Bewusstsein into the expression consciousness. Bewusstsein (Figs. 3, 6) is represented by the rule

42321: N10W → Bewusstsein

Consciousness (Figs. 4, 7) is represented by the rule

27951: N5F → consciousness

We define the equivalence of these two expressions by writing the bi-directional transformation T_x :

$$[42321]_g \leftarrow T_x \rightarrow [27951]_e$$

This is equivalent to writing a reversible transformation between the structures of Figure 8.

Figure 8

N10W		N5F
Bewusstsein	↔	consciousness

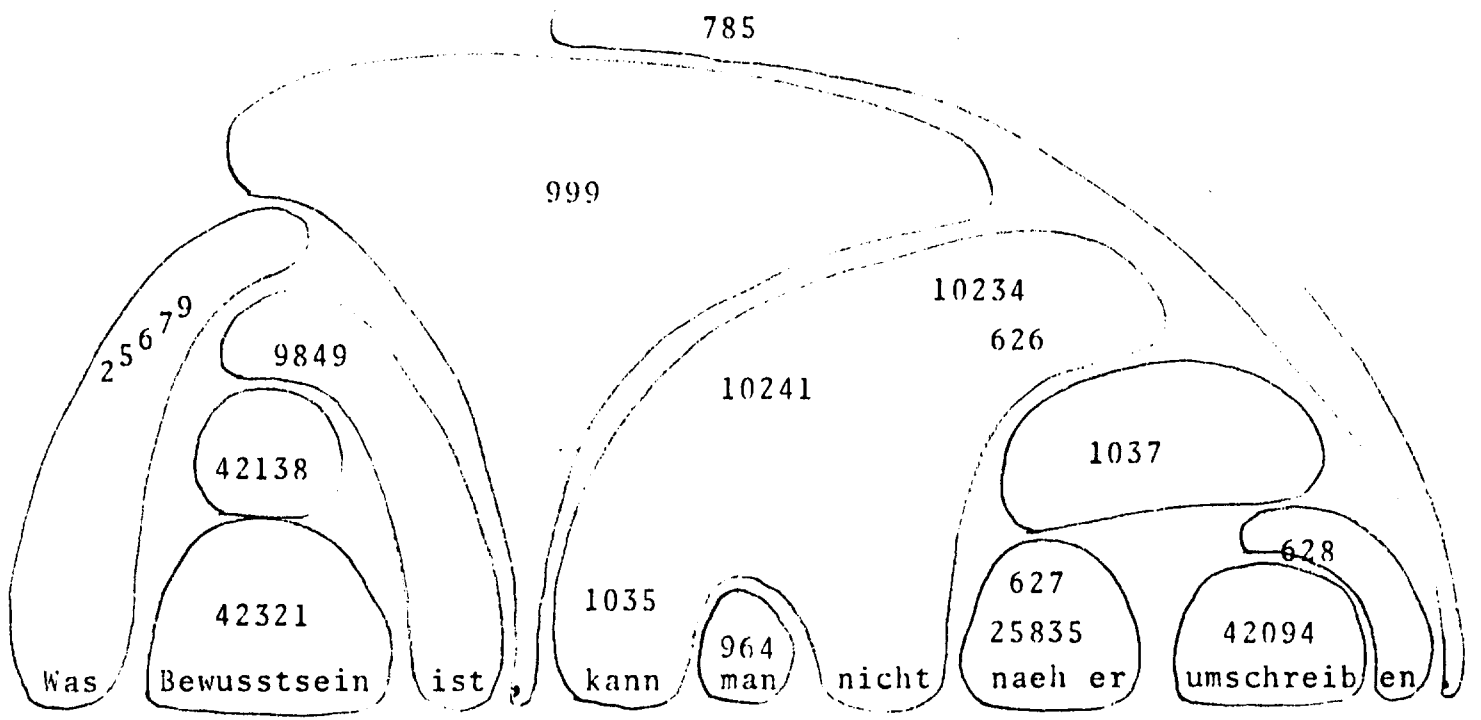


Figure 6

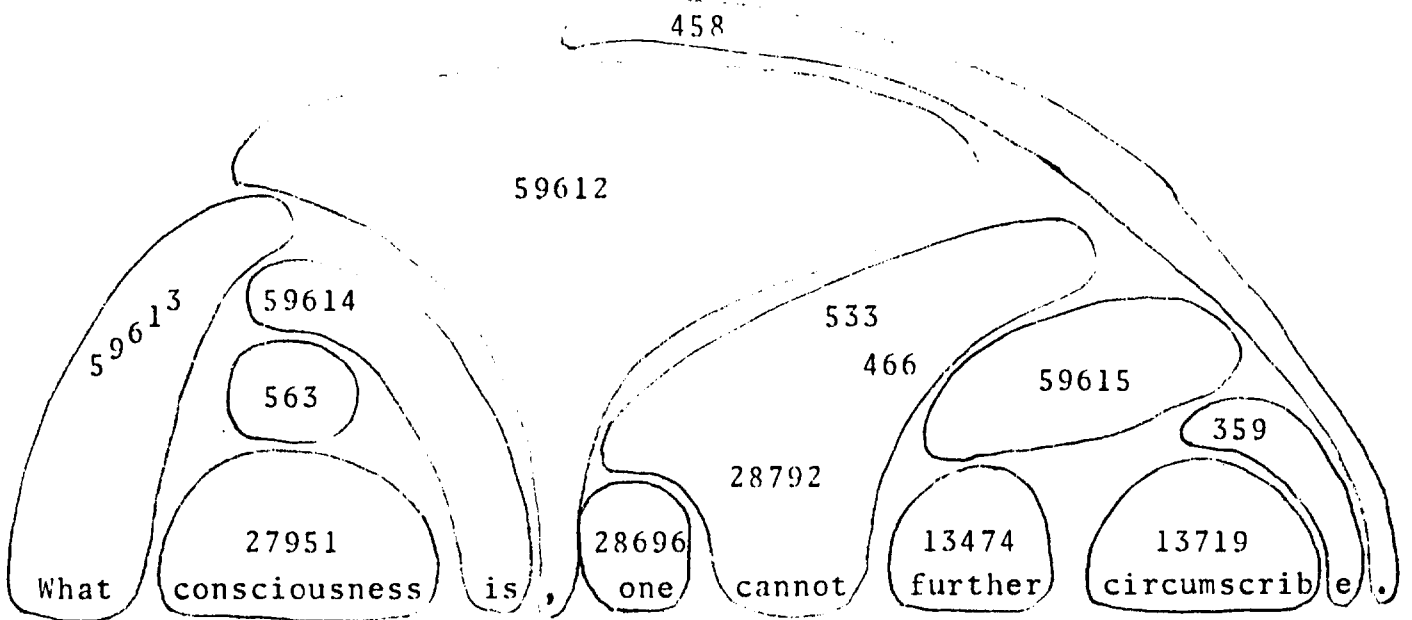


Figure 7

Similarly, we may translate from an infinitive construction in the one language into a corresponding construction in the other. The infinitive of umschreib- is formed with -en by the rule

628: INF/ACSTV \rightarrow V12A + en

The corresponding English construction is formed by the rule

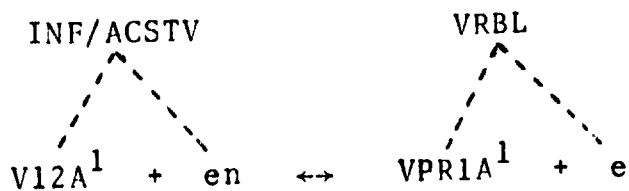
359: VRBL \rightarrow VPR1A + e

We record thus the transformation T_y

$[628]_g + T_y \rightarrow [359]_e$

to define the translation equivalence. This is equivalent to writing the transformation in Figure 9.

Figure 9



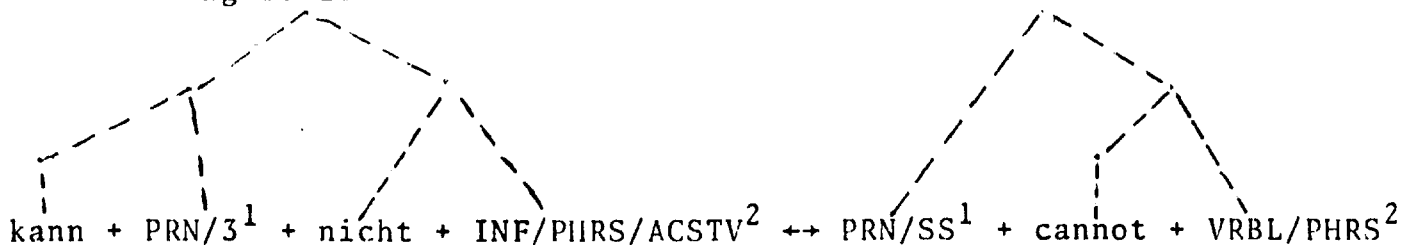
The foregoing examples are typical of the many rule-for-rule correspondences to be found in a pair of structurally similar languages.

Of greater interest are those transformations of pairs of structures which are dissimilar in terms of constituent rules. In Figure 6 the sub-string kann (man) nicht (naecher umschreiben) is analyzed in part by the rule sequence 10234 + 10241 + 1035 + 626. The sub-string consists, furthermore, of a subject-verb inversion characteristic of German syntax. We may transform this construction of four rules into the corresponding English construction (Fig. 7) of three rules 533 + 466 + 28792 by writing the transformation T_z :

$$[10234 + 10241 + 1035 + 626]_g \xrightarrow{T_z} [533 + 466 + 28792]_e$$

This is equivalent to writing a transformation on the structures in Figure 10.

Figure 10



The transformation brings us from the subject-verb inversion of German into the normal subject-verb order for English. Superscripts are associated with all class names in phrase structure rules in order to maintain proper order of content substitution during transformation from one structure to another [13, pp. 12f, 51-66].

TRANSLATION OUTPUT

After all translation data have been collected and compiled for a given test corpus, the next step is to verify the data in the computer system by attempting to carry out automatic translation. As in the case of automatic analysis, we expect translation output corresponding at least to the target language structures for which we have set up translation rules. That is, we expect in the case of successful translation an output which resembles within satisfactory limits the human translation given as the ideal goal. There may be, in addition, various alternative paraphrases, but the content should be essentially the same. The more likely case in the beginning stages, naturally, is partial success mixed with failure.

Our first output for German to English translation is given in Appendix C. The unsatisfactory quality

in this example is the result of a combination of program errors and inadequate linguistic data. Word-for-word output would produce results quite similar to this sample. Receiving such results, we referred back to the appropriate sentence diagrams and lists of translation rules to reconstruct the rules necessary for well-formed output. A subsequent run with the needed additional translation rules is displayed in Appendix D.³

If we compare the computer translation (Appendix D) with the human translation (Appendix B), they appear quite similar at first glance, as indeed we should hope they would be. A closer inspection, however, reveals numerous differences. Some of these result from weaknesses in description as limited by the model, while some result from the alternatives implicit in the descriptive data -- alternatives which the model is designed to cope with.

In the first or title paragraph, the German title is constructed in the framework of a prepositional phrase beginning with ueber. Since the human translation was prepared without a preposition, transformation rules were set up to delete the preposition accordingly in the computer version of the English output. This is probably not advisable, however, since in the syntactic model there

is no satisfactory way to distinguish contextually a prepositional phrase functioning as a title from its other uses. The implication is, then, that we should reformulate our transformation for this context to produce an English preposition like on.

The human and machine translations are identical in the first sentence of paragraph 1 denoted by the numbers 74 001 in the left margin (Appendices A, B, D). The German adverb allein, which is an element in the relative clause modifying the subject-noun head, has been transformed into the English adverb only, which now is a member of the corresponding English subject-noun head construction and not an element of the following relative clause. For the German clause das Problem...so verzweigt, we have transformed into the corresponding English clause the problem...so complex, inserting a copula verb is. Finally, in the last clause of the German sentence there is a passive construction which has been transformed into an equivalent English active construction. There are transformations of similar complexity throughout the remainder of the corpus.

There is an interesting difference between the last sentence of the human translation of paragraph 1 and the machine translation. In the human translation the sentence ends ...problem of a dependence of mental processes on the body. In the machine translation the sen-

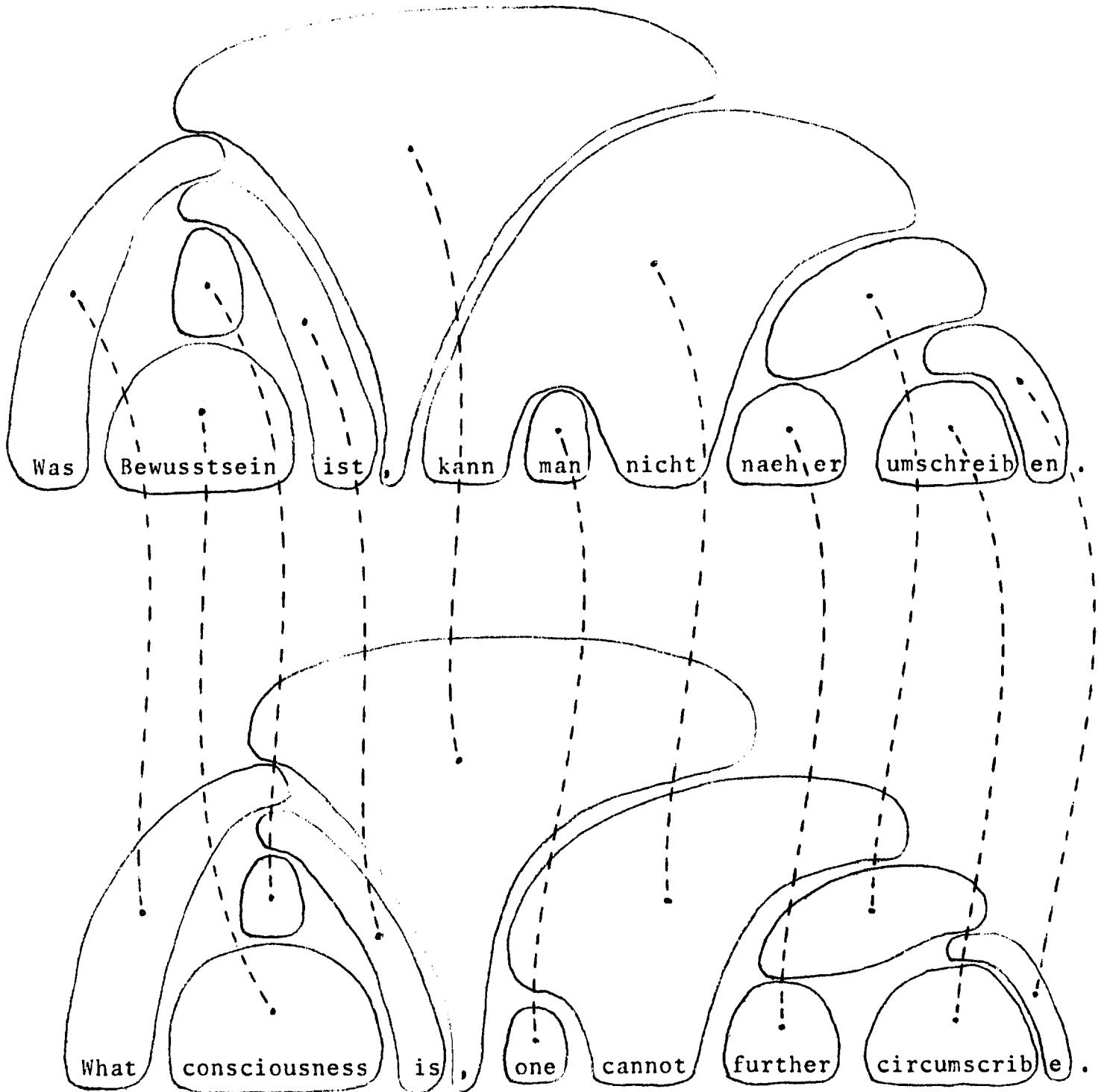


Figure 5

tence ends ...problem of a physical dependence of mental processes. Although all the necessary grammar rules and transformations were available to the translation system for producing an output identical with that of the human translation, it is interesting that the system picked instead an alternative paraphrase (and a potentially confusing one) which was more similar to the syntax of the original German input. The system's choice was made on the basis of certain probability parameters available to it and with which we are in continual experimentation. It is not surprising that the system selected such an alternative, for we expect such to be the case in the present model. What is interesting, however, is the fact that a choice was available even within the limited data set which we prepared for these few paragraphs. For this experiment the system had available to it dictionary data for the entire article of 32 paragraphs. With respect to syntactic data, however, it was quite limited since we supplied just the rules necessary to carry out analysis and/or synthesis of the six paragraphs involved in the experiment. Furthermore, we had limited ourselves in the transformation data to a choice of one syntactic output for each sentence -- the output identical with that of the human translation. Nonetheless, it is evident that in this small data set there are already sufficient

implicit relationships to permit unplanned for if not unexpected paraphrases.

LIMITATIONS IN THE MODEL

Paragraph 2 of Appendix D contains probably the most frequent and characteristic examples of deviation from an ideal output. The paragraph contains a number of aberrant pronominal forms. Since German contains the forms er, es, sie and all their variant case forms and since all these forms are ultimately correlatable with all forms of English he, she, it, it follows that we may generate any one of the English third singular pronouns from any one of the German third singular pronouns. In the model presented here, we have not, for instance, classified nouns on the basis of such features as gender, animateness, concreteness, etc. Thus, in the first sentence of paragraph 2, we have not classified either reader or brain as to referential gender. Consequently, at the moment when the translation system is prepared to generate a pronoun following the sequence ...at this moment when..., the English grammar is so constructed and tied into the transformation-transfer data that the system may generate (just the proper case form of) all three third singular pronouns. Which one is

generated depends on which rule has the highest probability value, in this case the rule producing the expression it, since this is the most frequent of the pronouns in the text.

It is not clear that the proper choice of English pronoun gender could be specified even if we included in the syntactic description such features as gender, animateness, etc. For some instances of pronoun-antecedent agreement will remain ambiguous, given two or more antecedents. The ambiguity occasionally cannot be resolved without resort to reference to the extra-linguistic environment. The first sentence of paragraph 2 is perhaps a case in point. Given the general context of psychology in which the test corpus was written, it is conceivable that either the pronoun he or it could refer back to the appropriate respective antecedents reader or brain.

In those cases where pronoun-antecedent agreement can be stated within the linguistic environment, we should of course be prepared to build such features as gender, animateness, concreteness, countableness, and a host of other such features into our grammars--features which have been difficult to account for systematically before the advent of stratificational, tagmemic and transformational techniques.

In the grammars we have undertaken so far for the several languages, we have tended to exclude such features from morpho-syntactic description.

EXPANDING THE MODEL

We shall include features such as lexical collocation (agent-action agreement) and transformations of semantic equivalence in a systematic description of a higher order which presupposes a morpho-syntactic description for each language [8, pp. 66-71]. The following analogy might be drawn: just as strings of alphabetic and other characters are taken as a body of data to be parsed and classified by a phrase structure grammar, we may regard the string of rule numbers generated from a phrase structure analysis as a string of symbols to be parsed and classified in a still higher order grammar [11; 13, pp. 67-83], for which there is as yet no universally accepted nomenclature. The term transformational strongly suggests itself and is widely used, but the term semantic may seem more appropriate to others.⁴

PROJECTIONS

During the coming year we shall proceed to expand syntactic description of all languages now under investigation. Sufficient transfer data will be compiled between pairs of languages to test the general validity of the model and the general adequacy of the system of programs we are now using. Several questions suggest themselves with respect to limitations of the model, among them: 1. how large will the syntactic description of a language be in terms of rules before the grammar converges on the languages, and 2. in what ways can we improve the quality of translation by using a more sophisticated model, say one in which there is a grammar of structural semantics? We shall be occupied primarily with these two questions in an effort to anticipate the need for modifying elements of the translation programs and in an effort to test empirically with a comprehensive data base some of the more recent theories and notions of linguistics.

FOOTNOTES

1. Research at the Linguistics Research Center is supported by the National Science Foundation, the U. S. Army Electronics Laboratories, the U. S. Air Force and the Latin American Institute of The University of Texas.
2. Recognition is due the entire LRC staff, present and past, for success in the results reported here. Among the linguists who contributed more immediately to the underlying data are: T. Baker, T. Git, M. Prince, K. Ryan, R. Stachowitz, A. Staves, C. Swinburn. Intensive preparation of test data for the demonstration covered the period from August, 1964 to January, 1965. General research and development of programs have been under way since May, 1959.
3. On comparing the computer and human versions of the English translation with the German version, the reader is reminded that nowhere are any corpus data stored explicitly in the translation system of programs. Only raw corpus data in the source language are fed in as input to the analysis programs in the system. The analysis and synthesis programs use grammatical descriptions in both languages with attendant transformation/translation rules to produce output in the target language from the analysis-transfer-synthesis cycle.
4. Perhaps a passing observation is in order. The term transformational, borrowed from mathematics, is a term generally applicable to any process of mapping equivalences of one structure onto another and so is applicable to all levels of linguistic description.

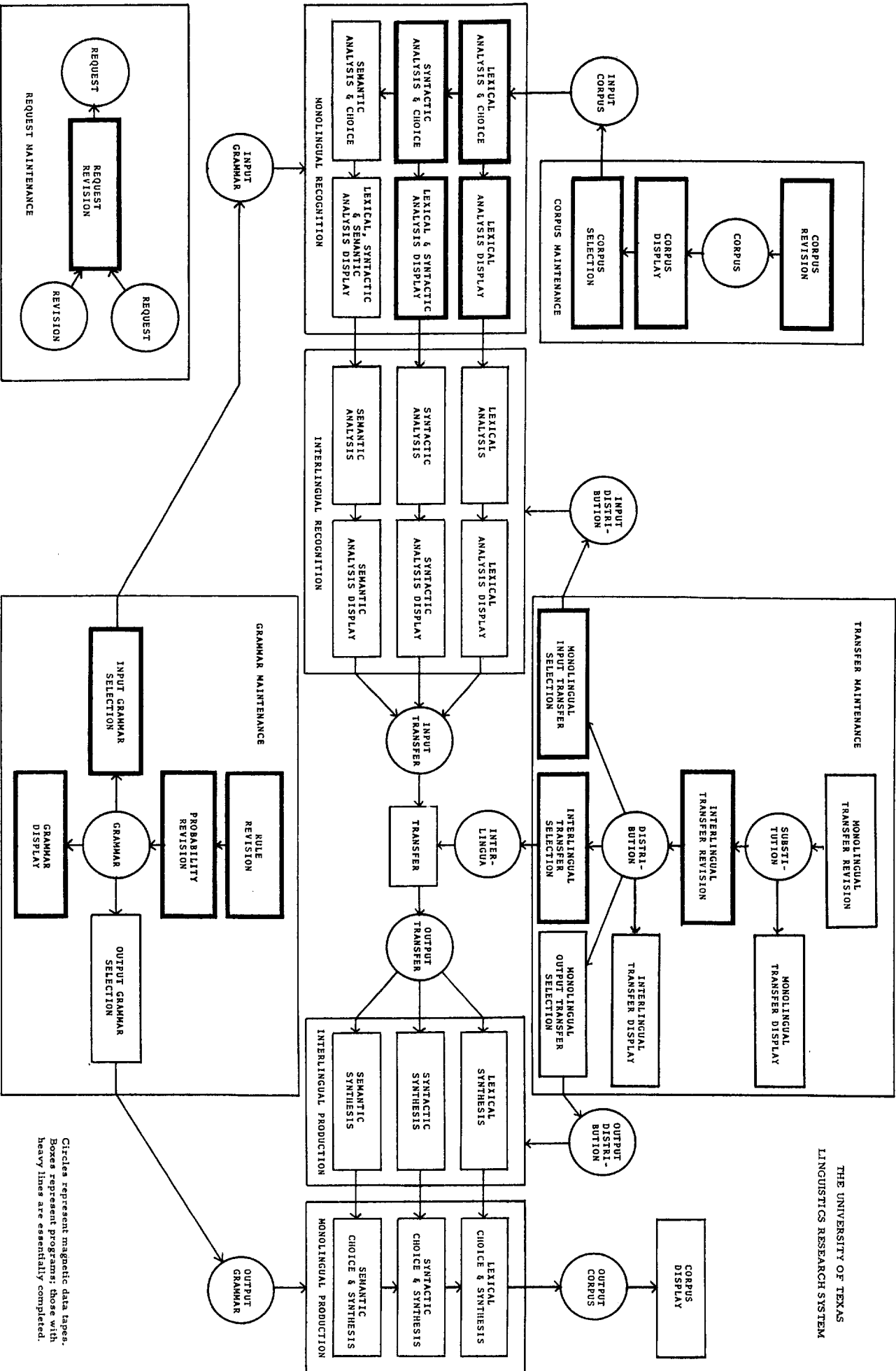
FOOTNOTES (CONTINUED)

It should not, therefore, be used to denote a particular level in a hierarchical structure. The term semantic, on the other hand, may perhaps come to be universally accepted as a hierarchical expression in some series like:

pragmatic
logical
semantic
syntactic
morphological
phono-/graphological

REFERENCES

- [1] A. G. Dale, N. Dale, E. D. Pendergraft, "A Programming System for Automatic Classification with Applications in Linguistic and Information Retrieval Research," presented to the International Study Conference on Classification Research, Elsinore, 1964 (mimeo)
- [2] R. Jernigan and A. G. Dale, "Set Theoretic Models for Classification and Retrieval," LRC 64 WTM-5 (Austin, 1964)
- [3] I. Langerhans, "Transformational Translation: English to German," Transformations and Discourse Analysis Projects 50 (Philadelphia, 1963)
- [4] W. P. Lehmann, "Computational Linguistics: Procedures and Problems," LRC 65 WA-1 (Austin, 1965)
- [5] W. P. Lehmann, "Towards Machine Translation," LRC 64-WA2 (Austin, 1964)
- [6] W. P. Lehmann and E. D. Pendergraft, "Structural Models for Linguistic Automation," VISTAS IN INFORMATION HANDLING, P. W. Howerton, ed. (Washington, 1963), pp. 78-91.
- [7] Linguistics Research Center, "Seventh Quarterly Progress Report," LRC 65 AMC-23 (Austin, 1965)
- [8] Linguistics Research Center, "Symposium on the Current Status of Research," LRC 63-SR1 (Austin, 1963)
- [9] E. D. Pendergraft, "A Generalized Computer System for Language Translation," LRC 64-WA1 (Austin, 1964)
- [10] H. Schaefer, "Ueber die physiologische Grundbedingung des Bewusstseins," UNIVERSITAS, Oktober 1959, 10.1079-1090.
- [11] W. Tosh, "Content Recognition and the Production of Synonymous Expressions," PROCEEDINGS OF THE NINTH INTERNATIONAL CONGRESS OF LINGUISTS, (The Hague, 1964), pp. 723-729.
- [12] W. Tosh, "Development of Automatic Grammars," to appear in LINGUISTICS.
- [13] W. Tosh, SYNTACTIC TRANSLATION, (The Hague, 1965)
- [14] W. Tosh, "System Requests Programming Reference Manual," LRC 65 TP-1 (Austin, 1965)



Circles represent magnetic data tapes.
Boxes represent programs; those with heavy lines are essentially completed.

GERMAN INPUT TEXT
CORPUS DISPLAY

74000001 74
74000002 UEBER DIE PHYSIOLOGISCHE GRUNDBEDINGUNG DES BEWUSSTSEINS
74000003 (AUFSATZ VON PROF. DR. HANS SCHAEFER, UNIVERSITAET HEIDELBERG,
74000004 IN/1/ /5/UNIVERSITAS/5/, OKTOBER 1959, 14. JAHRGANG, HEFT 10,
74000005 SEITE 1079-1090)
74000006
74001001 DIE KOERPERLICHEN BEDINGUNGEN, UNTER DENEN ALLEIN BEWUSSTSEIN
74001002 MOEGLICH IST, SIND RECHT MANNIGFALTIG, DAS PROBLEM DER KOPPLUNG
74001003 VON PSYCHISCHEM AN DIE STRUKTUR UNSERES GEHIRNS SO VERZWEIGT, DASS
74001004 IN EINEM AUFSATZ NUR EIN TEILPROBLEM HERAUSGEGRIFFEN WERDEN KANN.
74001005 WAS HIER BEHANDELT WERDEN SOLL, STELLT DAS (WIE ICH GLAUBE)
74001006 WESENTLICHSTE PROBLEM EINER KOERPERLICHEN BEDINGTHEIT
74001007 SEELISCHER VORGAENGE DAR.
74001008
74002001 I
74002002 DER ZUSTAND, DEN DAS GEHIRN DES LESERS IN DIESEM AUGENBLICK
74002003 AUFWEIST, WO ER SICH ENTSCLOSSEN HAT, EIN SO KOMPLIZIERTES THEMA
74002004 MIT DEM VERFASSER GEMEINSAM ZU BETRACHTEN, IST DER EINER WACHEN
74002005 AUFMERKSAMKEIT. IN IHM ... D.H. IN DEMJENIGEN TEIL SEINER PERSON,
74002006 DEN ER SEIN /5/ICH/5/ NENNT UND DER SEINER SELBSTBEOBACHTUNG IN DIES
74002007 EM MOMENT OFFENLIEGT, FINDET ER JETZT EINE REIHE VON UEBERLEGUNGEN
74002008 VOR, DIE TEILS MIT DEM GLEICH SIND, WAS DER VERFASSER ZUR ZEIT DER
74002009 ABFASSUNG DIESES AUFSATZES AUCH UEBERLEGTE. TEILS WEICHEN SEINE
74002010 GEDANKEN VON DENEN DES VERFASSERS EIN WENIG AB, WAS ALLEIN DADURCH
74002011 VERSTAENDLICH IST, DASS DER VERFASSER DIESE GEDANKEN PRODUZIERT, IM
74002012 UEBRIGEN AUCH FUER RICHTIG HAELT, DER LESER DAGEGEN DER
74002013 /5/NACHDENKENDE/5/, WEIL EMPFANGENDE IST UND DABEI HOFFENTLICH
74002014 NICHT GANZ DEN ZWANG LOSWIRD, BEIM NACHDENKEN DAS, WAS IHM GESAGT
74002015 WIRD, AUF SEINE /5/RICHTIGKEIT/5/ ZU UEBERPRUEFEN.
74002016
74003001 ALLES DAS ABER LAEUFT IM LESER ALS /5/BEWUSSTSEIN/5/ AB, ALSO
74003002 DORT, WO /5/ER SELBST/5/ ZU HAUSE IST. WAS BEWUSSTSEIN IST, KANN
74003003 MAN NICHT NAEHER UMSCHREIBEN. ES GIBT KEINE BESCHREIBUNGSMITTEL
74003004 FUER ETWAS, DAS SELBER EINER JEDEN BESCHREIBUNG ALLER DINGE
74003005 VORAUSGEHT. ALLES, WAS WIR BESCHREIBEN, SIND VORGAENGE, DIE IHRE
74003006 SPUR VORHER IN UNSER BEWUSSTSEIN EINGEGRABEN HABEN.
74003007
74004001 WENN WIR EINEN AUGENBLICK UNSERE AUFMERKSAMKEIT IM ZIMMER
74004002 UMHERWANDERN LASSEN, IN DEM WIR SITZEN/1/ VIELLEICHT HOEREN WIR
74004003 JETZT EINE UHR TICKEN, EIN GLOCKENTON MAG VON AUSSEN AN UNSER OHR
74004004 DRINGEN, ODER EIN KIND PLAPPERT VOR SICH HIN ... WOVON WIR VORHER
74004005 NICHTS WAHRGENOMMEN HABEN. WENN WIR AUFMERKSAME LESER SIND,
74004006 VERGESSEN WIR ALLES UM UNS HERUM, VIELLEICHT NICHT IMMER BEI EINEM
74004007 WISSENSCHAFTLICHEN TEXT WIE DIESEM, BEI DEM SO VIEL KONZENTRATION
74004008 ZU VIEL VERLANGT WAERE. WER ABER KENNT NICHT DEN LESER DES
74004009 KRIMINALROMANS, DER IN SICH VERSUNKEN DIE WELT
74004010 VERGISST ... Sogar das Donnern der Untergrundbahn, die er benutzen
74004011 will und die nun dem erschreckt auffahrenden bereits
74004012 davongefahren ist.
74004013
74005001 DIESE KURZE GEMEINSAME UEBERLEGUNG IST EINE ART EXPERIMENT MIT
74005002 UNS SELBST GEWESEN, UM DREI BEGRIFFE ZU KLAEREN/1/ BEWUSSTSEIN,

CORPUS DISPLAY

74005003 ALSO DAS, WAS WIR IN UNS UNMITTELBAR VORFINDEN/2/ AUFMERKSAMKEIT
74005004 ALS EIN WORT FUER EINE UNS ZUNAECHST UERKLAERLICHE KRAFT, DIE
74005005 UNSER BEWUSSTSEIN VON DEN MEISTEN GEGENSTAENDEN UNSERER UMWELT
74005006 WEGZIEHT UND EINEM EINZIGEN VORGANG ZUWENDET/2/ ENDLICH DINGE, DIE
74005007 ZWAR UNSERE SINNESORGANE TREFFEN (GERAEUSCHE Z.B.), VON DIESEN AUCH
74005008 MELDUNGEN IN DAS GEHIRN SCHICKEN, WIE WIR SICHER WISSEN, DOCH IN
74005009 UNSEREM GEHIRN NICHT IN DAS BEWUSSTSEIN DRINGEN, ALSO UNBEWUSST
74005010 VERBLEIBEN. SIE ENTGEHEN UNSERER AUFMERKSAMKEIT, HINTERLASSEN ABER
74005011 DOCH IHRE SPUREN, DENN NACHTRAEGLICH NACH DEM BEFRAGT, WAS WAEHREND
74005012 DER LEKTUERE DES KRIMINALROMANS UM UNSEREN VERTIEFTEN LESER VOR
74005013 SICH GING, WIRD ER SICH AN MANCHES ERINNERN, WENN AUCH UNDEUTLICH.
74005014 IN EINER HYPNOSE LASSEN SICH SOLCHE ERINNERUNGSSPUREN UNTER
74005015 UMSTAENDEN NOCH WEITER ERHELLEN UND INS LICHT DES BEWUSSTSEINS
74005016 HEBEN.
74005017
74006001 BEWUSSTSEIN IST ALSO ... VON INNEN GEGEHEN ... ETWAS, DAS AN
74006002 EINEN STROM VON ERREGUNGEN GEBUNDEN, AUS SINNESORGANEN UEBER NERVEN
74006003 IN ZENTRALNERVOESE STRUKTUREN EILEND, HIE UND DA AUFBLITZT, VON
74006004 EINEM TEILE DIESES STROMES BESITZ ERGREIFT UND JE NACH DER RICHTUNG
74006005 DER AUFMERKSAMKEIT BALD HIER BALD DORT ETWAS /5/WAHRNIMMT/5/.
74006006 /5/WAHRNEHMEN/5/ HAT MIT /5/NEHMEN/5/ ZU TUN UND DRUECKT EINEN
74006007 AKTIVEN ANTEIL UNSERES ICH AN DER AUSWAHL AUS DER SUMME MOEGLICHER
74006008 ERFAHRUNGEN AUS.
74006009

APPENDIX B
ENGLISH CORPUS DISPLAY

TOSH A-2

CORPUS DISPLAY

HUMAN TRANSLATION

74001001 THE ONLY BODILY CONDITIONS UNDER WHICH CONSCIOUSNESS IS POSSIBLE
74001002 ARE QUITE DIVERSE AND THE PROBLEM OF CONNECTING THE PSYCHIC WITH
74001003 THE STRUCTURE OF OUR BRAIN IS SO COMPLEX THAT IN AN ESSAY ONE CAN
74001004 ONLY SELECT A PARTIAL PROBLEM. THE SUBJECT TO BE CONSIDERED HERE
74001005 REPRESENTS (IN MY OPINION) THE MOST ESSENTIAL PROBLEM OF A
74001006 DEPENDENCE OF MENTAL PROCESSES ON THE BODY.
74001007

74002001 I.
74002002 THE CONDITION OF THE READER/6/S BRAIN AT THIS MOMENT WHEN HE HAS
74002003 DECIDED TO CONSIDER WITH THE AUTHOR SUCH A COMPLICATED SUBJECT IS
74002004 THAT OF WAKEFUL ATTENTIVENESS. IN IT, I.E. IN THAT PART OF HIS
74002005 PERSON WHICH HE CALLS HIS /5/EGO/5/ AND WHICH AT THIS MOMENT IS
74002006 OPEN TO HIS SELF-OBSERVATION, HE NOW DISCOVERS A SERIES OF
74002007 REFLECTIONS, WHICH ARE PARTLY IDENTICAL WITH THE AUTHOR/6/S
74002008 REFLECTIONS AT THE TIME THIS ESSAY WAS WRITTEN. PARTLY, HIS
74002009 THOUGHTS DIFFER A LITTLE FROM THE AUTHOR/6/S, WHICH IS
74002010 UNDERSTANDABLE MERELY THROUGH THE FACT THAT THE AUTHOR PRODUCED
74002011 THESE THOUGHTS, AND FURTHERMORE CONSIDERS THEM CORRECT, WHILE THE
74002012 READER IS THE RECEIVING PARTY AND THEREFORE THE /5/MEDITATOR,/5/
74002013 AND, HOPEFULLY, DOES NOT IN THE PROCESS LOSE THE COMPULSION TO
74002014 EXAMINE WHAT HE IS BEING TOLD AS TO ITS CORRECTNESS.
74002015

74003001 ALL THIS, HOWEVER, PROCEEDS IN THE READER AS /5/CONSCIOUSNESS,/5/
74003002 I.E. IN THAT AREA WHERE /5/HE HIMSELF/5/ IS AT HOME. WHAT
74003003 CONSCIOUSNESS IS, ONE CANNOT FURTHER CIRCUMSCRIBE.] THERE IS NO
74003004 MEANS OF DESCRIPTION FOR SOMETHING WHICH ITSELF PRECEDES ANY
74003005 DESCRIPTION OF ALL THINGS. EVERYTHING WE DESCRIBE CONSISTS OF
74003006 PROCESSES WHICH HAVE FIRST ENGRAVED THEIR TRACES IN OUR
74003007 CONSCIOUSNESS.
74003008

74004001 IF WE LET OUR ATTENTION ROAM ABOUT FOR A MOMENT IN THE ROOM IN
74004002 WHICH WE ARE SITTING/1/ MAYBE WE NOW HEAR THE TICKING OF A CLOCK,
74004003 THE PEAL OF A BELL MAY REACH OUR EARS FROM OUTSIDE, OR A CHILD
74004004 BABBLES TO HIMSELF ... NOTHING OF WHICH WE PERCEIVED EARLIER. IF WE
74004005 ARE ATTENTIVE READERS, WE WILL FORGET EVERYTHING AROUND US, MAYBE
74004006 NOT ALWAYS WITH A SCIENTIFIC TEXT LIKE THIS ONE, WHERE SUCH
74004007 CONCENTRATION WOULD BE TOO MUCH TO EXPECT. BUT WHO DOES NOT KNOW
74004008 THE READER OF A DETECTIVE STORY WHO, LOST IN HIMSELF, FORGETS THE
74004009 WORLD ... EVEN THE THUNDER OF THE SUBWAY WHICH HE WANTED TO TAKE
74004010 AND WHICH NOW THE STARTLED READER, JUMPING UP, HAS ALREADY MISSED.
74004011

74005001 THIS SHORT JOINT REFLECTION HAS BEEN A KIND OF EXPERIMENT WITH
74005002 OURSELVES IN ORDER TO CLARIFY THREE CONCEPTS/1/ CONSCIOUSNESS,
74005003 I.E. THAT WHICH WE FIND DIRECTLY IN OURSELVES/2/ ATTENTIVENESS AS
74005004 A TERM FOR A FORCE WHICH IS AT FIRST INEXPLICABLE, WHICH DRAWS
74005005 AWAY OUR CONSCIOUSNESS FROM MOST OBJECTS OF OUR ENVIRONMENT AND
74005006 DIRECTS IT TOWARD A SINGLE PROCESS/2/ FINALLY, THINGS WHICH MEET
74005007 OUR SENSE ORGANS (E.G. NOISES) AND, AS WE DEFINITELY KNOW, SEND

CORPUS DISPLAY

74005008 REPORTS FROM THEM TO OUR BRAIN, BUT DO NOT PENETRATE INTO
74005009 CONSCIOUSNESS WITHIN OUR BRAIN, AND THUS REMAIN UNCONSCIOUS.
74005010 THEY ESCAPE OUR ATTENTION BUT LEAVE THEIR TRACES, FOR IF ASKED
74005011 SUBSEQUENTLY ABOUT THAT WHICH TOOK PLACE AROUND OUR ABSORBED
74005012 READER WHILE HE WAS READING THE DETECTIVE STORY, HE WILL REMEMBER
74005013 SOME THINGS, IF ONLY DIMLY SO. IN CERTAIN CASES SUCH MEMORY TRACES
74005014 MAY BE ILLUMINATED EVEN FURTHER UNDER HYPNOSIS AND MAY BE RAISED
74005015 INTO THE LIGHT OF CONSCIOUSNESS.
74005016

74006001 CONSCIOUSNESS ... SEEN FROM WITHIN ... IS THUS SOMETHING TIED TO A
74006002 STREAM OF STIMULI, WHICH RUSHES FROM OUR SENSES BY WAY OF OUR
74006003 NERVES INTO CENTRAL NERVOUS STRUCTURES, LIGHTS UP HERE AND THERE,
74006004 TAKES POSSESSION OF A PART OF THIS STREAM AND, DEPENDING ON THE
74006005 PARTICULAR DIRECTION OF THE ATTENTIVENESS, PERCEIVES SOMETHING HERE
74006006 AND THERE. /5/TO PERCEIVE/5/ HAS TO DO WITH /5/TO TAKE,/5/ AND
74006007 EXPRESSES AN ACTIVE INTEREST OF OUR EGO IN SELECTING FROM THE SUM
74006008 OF POSSIBLE EXPERIENCES.
74006009

FINAL DISPLAY FIRST LRC MT OUTPUT 13 JAN 65

74000001 74UEBER THE PHYSIOLOGISCHE GRUNDBEDINGUNG DES CONSCIOUSNESS (

74000002 ESSAY VON PROF. DR. HANS SCHAEFER, UNIVERSITAET HEIDELBERG, IN/1/

74000003 /5/ UNIVERSITAS/5/, OKTOBER 1959, 14. JAHRGANG, HEFT 10, SEITE

74000004 1079-1090)

74001001 THE BODYEN BEDINGUNGEN, UNTER DENEN MERELY CONSCIOUSNESS

74001002 POSSIBLE IST, SIND RECHT DIVERSE, THE PROBLEM THE KOPPLUNG VON

74001003 PSYCHICEMAN THE STRUCTUR WE BRAINS SO COMPLEX, DASS IN AN

74001004 ESSAY ONLY A PARTIAL PROBLEM SELECT WERDEN CAN. WAS HERE

74001005 CONSIDERIT WERDEN SOLL, STELLT THE (WIE ICH GLAUBE) ESSENTIALSTE

74001006 PROBLEM EINER BODYEN BEDINGTHEIT MENTALER PROCESSE DAR.

74002001 I THE ZUSTAND, DEN THE BRAIN THE READERS IN THESE MOMENT

74002002 AUFWEIST, WO HE DECIDED HAT, EIN SO COMPLICATED SUBJECT A MIT

74002003 THE AUTHOR GEMEINSAM ZU BETRACHTEN, IST DER EINER WAKEFULEN

74002004 ATTENTION. IN IT ... D.H. IN DEMJENIGEN TEIL ITER PERSON, DEN HE

74002005 IT /5/ EGO/5/ CALLT UND DER ITER SELF-OBSERVATION IN THESE

74002006 MOMENT OFFENLIEGT, FINDET HE NOW EINE SERIES VON UEBERLEGUNGEN

74002007 VOR, DIE PARTLY MIT DEM GLEICH SIND, WAS THE AUTHOR ZUR ZEIT THE

74002008 WRITT THESE ESSAYES AUCH REFLECTIONTE. PARTLY WEICHEN ITE

74002009 THOUGHTN VON DENEN THE AUTHORS A LITTLE AB, WAS MERELY DADURCH

74002010 UNTERSTANDABL IST, DASS THE AUTHOR THESE THOUGHTN PRODUCTE,

74002011 FURTHERMORE FOR CORRECT HAELT, THE READER DAGEGEN THE /5/

74002012 MEDITATORE/5/, WEIL EMPFANGENDE IST UND DABEI HOPEFULLY NICHT

74002013 GANZ THE COMPULSION LOSWIRD, BEIM NACHDENKEN DAS, WAS IT GE

74002014 TOLDT WIRD, AUF ITE /5/ CORRECTNESS/5/ ZU UEBERPRUEFEN.

74003001 ALLES DAS ABER LAEUFT IN THE READER ALS /5/ CONSCIOUSNESS/5/ AB,

74003002 ALSO DORT, WO /5/ HE HIMSELF/5/ AT HOME IST. WAS CONSCIOUSNESS

74003003 IST, CAN ONE NICHT FURTHER CIRCUMSCRIBEN. ES GIBT NO

74003004 BESCHREIBUNGSMITTEL FOR SOMETHING, DAS SELBER EINER JEDEN

74003005 DESCRIPTION ALLER THINGE PRECEDES. ALLES, WAS WE DESCRIBEN, SIND

74003006 PROCESSE, DIE THEIRE TRAC FIRST IN WE CONSCIOUSNESS

74003007 EINGEGRABEN HABEN.

74004001 WENN WE FOR A MOMENT WE ATTENTION IM ROOM ABOUT ROAMN LET,

74004002 IN DEM WE SITTEN/1/ MAYBE HEAREN WE NOW A CLOCK TICKEN, EIN

74004003 GLOCKENTON MAY FROM OUTSIDE AN OUR EARS DRINGEN, ODER A CHILD

74004004 BABBLT TO HIMSELF ... WOVON WE FIRST NOTHING WAHGENOMMEN

74004005 HABEN. WENN WE ATTENTIVEE READER SIND, FORGETEN WE ALLES UM UNS

74004006 HERUM, MAYBE NOT ALWAYS WITH EINEM SCIENTIFICEN TEXT WIE

74004007 THESE, WITH DEM SUCH CONCENTRATION ZU VIEL VERLANGT WAERE. WER

74004008 ABER KNOWT NICHT THE READER THE DETECTIVE STORYS, DER IN SICH

74004009 VERSUNKEN THE WORLD FORGETT ... EVEN THE THUNDER THE

74004010 SUBWAY, DIE HE BENUTZEN WILL UND THE NOW THE STARTLT JUMPING

74004011 UPEN ALREADY DAVONGEFahren IST.

74005001 THESE KURZE GEMEINSAME UEBERLEGUNG IST EINE ART EXPERIMENT MIT

74005002 UNS SELBST GEWESEN, UM DREI BEGRIFFE ZU KLAEREN/1/

74005003 CONSCIOUSNESS, ALSO DAS, WAS WE IN UNS UNMITTELBAR VORFINDEN/2/

74005004 ATTENTION ALS EIN WORT FOR EINE UNS ZUNAECHST UNERKLAERLICHE

74005005 KRAFT, DIE WE CONSCIOUSNESS VON THE MEISTEN GEGENSTAEENDEN

74005006 UNSERER UMWELT WEGZIEHT UND EINEM EINZIGEN VORGANG ZUWENDET/2/

74005007 ENDLICH THINGE, DIE ZWAR WE SINNESORGANE TREFFEN (GERAEUSCHE

74005008 Z.B.), VON DIESEN AUCH MELDUNGEN IN THE BRAIN SCHICKEN, WIE WE

74005009 SICHER WISSEN, DOCH IN UNSEREM BRAIN NICHT IN THE

74005010 CONSCIOUSNESS DRINGEN, ALSO UNBEWUSST VERBLEIBEN. SIE ENTGEHEN

74005011 UNSERER ATTENTION, HINTERLASSEN ABER DOCH THEIRE TRAGEN, DENN

74005012 NACHTRAEGLICH NACH DEM BEFRAGT, WAS WAEHREND THE LEKTUERE THE

74005013 DETECTIVE STORYS UM UNSEREN VERTIEFTEN READER VOR SICH GING, WIRD

74005014 HE SICH AN MANCHES ERINNERN, WENN AUCH UNDEUTLICH. IN EINER

74005015 HYPNOSELASSEN SICH SOLCHE ERINNERUNGSSPUEN UNTER UMSTAEENDEN NOCH

74005016 WEITER ERHELLEN UND INS LICHT THE CONSCIOUSNESSS HEBEN.

FINAL DISPLAY

MACHINE TRANSLATION OUTPUT

74000001 74 THE PHYSIOLOGICAL BASIS OF CONSCIOUSNESS (ESSAY BY PROF.
74000002 DR. HANS SCHAEFER, UNIVERSITY OF HEIDELBERG, IN/1/ UNIVERSITAS,
74000003 OCTOBER 1959, VOLUME 14, NUMBER 10, PAGES 1079 - 1090.)

74001001 THE ONLY BODILY CONDITIONS UNDER WHICH CONSCIOUSNESS IS
74001002 POSSIBLE ARE QUITE DIVERSE AND THE PROBLEM OF CONNECTING THE
74001003 PSYCHIC WITH THE STRUCTURE OF OUR BRAIN IS SO COMPLEX THAT IN AN
74001004 ESSAY ONE CAN ONLY SELECT A PARTIAL PROBLEM. THE SUBJECT TO BE
74001005 CONSIDERED HERE REPRESENTS (IN MY OPINION) THE MOST ESSENTIAL
74001006 PROBLEM OF A PHYSICAL DEPENDENCE OF MENTAL PROCESSES.

74002001 I THE CONDITION OF THE READER/6/S BRAIN AT THIS MOMENT WHEN IT
74002002 HAS DECIDED TO CONSIDER WITH THE AUTHOR SUCH A COMPLICATED
74002003 SUBJECT IS THAT OF WAKEFUL ATTENTIVENESS. IN IT, I.E. IN THAT
74002004 PART OF ITS PERSON WHICH IT CALLS ITS /5/EGO/5/ AND WHICH AT THIS
74002005 MOMENT IS OPEN TO ITS SELF-OBSERVATION, IT NOW DISCOVERS A SERIES
74002006 OF REFLECTIONS, WHICH ARE PARTLY IDENTICAL WITH THE AUTHOR/6/S
74002007 REFLECTIONS AT THE TIME WHEN THIS ESSAY WAS WRITTEN. PARTLY, ITS
74002008 THOUGHTS DIFFER A LITTLE FROM THE AUTHOR/6/S, WHICH IS
74002009 UNDERSTANDABLE MERELY THROUGH THE FACT THAT THE AUTHOR PRODUCED
74002010 THESE THOUGHTS, AND FURTHERMORE CONSIDERS THEM CORRECT, WHILE THE
74002011 READER IS THE RECEIVING PARTY AND THEREFORE THE /5/MEDITATOR,/5/
74002012 AND, HOPEFULLY, DOES NOT IN THE PROCESS LOSE THE COMPULSION TO
74002013 EXAMINE WHAT HE IS BEING TOLD AS TO ITS CORRECTNESS.

74003001 ALL THIS, HOWEVER, PROCEEDS IN THE READER AS
74003002 /5/CONSCIOUSNESS,/5/ I.E. IN THAT AREA WHERE /5/HE HIMSELF/5/ IS
74003003 AT HOME. WHAT CONSCIOUSNESS IS, ONE CANNOT FURTHER
74003004 CIRCUMSCRIBE. THERE IS NO MEANS OF DESCRIPTION FOR SOMETHING
74003005 WHICH ITSELF PRECEDES ANY DESCRIPTION OF ALL THINGS. EVERYTHING
74003006 WE DESCRIBE CONSISTS OF PROCESSES WHICH HAVE FIRST ENGRAVED ITS
74003007 TRACES IN OUR CONSCIOUSNESS.

74004001 IF WE LET OUR ATTENTION ROAM ABOUT FOR A MOMENT IN THE ROOM IN
74004002 WHICH WE ARE SITTING/1/ MAYBE WE NOW HEAR THE TICKING OF A CLOCK,
74004003 THE PEAL OF A BELL MAY REACH OUR EARS FROM OUTSIDE, OR A CHILD
74004004 BABBLES TO HIMSELF ... NOTHING OF WHICH WE PERCEIVED EARLIER. IF
74004005 WE ARE ATTENTIVE READERS, WE WILL FORGET EVERYTHING AROUND US,
74004006 MAYBE NOT ALWAYS WITH A SCIENTIFIC TEXT LIKE THIS ONE, WHERE SUCH
74004007 CONCENTRATION WOULD BE TOO MUCH TO EXPECT. BUT WHO DOES NOT KNOW
74004008 THE READER OF A DETECTIVE STORY WHO, LOST IN HIMSELF, FORGETS THE
74004009 WORLD ... EVEN THE THUNDER OF THE SUBWAY WHICH IT WANTED TO TAKE
74004010 AND WHICH NOW THE STARTLED READER, JUMPING UP, HAS ALREADY
74004011 MISSED.

74005001 THIS SHORT JOINT REFLECTION HAS BEEN A KIND OF EXPERIMENT WITH
74005002 US IN ORDER TO CLARIFY THREE CONCEPTS/1/ CONSCIOUSNESS, I.E. THAT
74005003 WHICH WE FIND DIRECTLY IN US/2/ ATTENTIVENESS AS A TERM FOR A
74005004 FORCE WHICH IS AT FIRST INEXPLICABLE WHICH DRAWS AWAY OUR
74005005 CONSCIOUSNESS FROM MOST OBJECTS OF OUR ENVIRONMENT AND DIRECTS IT
74005006 TOWARD A SINGLE PROCESS/2/ FINALLY, THINGS WHICH MEET OUR SENSE
74005007 ORGANS (E.G. NOISES) AND, AS WE DEFINITELY KNOW, SEND MESSAGES
74005008 FROM THEM TO OUR BRAIN, BUT DO NOT PENETRATE INTO CONSCIOUSNESS
74005009 IN OUR BRAIN, THUS REMAIN UNCONSCIOUS. THEY ESCAPE OUR ATTENTION

74005010 BUT LEAVE ITS TRACES, FOR IF ASKED SUBSEQUENTLY ABOUT THAT WHICH
74005011 TOOK PLACE AROUND OUR ABSORBED READER WHILE HE WAS READING THE
74005012 DETECTIVE STORY, IT WILL REMEMBER SOME THINGS, IF ONLY DIMLY SO.
74005013 IN CERTAIN CASES SUCH MEMORY TRACES MAY BE ILLUMINATED EVEN
74005014 FURTHER UNDER HYPNOSIS AND MAY BE RAISED INTO THE LIGHT OF
74005015 CONSCIOUSNESS.

74006001 CONSCIOUSNESS ... SEEN FROM WITHIN ... IS THUS SOMETHING TIED
74006002 TO A STREAM OF STIMULI, WHICH RUSHES FROM OUR SENSES BY WAY OF
74006003 OUR NERVES IN CENTRAL NERVOUS STRUCTURES, LIGHTS UP HERE AND
74006004 THERE, TAKES POSSESSION OF A PART OF THIS STREAM AND, DEPENDING
74006005 ON THE PARTICULAR DIRECTION OF THE ATTENTIVENESS, PERCEIVES
74006006 SOMETHING HERE AND THERE. /5/TO PERCEIVE/5/ HAS TO DO WITH /5/TO
74006007 TAKE,/5/ AND EXPRESSES AN ACTIVE INTEREST OF OUR EGO IN SELECTING
74006008 FROM THE SUM OF POSSIBLE EXPERIENCES.

APPENDIX E

Rules Used in Figure 1 (German)

SNTNC → CLS + .

CLS → CLS/S-O-R + , + PRDCT/D1/ACSTV

CLS/S-O-R → was + PRDCT/D2/3

PRDCT/D2/3 → NO/NTR/NDA + ist

NO/NTR/NDA → N1OW

N1OW → Bewusstsein

PRDCT/D1/ACSTV → PRDCT/D1/INF + INF/PHRS/ACSTV

PRDCT/D1/INF → MDL/3 + PRN/3

MDL/3 → kann

PRN/3 → man

INF/PHRS/ACSTV → nicht + INF/PHRS/ACSTV

INF/PHRS/ACSTV → ADV + INF/ACSTV

ADV → A1A + er

A1A → naeh

INF/ACSTV → V12A + en

V12A → umschreib

APPENDIX E (Continued)

Rules Used in Figure 2 (English)

SNTNC → CLS + .

CLS → CLS/SBSTNT + , + CLS

CLS/SBSTNT → what + BE/SNGLR/PRSNT

BE/SNGLR/PRSNT → NMNL/A/' + is

NMNL/A/' → NSF

NSF → consciousness

CLS → PRN/SS + VRBL/MDL/PHRS

PRN/SS → one

VRBL/MDL/PHRS → MDL + VRBL/PHRS

PRN/SS → one

VRBL/MDL/PHRS → MDL + VRBL/PHRS

MDL → cannot

VRBL/PHRS → ADVB/A + VRBL

ADVB/A → further

VRBL → VPR1A + e

VPR1A → circumscrib