

ESPRIT Project 2315
April 1989 - March 1992

Translator's Workbench

**Translational Evaluation
of the Termbank**

Universität Heidelberg

M. Albl, K. Kohn, H. Mikasa,

März 1992

Triumph-Adler AG
Fraunhofer Society, IAO
L-Cube
Mercedes-Benz AG
Siemens AG
Siemens CDS
University of Barcelona
University of Heidelberg
University of Stuttgart
University of Surrey

European Strategic Programme for Research and Development
in Information Technology

ESPRIT Project No. 2315: TWB

Translator's Workbench

**Translational Evaluation
of the Termbank**

Workpackage 1.4

"Verifying and explicating terminology"

Universität Heidelberg

Institut für Übersetzen und Dolmetschen

Michaela Abl, Kurt Kohn, Hans Mikasa

March 1992

CONTENTS

1	Objectives	1
2	Pilot Study	3
2.1	Aims and Methods.....	3
2.2	Interview Findings	5
2.2.1	Technical Texts	5
2.2.2	Specialist Lexica.....	7
2.2.3	Extended Termbank Information	8
2.3	Preliminary Conclusions.....	10
3	Method of Evaluation	13
3.1	Thinking Aloud.....	13
3.2	Elicitation Design	16
3.2.1	Evaluation text	16
3.2.2	Prototype Termbank	18
3.2.3	Test Specification.....	21
4	Analysis of Information Categories	23
4.1	Core Termbank.....	23
4.1.1	Meaning Definitions.....	23
4.1.2	Synonyms.....	24
4.1.3	Variants	24
4.1.4	Collocation.....	25
4.1.5	Compounds	26
4.2	Extended Termbank Information	28
4.2.1	Usage Comments.....	28
4.2.2	Transfer Comments.....	29
4.2.3	Encyclopaedic Units.....	31

5	Analysis of Search Strategies	36
5.1	The Translator's Paradox	36
5.1.1	The Discrepancy between Theory and Practice	36
5.1.2	The Time Factor	38
5.1.3	Translating Along the Text Surface	39
5.2	Relevance of Subject Knowledge.....	40
5.3	Search-efficient Representation of Information.....	46
5.3.1	Termbank Design	46
5.3.2	Linking up Extended Termbank Information.....	52
5.3.3	User Orientation.....	56
6	Conclusion.....	58
	References.....	60

1

Objectives

In the following, the Heidelberg Working Group on Translational Termbank Design will report about its investigations concerning the **translational relevance of the TWB-Termbank**.

In keeping with the project's overall interest in a termbank designed to cater for the specific terminological needs of translators, we wanted to know to what extent translators were indeed able to profit from the information provided by the TWB-Termbank. As our concern was with the translational orientation of the Termbank, particular attention was given to those information categories which were considered to be useful in the translation context - i.e. Definition, Synonym, Variant, Collocation - and to those which were created from the very beginning with the translator's more refined needs in mind, i.e. usage comments, transfer comments, and encyclopaedic units.

In our evaluation of the TWB-Termbank, we were not interested in tool evaluation as such, this being the more comprehensive task of the TWB evaluation team (Mercedes-Benz AG). What we wanted to know was whether translators were able to make good use of the terminological information provided, when confronted with textual and/or translational problems of terminology. We were, thus, interested in evaluating the tool only in so far as it was designed to deliver certain **types of terminological information** to the translator. By the same token, it was not with respect to content quality that we wanted to analyse the Termbank. It goes without saying that an incorrect or badly written, say, meaning definition is hardly of any use for the translator. Our interest was clearly centred on information types in connection with their translational relevance.

The evaluation was carried out as a **case study** on the basis of Thinking Aloud Protocols; our subjects were student translators with different grades of terminological and translational experience in the field of catalytic converters. Again, these features of the evaluation design fit in well with the prime objective of evaluating the Termbank with respect to information types: As no comparison between different user groups was intended, a case-study design was certainly the best choice. Because of the complexities of translational behaviour, an

investigation into the question of the translational relevance of terminological information should be based on a sufficiently wide range of interrelated data for each individual subject. This, however, can only be feasibly achieved - within the limits of our project at least - on the basis of a case study. Moreover, such an approach offers the advantage of making use of Thinking Aloud Protocols, i.e. a type of data particularly useful for studying the relevance of terminological information in the context of translation. That student translators were used as subjects in our evaluation study, is in part due to the fact that the overall evaluation of TWB-tools was carried out with practising translators. In addition, the training context made it easier to set the evaluation task on the intended target, i.e. the translational appropriateness of the TWB-Termbank with respect to terminological information types.

2

Pilot Study

2.1 Aims and Methods

In an effort to prepare our user tests, we carried out a pilot study to gain initial insights into Extended Termbank information. We wanted to see in what way and to what extent the Extended Termbank supplemented other information sources (and categories) in supporting users to acquire translationally relevant terminological knowledge and what aspects we should focus upon in the ensuing Termbank evaluation.

The background chosen for our pilot study was a linguistics seminar on *Translation-oriented analysis of terminology*. Participants were called on to write seminar papers on linguistic phenomena in LSP, i.e. synonymy, collocation, word family, hyponymy, the relevance of meaning definitions, meaning specification in technical texts and transfer problems. The special area they were dealing with was 'catalytic converter technology', so that the "subjects" perfectly suited our experimental purposes. Just like translators, who are often confronted with subject-matters that are altogether unfamiliar to them, these students - most of them beginners in the field of catalytic converter technology - had to acquire an understanding of a given technical domain and to analyse terminological problems and gather linguistic and factual information. Our idea was that it would be revealing to observe how they would go about their task, i.e. what use they would make of domain-specific material in order to become acquainted with the new LSP-domain and in order to investigate the terminological aspects and problems. The issue at hand, therefore, was to see what obstacles and problems they would be confronted with and to what extent different information sources might assist in the overcoming of them.

It was to this end that the students were given the following material:

- (1) real-life technical **texts** and information brochures in the field of catalytic converter technology

- (2) specialist **lexica**, in particular the *Dictionary of Automotive Emission Control* by Peter Schmitt (1986)
- (3) Extended **Termbank** information, i.e. encyclopaedic units and transfer comments elaborated within the framework of our work-package (cf. Albl et al. 1991)

We asked them to keep track of how they read their way into their task, what use they made of the material at their disposal, and to what extent it assisted their purpose. The study was carried out at a point when the students had not yet completed their seminar papers but had concluded all the necessary inquiries and were already engaged in the process of writing up their findings.

We interviewed the students in a non-standardized way, bearing in mind, at the same time, our interest in finding out more about the following questions:

1. In what way do technical **texts** (including company-specific brochures) provide factual and terminological information? What types of information can be extracted from the texts (e.g. collocation, synonyms, hyponyms etc.)? What problems arise in LSP-texts, i.e. what information is needed to deal with the texts of a given domain, and what answers can texts provide, i.e. what knowledge do texts convey?
2. To what extent is this information represented in **lexica**? Is the information found in communicative texts made more concise and to-the-point in the lexica? Do lexica supplement texts by giving more comprehensive information? Are they therefore a sufficient source of information for the reception and production of texts?
3. What conclusions could be drawn as to the representation of these items of information in a **termbank**? What (core and extended) information categories are needed? What should be their content and structure? What navigation possibilities should be provided for, so as to guarantee the accessibility of the information? How relevant and helpful is encyclopaedic and translation-specific information?

2.2 Interview Findings

According to the three **sources of information** that had served as the basis of our study, the interviews revealed interesting observations with respect to technical texts, dictionaries and (extended) termbank information.

2.2.1 Technical Texts

In the interviews, the following features were found to be characteristic of technical texts - e.g. books, articles, reports, company brochures and encyclopaedic entries:

- * Texts tend to be heterogeneous, they vary in the usage of terms and often reflect a **lack of precision and standardization**.

In technical texts one finds phrases such as *emission control requirements for nitrogen oxide emissions* or *emission control requirements of 1 gram per mile for NOx* (Taylor 1984, 1). Although the explicit mentioning of the pollutants clearly indicates that *standards* are referred to here, the term *requirements* is actually used. Another very irritating feature was the fact that for the indication of the stoichiometric mixture, texts rather arbitrarily used the figures 14.6, 14.7, or 14.8.

- * Texts contain differing definitions and **contradictory information**.

Irritating differences in meaning definitions are particularly striking in the case of *dual bed/dual converter*. Bosch's "Kraftfahrtechnisches Taschenbuch" (Bosch GmbH 1987, 445) uses the expressions *Doppelbett-Katalysator* and *Zweibett-Katalysator* synonymously for two catalysts in one bed: "Die Zusammenfassung von Reduktions- und Oxidationskatalysator in einem Gehäuse oder 'Bett' bezeichnet man als 'Doppelbett-' oder 'Zweibett-Katalysator'". Bosch's "Technische Unterrichtung" (Bosch GmbH 1985, 26), however, depicts the *Doppelbettkatalysator* as a system with two beds as opposed to the oxidation-catalyst and the three-way-catalyst with only one bed.

- * Texts are full of factual as well as terminological **unclarities and difficulties**.

Non-catalyst technology is a recurrent phrase not found in any dictionary; it is far from clear what measures it comprises. *Net oxidizing conditions* is another expression which cannot be looked up, as there is no conventional equivalent. What is more, this expression does not give a clue to what it might mean, i.e. it is not motivated as is non-catalyst technology. It is therefore extremely difficult for users to come across an appropriate suggested translation, e.g. *bei Luftüberschuß* (as proposed by the department's translation expert in the field).

- * It is especially in the **translation** process that these **irregularities** become most strikingly apparent; English texts in particular are vague and non-exact, while the German versions often require rather more specification or elaboration of the facts and expressions in the source text. Unless comprehensive and detailed **background knowledge** is available or provided, an adequate translation decision cannot be taken.

The engine-out exhaust emissions ... vary as a function of air-fuel ratio as well as several other parameters such as ignition timing or EGR (Taylor 1984, 3). In this case, EGR or exhaust gas recirculation does not refer to the system or concept as such, but to the volume of recirculated exhaust, so that the translation must be: ... Parameter wie eingestellter Zündzeitpunkt sowie rückgeführte Abgasmenge. Such a solution would never occur to a lay user, nor would it be found in a conventional source of information.

The observed vicissitudes in technical communication and the stumbling-blocks they constitute for a reader/translator are often due to **author-, company- and country-specific variations**:

- (a) Experts naturally differ in their views, insights and perspectives as is demonstrated in their utterances.

In phone calls with experts our interviewees observed that technical communication is a rather subjective matter. They found that there was wild disagreement amongst the experts from Opel, Mercedes Benz and Ford on whether to use *g/km* or *g/m* as a measurement for emissions (gram per mile). Funnily enough, older experts tended to prefer the American version *g/m*, while the younger ones rather used *g/km*; the latter use would, of course, involve an additional effort on the part of the translator, who would then have to convert all the figures given in a text.

- (b) Apart from company-specific peculiarities (e.g. *Thermactor secondary air control* (Ford), *K-Jetronic injection system* (Bosch)), inconsistencies arise from differences in the various countries not only with respect to terminology, but to technology itself.

In American texts the terms *dual-bed converters* and *dual converters* denote three-way catalysts followed by an oxidation catalyst: "(Some manufacturers employ dual converters. The distinction between dual bed and dual converters is whether the two catalysts are housed in the same container or in separate containers.) Both contain a three-way catalyst..." (Taylor 1984, 1). In Germany such systems are not produced at all. German vehicles are equipped either with a single bed oxidation catalyst, or else with a single bed three-way catalyst, or with a dual bed catalyst containing an oxidation and a reduction catalyst in one container with two beds. Three-way catalysts never occur in combination with an oxidation catalyst as is typical in the United States. It is therefore very irritating and a major transfer problem to have the

similar terms *dual-bed converter* (and *dual converter*) as well as *Doppelbett-/Zweibett-Katalysator* not referring to the same types of catalysts.

Technical texts turn out to be precise, reliable and sufficiently conclusive only in places. Contrary to what is often assumed, such features cannot be considered as general characteristics of LSP texts, since the quality of a text always depends on the circumstances under which it is produced.

2.2.2 Specialist Lexica

The interviewees agreed that lexica are not capable of coping with the multitude of evolving and changing variations and peculiarities typical of technical communication. Even the **Schmitt Dictionary**, which gives a much more comprehensive view of the domain in question than do conventional specialist lexica and was regarded as being rather helpful is also reported to be insufficient. The following reasons were given:

- * **lexica** were often found to lack major points of difficulty and specific idiosyncracies; they do not go into sufficient depth, presenting, as a consequence, misleading or even incorrect information.

The Schmitt dictionary, for example, was found to lack phrases such as *net oxidizing conditions*, which have no obvious German equivalent as well as more general phrases such as *non-catalyst technology*, *catalyst systems*, *catalytic control*, or culture-specific differences such as the dual bed/dual converter vs. Doppel-/Zweibettkatalysator issue (illustrated in chap. 2.2.1). If compounds are treated, only the basic meaning is given, e.g. *control systems* meaning *Steuer- und Regelsystem*. The dictionary fails to mention that control system can be a clipping of emission control system [Abgasreinigungssystem] (Schmitt 1986, 16).

- * **definitions**, in particular, were felt to be of little use in covering all the aspects necessary for the comprehension of technical expressions; they are rather loosely knit, omitting a host of relevant points. They certainly do not provide the detailed in-depth knowledge needed not only for the understanding of individual terms but also for the analysis of the complex expressions and relations that make up technical communication; the quasi-definitions or explanations integrated in texts were felt to be often far more revealing.

While in the Schmitt dictionary *dual converter* is wrongly presented as a synonym for *dual bed converter*, with the same definition given for both terms, a much more adequate description can be found in the text by Taylor (1984, 5/6): "The catalytic converters used on most 1981 model year vehicles fall into two general categories: three-way converters and dual-bed converters. (Some manufacturers employ dual

converters. The distinction between dual bed and dual converters is whether the two catalysts are housed in the same container or in separate containers.) Both contain a three-way catalyst, but with the dual bed converter (and dual converter) the three-way catalyst is followed by an oxidation catalyst to provide increased oxidation capability...". The fact that such three-way and oxidation catalyst combinations do not exist in Germany, and that the corresponding terms *Doppelbettkatalysator/ Zweibettkatalysator* refer to a different system, is not explained by any definition.

- * **synonyms and equivalents** as given in conventional lexica fail to be fully instructive, since they are listed in isolation and without any additional comments on the conditions under which they apply and on how they should be used. The Schmitt dictionary gives some but certainly not sufficient additional information and explanations.

Schmitt (1986, 21) wrongly lists as synonyms *Gemischbildung* and *Gemischaufbereitung* and fails to provide a definition or a comment on the difference between the two. (Bosch GmbH 1987, 439), on the other hand, presents *Gemischaufbereitung* as a wider term comprising *Gemischbildung* as well the resulting quality of the mixture: "Gemischaufbereitung umfaßt neben der Einstellung auch die Qualität, mit der das Gemisch den Brennraum erreicht. Homogenität bzw. Schichtung des Kraftstoffs zum Zeitpunkt der Entflammung, sowie die Gemischtemperatur haben einen wesentlichen Einfluß auf die Entflammbarkeit, den Verbrennungsablauf und damit auf die Abgaszusammensetzung."

- * **collocations** do not, in most cases, even appear. Again the Schmitt dictionary has a much better record than conventional lexica, but it should be elaborated upon in the Translator's Workbench/Termbank.

Although the collocations in connection with *stoichiometric* (e.g. *to maintain stoichiometric operation, stoichiometric exhaust gas, close to the stoichiometric composition, at a stoichiometric mean A/F, to cross the stoichiometric point* etc.) are instrumental for catalytic converter technology and are very difficult to translate given the German expression *lambda*, which is often used instead, the only collocation found in Schmitt is that of a *stoichiometric A/F mixture* (Schmitt 1986, 56).

2.2.3 Extended Termbank Information

The Extended Termbank was designed to account for the plethora of domain-related, translation-specific and other terminological problems in which in all technical texts abound but which are, in many cases, not catered for in lexica and dictionaries. Complementing the traditional information categories, additional information is provided in the form of **extended**

usage (for the adequate understanding and use of terms), **transfer comments** (for specific translation-related bilingual problems) and **encyclopaedic units** (for the necessary domain-specific background knowledge).

We had handed out encyclopaedic units and transfer comments to the interviewees in order to find out what use would be made of Extended Termbank information as a means of introduction to the domain and, more specifically, as an instrument in the process of their linguistic analyses.

Encyclopaedic units and transfer comments were found to be, on the whole, extremely helpful. With regard to **encyclopaedic units**, the following observations were made:

- * It was felt strongly that they saved the time-consuming analysis of comparative texts and the otherwise inevitable innumerable phone calls to experts in the field. Other than conventional information sources, they were the only sources to help overcome the hurdles presented by the technical details of texts (such as the implications of a stoichiometric mixture, for example, or the facts related to control mechanisms) and to bridge gaps in terminological and factual knowledge and integrate the two.

Non-catalytic technology, a term occurring in many texts, is not normally found in dictionaries or termbanks. The unit ABGASREINIGUNGSVERFAHREN was able to provide the background knowledge necessary to the translator, e.g. the fact that there are engine-internal and engine-external measures to control emissions, whereby the latter comprise thermal and catalytic afterburning; *non-catalytic-technology* therefore refers to engine-related measures plus thermal afterburning.

- * At the same time, the encyclopaedic units were said not to suffice as a means of becoming fully acquainted with the the factual information related to the domain in question. An introduction to a domain seems to require certain redundancies. In order to be able to grasp the meaning thoroughly, the uninitiated reader, at least, needs time to get into the unfamiliar subject-matter and prefers to read things in a paraphrased and recurrent rather than a condensed form. However, this is neither the intended nor the anticipated purpose of the Encyclopaedia; rather, encyclopaedic units were felt to be particularly useful when allowing the user to look up specific problems.
- * Encyclopaedic units present collocations and other types of terminological information in their natural environment. They thus provide some of the terminological information needed for target text production. As units are written in a concise and to-the-point way, any aspect of a given domain is always expressed only once. Thus the terminological information implicitly provided is highly selective. Users, however, like to have several

idiomatic ways of putting the same thing, i.e. a variety of linguistic forms and expressions. Units should therefore be seen in interaction with the terminological information provided by the Termbank proper.

With respect to **transfer comments** it was pointed out they indeed helped understand and solve the really tricky problems arising from the transfer process. It was found, however, that their number should be increased to cover a larger and more differentiated variety of translationally relevant problems should be covered.

Stoichiometry is one of the central terms in catalytic converter technology. However, the German noun *Stöchiometrie* only exists in the chemical domain, so that with respect to catalytic converters expressions such as *der stöchiometrische Punkt* or $\lambda = 1$ have to be used instead. Such an important comment on the translation of the term does not occur in lexica; in the Schmitt Dictionary (1986, 56) the noun itself is not treated at all; the user will find only the adjective *stoichiometric*.

At the same time, transfer comments were sometimes found to be difficult to understand, as they are treating very specific and intricate problems. For this reason, they require a careful analysis of the problem referred to and a well-structured presentation of the relevant information.

2.3 Preliminary Conclusions

The following conclusions can be drawn from the interviews and serve as a basis for the evaluation of the Termbank's contents:

- * The **Encyclopaedia** is not a *comprehensive text book* for the initiation in a domain, nor, as was originally proposed, should it be considered as a *rough outline of the domain*. Rather, it must be regarded as a **supplement to conventional terminological information categories** clarifying particular points of difficulty by providing the necessary background information.
- * It is therefore very important that the Encyclopaedia be thoroughly **interlinked with the other information categories**, in particular, meaning definitions, TCs and collocations. **Flexible navigation** possibilities alone can provide a tool offering a wide range of suggestions and solutions.

- * The degree of variation, redundancy and detail in description depends on the target group, i.e. the users' background knowledge. The Encyclopaedia is therefore not an absolute entity which is to be written in such and such a way and not in any other. Rather, it is **flexible in its contents and scope**, i.e. depending on the user group it should be more or less specific and more or less comprehensive. Given this insight, **guidelines** should be drawn up on how to create, elaborate, link up and navigate encyclopaedic information. The actual contents filling the architecture should be left to the individual end-user group.
- * The repeated call for the units to include or deal with all the problematic and tricky points of sublanguages, such as in-company-specific peculiarities, which are not translation-, i.e. interlanguage-specific and are therefore not taken up by the transfer comments, was due to lack of extended usage information. This clearly demonstrates that a **comprehensive usage category** must be part and parcel of an Extended Termbank aiming at solving real-life translating problems.
- * To help make the **transfer comments** a truly efficient tool, it might have to be checked in terms of content or structure as to how their intricacy could be reduced.
- * Having interviewed the students in the process of completing their seminar papers, the striking thing we found was that within a very short time (4-8 weeks) of working with the given material (as described in chap. 2.1), their attention had been drawn to profound terminological complexities, which we - in writing the encyclopaedic units and transfer comments - had taken much longer to realize, analyse and understand. The thorough **in-depth information** of the Extended Termbank information had paved the way right from the beginning to an understanding of the subject matter which could not have been gained with the same degree of efficiency from reading and comparing texts or from consulting conventional sources. A rather **positive picture** of the encyclopaedic units and the transfer comments had emerged, which needed to be examined by the tests.
- * On the other hand, the students had come across some very complicated and controversial points, which they were able to clarify only - apart from certain cases where no straightforward answer could be given - by phoning up experts. The value of expert opinion, however, also has to be judged with caution: Concerning the controversy around *dual* and *dual bed converter* mentioned above, German engineers at Opel, Mercedes Benz and Eberspächer, for instance, were fully unfamiliar with the definition given in Taylor (1984). In the case of *net oxidizing conditions* none of the

experts available was of any help; it was the translation department's expert dealing with technical translations in the field of catalytic converter technology who came up with the right solution, i.e. the suggested translation phrase *bei Luftüberschuß*.

- * It seemed that there were indeed **complexities** and intricate problems in the terminology and in the technical texts of a given domain that needed to be taken up and discussed by **Extended Termbank comments**.

3

Method of Evaluation

3.1 Thinking Aloud

Introspection

Since the focus of interest in translation shifted from translation as a product to translation as a process, the cognitive strategies involved in translating have been approached by means of **introspective methods**. Introspection was first employed in psychology to investigate human cognition. The classical type of introspection is self-observation, which involves the verbalization of thoughts either while an action is being carried out or else upon its completion (concurrent verbalization versus retrospective probing). The verbal data can then be subjected to analysis.

This of course begs the question as to how far our cognitive experience can be formulated in words. A distinction must here be made between our automatized cognitive processes, which use information already processed by the mind, and our non-automatized cognitive processes, which use information just entering consciousness. Although the former kind appear not to admit of verbalization, the latter, being present in consciousness, may be expressed in words. It is these latter processes that are the subject of the present investigation.

In a seminal work in this field, Krings (1986) has demonstrated **thinking aloud** to be the most important of introspective methods which may be brought to bear upon the process of translation.

In thinking aloud, subjects simply say aloud everything that is going through their minds while they are engaged upon a task of some kind. The verbal data thus produced are recorded and usually transcribed in the form of a Think-Aloud Protocol (TAP). A concurrent approach, thinking aloud was developed at the beginning of the 20th century by Bühler and first employed to investigate cognition by Clapadère and Duncker in the 'thirties. Although rejected by behaviorists on account of its lack of intersubjective testability, thinking aloud provides data to which we would otherwise have no access, and which allow us to draw conclusions as to the way the human mind works.

The validity of the verbal data obtained varies in accordance with the following criteria (cf. Krings 1986, 63):

- * **The lapse of time** between the performance of the allotted task and the verbalization of the thoughts it gives rise to. Verbal data obtained by retrospective probing will deviate more from what actually occurred in the mind than data obtained by concurrent verbalization because they inevitably undergo a process of reshaping and contextualization.
- * **The level of structure** in the data. Subjects thinking aloud let their thoughts flow without abstraction or selection. The resulting data are (relatively) complete, unstructured and not distorted by the recall of automatized data.
- * **The subject of the verbalization.** The results will vary according to the amount of non-automatized processes involved.
- * **The form of information** to be verbalized. Best results are attained with linguistic data.

Advantages

On the basis of the criteria listed above, we can conclude that thinking aloud is an effective method of examining the translation process, and this for the following reasons:

- * Translating is in itself a linguistic action.
- * A phenomenon called "inner speech", i.e. the formulation of phrases which never come to be uttered aloud, is frequently observed in the process of translating. Verbalization can thus almost be considered a natural action in this context.
- * Translation is mainly non-automatized, many of the mental processes to which it gives rise can therefore be verbalized.
- * Automatized structures derived from general knowledge are of course used in the process of translation. Bottom-up and top-down considerations interact permanently. They serve only, however, to solve the translation task and do not affect the task of verbalization. Another advantage is incorporated here. TAPs help us to discover the hidden cognitive strategies involved in translating (e.g. inferencing, the use of auxiliary tools.)

A brief comparison will now be made of the introspective method of thinking aloud and non-introspective approaches whose aim is also to investigate the process of translating. As Kohn (1988, 44) emphasizes, neither a **translation analysis** (which is limited to a comparison of the translated text - i.e. the product - with the source text) nor an **external process analysis**

(which adds to the comparison of texts an analysis of the visible steps taken by the translator) appear to reveal very much of what actually goes on in the translator's mind.

Disadvantages and Limits

Criticisms to thinking aloud are based chiefly on the inconsistency between data and action, on the incompleteness of the data, and on the effects of verbalization upon the action.

As already explained, concurrent verbalization is less subject to inconsistency than is retrospective probing. Nevertheless, thinking aloud cannot accurately reflect the internal structure of cognitive processes. It is therefore vital to develop a **theoretical model** as the basis for the analysis of verbal data. In this context, the question arises as to whether verbal report data can be generalized at all. To answer it, a large amount of data would have to be analysed.

The verbal data are incomplete because some ideas passing through the mind disappear before they can be put into words. Moreover, subjects tend to verbalize incompletely or even to cease verbalizing when the cognitive task to be solved is difficult. It would, however, be inappropriate to oblige subjects to verbalize, because deviations from their normal sequence of thoughts would then most probably occur.

Termbank Evaluation

Thinking-Aloud Protocols (TAPs) include two types of data of prime relevance to the Termbank. The first type is the strategies involved in the use of auxiliary tools for the comprehension of the source text and the production of the target text. The concept of the Termbank should be elaborated on the basis of these strategies so as to meet the needs of the translator (cf. Kohn 1988, 52). The second type is the data which help to evaluate the effectiveness of the Termbank. The factors most important for the evaluation are as follows:

- * TAPs reveal not only the steps taken during the course of the work with the Termbank but also **the train of thoughts and reasons** which lead the translator to take those very steps. From this alone, it is clear that an external process analysis would of itself be insufficient. One would obtain a chronological list of termbank consultations and their results, but miss out on the most important data, namely the links between the individual steps.
- * The verbalizations reveal **to what extent the Termbank contributes to put forward and verify meaning hypotheses** and can thus give us clues as to the appropriateness of

the Termbank structure, the order of the information presented and the value of the information.

- * TAPs document **the parts of the Termbank information actually used** (with the proviso that they be read out aloud by the subject). One can thus determine the criteria responsible for the use, acceptance or neglecting of information (e.g. the location or display of the information in the Termbank).
- * Taken together, the TAP and the translation permit an evaluation of **whether the Termbank information has been dealt with and processed correctly** and hence an evaluation of the contents of the various fields and sections of the Termbank.
- * **The textual problem which occasions a consultation of the Termbank** can also be observed in a TAP and used to check whether the modes of access to the Termbank are appropriate. A good way of evaluating retrieval is the study of the manner in which complex linguistic structures are tackled.

The cases in which the Termbank did not help the subjects solve their problems should be analysed thoroughly. For it is precisely at these "trouble spots" that the process is least automatized, and many of the thoughts passing through the translator's mind should thus be susceptible to verbalization. One can then determine whether the problems are related to weaknesses in the Termbank.

One problem of particular significance for the testing of the Termbank is the fact that subjects do not necessarily classify their problem as being the absence of a target language term, a translation problem, etc. Nor can one always tell from a TAP whether subjects were actually convinced as to the correctness of their decisions. The reliability of the verbal data varies considerably. Some of these disadvantages might, however, be avoided by preparing the subjects carefully for their allotted tasks.

One must not, however, forget that no alternative to thinking aloud as yet exists.

3.2 Elicitation Design

3.2.1 Evaluation text

For our translational evaluation of the Termbank we decided to have a **text** on automotive emission control translated from English into German. It soon became clear that it was anything but easy to find a text that was long enough to cover the central aspects of catalytic converter technology and, at the same time, short enough to be translated within a reasonable

period of time. This is why we "assembled" a source text from various passages taken from Taylor (1984): *Automobile Catalytic Converters*. The text we finally came up with was checked for coherence by a lecturer at our institute who is a translation and terminology expert in the field of automotive emission control and other technical areas.

The greater part of the evaluation text (lines 1-18) is taken from the introduction to the above book "Automobile Catalytic Converters". It therefore represents a cross section of catalytic emission control up to 1984, albeit one that only depicts the situation in the United States. The evaluation text runs as follows:

AUTOMOBILE CATALYTIC CONVERTERS

Catalysts have been widely used to lower the emissions of carbon monoxide (CO) and hydrocarbons (HC) in the exhaust of automobiles in the United States since the introduction of 1975 models. These catalysts, contained in so-called catalytic converters in the exhaust system of automobiles, promote the oxidation of CO and HC to CO₂ and H₂O under net oxidizing conditions (eg. A/F > 14.7). Until 1978, emission control requirements for nitrogen oxide (NOx) emissions were met through non-catalyst technology, primarily exhaust gas recirculation (EGR). Starting with some vehicles sold in California in 1977, NOx emissions from gasoline engines have been subject to catalytic control. The catalyst here has the additional function to promote the reduction of NO to N₂ via reaction of NO with hydrogen or CO. Catalyst systems designed to reduce NOx are considerably more complex than the earlier control systems. For example, the control system introduced by General Motors on some 1978 model year cars has closed-loop air-fuel ratio control (closed-loop fuel metering system, exhaust gas oxygen sensor and an electronic control unit) as well as a three-way catalyst which simultaneously promotes the conversion of HC, CO and NOx. Stringent federally mandated emission control requirements of 1 gram per mile (g mi⁻¹) for NOx have led to the further application of three-way catalysts.

The exhaust emissions of CO, HC and NOx vary as a function of air-fuel ratio as well as several other parameters such as ignition timing and EGR. Engine operation at lean A/F (net oxidizing condition) results in lower HC and CO emissions and more O₂. These conditions favor subsequent catalytic oxidation reactions. At lean A/F the exhaust contains insufficient reducing agents to react with all the O₂ and all the NO. By operating closer to the stoichiometric A/F more NO can react but the volume of CO which must be oxidized increases. The stoichiometric A/F occurs at about 14.7. Three-way catalysts operate in a narrow A/F band between 14 and 15.

Operation of the three-way catalyst requires that the exhaust A/F be controlled close to the stoichiometric composition. A closed-loop feed-back control system holds the A/F in a narrow region near stoichiometry.

The text was examined by an expert of translation in the field who gave the following assessment:

- * it represents the central **technical and chemical knowledge** of the domain;
- * it very well reflects the **word formation patterns of LSP** (e.g. consisting of a great number of multiple compounds);
- * it reflects the common practice of employing **abbreviations and acronyms** (e.g. *CO, A/F, EGR*);
- * it contains a number of "**semi-LSP words**", which must not be rendered into German in their familiar general language meaning. This is a characteristic syntactic pattern of LSP texts (e.g. *to promote*);
- * it is full of short forms of multiple compounds, i.e. so-called "**clippings**" - a typical feature of the domain and a major difficulty for the translator (e.g. *control system* instead of *emission control system*);
- * it contains terms, which can be understood only against the background of the **conceptual hierarchy** they belong to (e.g. *non-catalyst technology*);
- * it abounds in **collocations** typical of automotive emission control (e.g. *to control the exhaust A/F close to the stoichiometric composition*);
- * it covers a great diversity of **domain problems and translational difficulties**.

It was concluded that for both lexical and syntactical reasons, there is enough evidence that the evaluation text is an **appropriate touchstone for the evaluation** of the Termbank's strengths and weaknesses.

3.2.2 Prototype Termbank

Former students who had written their thesis on translational problems in the field of automotive emission control worked out, under the supervision of the department's translation and terminology expert, a Prototype Termbank covering the central aspects of the area under consideration. Over a period of about three months terminological data for the Core Termbank were gathered and Extended Termbank information was elaborated upon. The Core Termbank was supplemented by information made available to us by the terminology team of the University of Surrey. The Extended Termbank information had already been drawn up in the course of our second deliverable (Albl et al. 1991). Due to software and hardware problems, we were unfortunately unable to run the test on the SUN. This was not, however, a major drawback, as our aim was not to test the Termbank's architecture but solely its design in terms of types of contents. As an alternative, the tests were run on MULTITERM - a multilingual terminology processor. All the available data was fed into the MULTITERM

termbank system, in a way that allowed for the most flexible navigation possible. The following entry record shows the information our users were provided with:

Lemma

air/fuel ratio

Abbreviation

A/F

Variant

air-fuel ratio

A/F ratio

Usage

A/F is often used short for *A/F ratio* as well as *A/F mixture*.

Meaning Definition

The proportion of air to fuel in the working charge of an internal-combustion engine or in other combustible mixtures, expressed by weight for liquid fuels and by volume for gaseous fuels.

Synonym

A/F mixture

Equivalent:

Luft-Kraftstoff-Verhältnis

Mischungsverhältnis/ Gemischverhältnis (neutral)

Mischungsverhältnis von Kraftstoff und Luft (erkl.)

Verbrennungsluft-Gemischverhältnis (press., umgang.)

Transfer comment

Im Englischen werden die Mehrwortbenennungen *A/F ratio* (1) und *A/F mixture* (2) häufig im Text zu *A/F* gekürzt. Bei der Übersetzung muß beachtet werden, daß diese Kürzung die Bedeutung der vollen Mehrwortbenennungen trägt:

(1) *A/F* => dt.: *Luft-Kraftstoff-Verhältnis* oder *Gemischverhältnis* oder *Mischungsverhältnis*

(2) *A/F* => dt.: *Luft-Kraftstoff-Gemisch*

Für das englische Akronym *A/F* existiert im Deutschen keine äquivalente Kürzung.

Bei Zahlenangaben zum *Luft-Kraftstoff-Verhältnis* kann auch im Deutschen das Akronym *A/F* entlehnt werden ($A/F > 14.7$).

Im Gegensatz zum *Luft-Kraftstoff-Gemisch* bezeichnet das *Luft-Kraftstoff-Verhältnis* einen Wert (z.B. ein *Luft-Kraftstoff-Verhältnis* von 15 kg Kraftstoff zu 1 kg Luft). Beide Termini können jedoch in den meisten Kontexten gebrauchssynonym verwendet werden (nicht jedoch z.B. bei *Verbrennung des Luft-Kraftstoff-Gemischs*).

Luft-Kraftstoff-Verhältnis ist *Kraftstoff-Luft-Verhältnis* vorzuziehen, da in der Literatur bei der Erklärung des Begriffs fast ausschließlich in dieser Reihenfolge verfahren wird, z.B.: "*Luft-Kraftstoff-Verhältnis*, d.h. 14,7 kg Luft zu 1 kg Kraftstoff".

air ratio darf nicht mit *air/fuel ratio* verwechselt werden! *air/fuel ratio* => dt.: *Luft-Kraftstoff-Verhältnis*; *air ratio* => dt.: *Luftverhältnis Lambda*. Das Luft-Kraftstoff-Verhältnis ist das Verhältnis der Luftmasse zur Kraftstoffmasse. Die Luftzahl bzw. das Luftverhältnis Lambda dagegen gibt an, inwieweit die zugeführte Luft vom theoretischen Luftbedarf abweicht: $\text{Lambda} = \text{zugeführte Luftmasse} : \text{theoretischer Luftbedarf}$.

(--> Unit: STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS)

Hierarchy

-

Collocation

- to (continuously) correct/modulate/alter the A/F toward(s) a stoichiometric (A/F) mixture/ratio
- the # has to be controlled within +/- 0.1 A/F of stoichiometry
- the # cycles around the stoichiometric point
- maintain the # within the limits for efficient catalysis
- trim the A/F mixture (reaching the engine)
- stoichiometrically balanced #
- an # close to stoichiometry
- the A/F mixture is rich/lean of stoichiometry; richer/leaner than stoichiometry
- changes in # either rich or lean from stoichiometry
- at a stoichiometric mean A/F

Domain

[cat-tec]

Wordfamily

-

Grammar

noun

Encyclopaedia

STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS

[ideales Mischungsverhältnis; $\lambda = 1$; $\lambda > 1$; $\lambda < 1$; lambda; Luftmangel; Luftüberschuß; Luftverhältnis; Luftzahl; mageres Luft-Kraftstoff-Gemisch; Sauerstoffmangel; Sauerstoffüberschuß; stöchiometrisch; stöchiometrisches Luft-Kraftstoff-Verhältnis; überstöchiometrischer Bereich; unterstöchiometrischer Bereich]

Der im Tank von Kraftfahrzeugen in flüssiger Form mitgeführte Kraftstoff muß für die Verbrennung im Ottomotor aufbereitet, d.h. mit einer bestimmten Menge Luft gemischt werden. Das mit dem griechischen Buchstaben "lambda" (λ) bezeichnete **Luftverhältnis** (auch **Luftzahl** genannt) beschreibt das Verhältnis zwischen tatsächlich zugeführter Luftmenge und dem für die vollständige Verbrennung des Kraftstoffs theoretisch notwendigen Luftbedarf ($\lambda = L:L_{th}$). 1 kg Kraftstoff (ca. 1,4l) benötigt zu seiner vollständigen Verbrennung etwa 14,7 kg Luft (11,5 m³). Dieses ideale Mischungsverhältnis von 1:14,7 wird als **stöchiometrisches Luft-Kraftstoff-Verhältnis** bezeichnet. Für das Luftverhältnis λ gilt in diesem Fall $\lambda = 1$. Dieser Wert muß möglichst genau eingehalten werden, da bei $\lambda = 1$ die Konversionsrate für die im Abgas enthaltenen Schadstoffe am höchsten ist.

Je nach Betriebszustand des Motors weicht das praktische Mischungsverhältnis vom idealen ab. Wird mehr Luft zugeführt, als zur vollständigen Verbrennung benötigt wird (**Sauerstoffüberschuß**), ist also $\lambda > 1$, spricht man von einem **mageren Luft-Kraftstoff-Gemisch**; der Motor wird dann im **überstöchiometrischen Bereich** betrieben. Bei **Sauerstoffmangel** ($\lambda < 1$) wird das Gemisch als 'fett' bezeichnet; in diesem Fall wird der Motor im **unterstöchiometrischen Bereich** betrieben. Ottomotoren haben bei **Luftüberschuß** ihren niedrigsten Kraftstoffverbrauch, bei **Luftmangel** ihre größte Leistung.

3.2.3 Test Specification

In compliance with the overall aim of evaluating types of terminological information rather than comparing different user populations, our investigation was designed as a **case study**. It should be noted that because of the vast amount of **thinking-aloud** data to be elicited, transcribed and analysed for each individual user, any other approach covering large numbers of subjects would not have been feasible with the scope of the project. Our **test subjects** (hereafter referred to as S1, S2, etc.) were recruited from the group of advanced students at our institute. In order to introduce a necessary element of variation, we took care to distribute them across four different groups according to their previous experience with technical translations:

- (A) students who had done no technical translation whatsoever, i.e. the absolute beginners (S1 & S2);
- (B) students who had done technical translation, but not in the field of automotive emission control, and therefore had no domain knowledge (S3 & S4);
- (C) students who had participated in a translation class on automotive emission control for a period of about five weeks and therefore had some knowledge of this field (S5, S6 & S7);
- (D) students who had participated in the above-mentioned translation class for several terms and who can thus be considered as "experts" in this area (S8 & S9).

In way of preparation for the test, the students were given a **handout** explaining MULTITERM and the Termbank as well as the possibilities of navigation. The tests were finally carried out during a period of nine days, with each subject testing the Termbank individually.

At the outset, the subjects were given a **practical interactive introduction** to the Termbank and its navigation possibilities as a complement to the handout they had read. They then had the opportunity to read and **analyse the evaluation text** before trying to translate it with the help of the Termbank. The method of **thinking aloud** was applied in the process of translating and working with the Termbank. The subjects uttered anything that was going through their minds and verbalized each step performed and the strategies applied. All their

utterances were recorded on tape. In case problems should arise, one of the experimentors was always present to provide help. After they had completed their translations, subjects had another opportunity to give their assessment of the Termbank and its contents in a short **interview**. Note that there was no time limit to the overall process.

The recordings were transcribed in the form of **Thinking Aloud-Protocols (TAPs)**, which were then analysed and evaluated with a view to information categories as well as search strategies and routines. Our **findings** with respect to termbank design and retrieval behaviour are based on 184 pages of transcripts (including the interviews), and will be discussed in the following chapters.

4

Analysis of Information Categories

4.1 Core Termbank

4.1.1 Meaning Definitions

Definitions were found to be helpful insofar as they provided an initial insight into a previously unknown term. This in itself, however, turned out to be **insufficiently instructive**. In the text, for instance, our subjects were confronted with the following passage: . . . *the control system has . . . closed-loop air-fuel ratio control (closed-loop fuel metering system, exhaust gas oxygen sensor and an electronic control unit)*. On looking up and finding all these terms with relatively unambiguous equivalents in the Termbank, it became quite clear that the terms denoted the various parts which constitute the *closed-loop air-fuel ratio control* [*Gemischregelung*]. Yet this did not suffice to know how they were related to one another. Especially in the case of *Kraftstoffdosiersystem* [*closed-loop fuel metering system*], further encyclopaedic background information on how it fitted in with the other parts was considered desirable after reading the definition (S4).

Subjects tended to use not the definitions of the source term to clarify uncertainties, but rather the target language definition. In doubt as to whether *reduction* (in *the reduction of NO to N₂ via reaction of NO*) could actually mean *conversion* und could therefore be translated by *Umwandlung*, subject 6, for example, switched to the German Termbank to look up the definition of the entry *Umwandlung*, and indeed found that *Umwandlung* comprises a *Reduktionsvorgang, wobei NO in N₂ umgewandelt wird* [*reduction reaction converting NO into N₂*].

On the whole, definitions were rarely used. The reluctance to make use of definitions may also have been slightly due to the organization of entries in MULTITERM, where definitions appeared towards the end of a record.

4.1.2 Synonyms

Synonyms were very much welcomed, especially when accompanied by (short) usage notes (S1, S4, S7).

The lemmatization of synonyms is particularly important for the general **accessibility** of entries and the information they offer in the various categories. (*A/F region* used synonymously in the evaluation text for *A/F band*, for instance, was not included as a synonym, but appeared only in the form of an unlemmatized collocation, so that the overall rich information under *band* could not be retrieved.

Consequently, an **exhaustive treatment** of synonyms is highly desirable. It is particularly important in view of the lack of standardization in texts. It seems that, in technical communication, compounds, for example, are created quite freely, which easily results in discrepancies between texts and the Termbank. In our text, for instance, *emission control requirements* is used, but the Termbank only contains *exhaust emission requirements*. Our subject (S4) was in some uncertainty as to whether the two terms were synonymous, and concluded that due to a general variability in texts all possible synonymous options, particularly the various textual realizations of compounds should be included in the Termbank.

It should be noted, however, that it is not enough to give a comprehensive list of synonyms. Elaborated (long) **usage comments** should be provided whenever the differentiation with respect to how synonymous terms are used is needed. In the case of, e.g., *Reduktionsmittel* and *reduzierende Komponenten/Reduktionsgase* [*reducing agents*], or *Zündzeitpunkt* and *Zündwinkel* [*ignition timing*], usage notes such as [*chem*] and [*KatTech*] do not suffice; usage comments are needed to explicate the relevant conditions of use in detail.

4.1.3 Variants

Our tests have confirmed the importance of having abbreviations lemmatized. It would otherwise not be possible to guarantee the **accessibility** of, e.g., chemical formulae such as *HC*, *CO*, or *NO* (S7). Another interesting case is that of *A/F*, which is an abbreviation of *air-fuel* and, at the same time, a clipped variant of *A/F mixture* and *A/F ratio*. Our tests showed that unless *A/F* is made a lemmatized variant of both entry terms it cannot be found by the user. For, as *A/F* often occurs in its clipped form (e.g. *A/F > 14.7, at lean A/F, the stoichiometric A/F*), some users (S1, S4, S6), unfamiliar with the term and what it stands for, took it as a full variant, and not as a clipped form. They therefore searched for *A/F* only, instead of typing in *A/F*+ "wild card", and thus did not find the term.

It was also thought (S4) that it is important to have all variants **exhaustively treated** in the Termbank, since in most cases there are not only the common ones such as H_2O and CO_2 , but also others such as HC and CO , with which some users are less familiar. Users were always delighted on finding terms together with their respective abbreviations, such as *exhaust gas recirculation/EGR*.

Another argument that makes for the lemmatization of variants is the accessibility of all possible **spelling variants**, such as *oxidize/oxidise* or *carburetor/carburetter*.

4.1.4 Collocation

An interesting finding in our tests was that users had quite clear ideas about what should be a compound and therefore lemmatized and what is a collocation and is therefore not expected to be found as a lemma.

Subjects (S1, S4) were, for example, surprised to find *lean* as an independent lemma, especially in view of its general language character. In addition, *lean* in technical language refers exclusively to the composition of the *air/fuel ratio* and mostly occurs in the form of *lean A/F*. It was therefore felt that *lean* should appear as a collocation of *A/F*. A similar case was that of *to operate in a narrow (A/F) band* [*Einhaltung des engen Regelbereichs*]. *Narrow A/F band* is treated as a lemma in the Termbank. It did not occur to the subjects (S4), however, that this should be a separate lemmatized entry term. It was instead seen as a collocation.

Net oxidizing conditions or *reducing agents*, on the other hand, were, even when the phrase had never been heard of before - as was the case with the beginners (S1) - intuitively taken as tightly-knit compounds which were expected to have lemma status. It seems that rather loosely collocated phrases (e.g. those beginning with general language adjectives) might best fall under collocations, while nominalized phrases, which appear to be fixed compounds, should be lemmatized as independent entry terms.

An argument in favour of the **non-lemmatization of collocations** is the fact that they often comprise extremely long phrases which cannot be shortened without losing their value, e.g. *to promote the reduction of NO to N_2 via reaction of NO with...* or *a closed-loop feed-back control system holds the A/F in a narrow region near stoichiometry* (S6).

In the case of non-lemmatization, however, accessibility is somewhat more difficult. If, as in our Prototype Termbank, a phrase like *widely used* [*serienmäßig eingesetzt*] appropriately appears as a collocation of *catalytic converter* but not of *catalyst* (for, strictly speaking, it is

the *converter* not the *catalysing substance* that is "*widely used*"), the user will not be able to retrieve the collocation. Therefore the undogmatic, or undefined, use of phrases in texts, e.g. *catalysts that are widely used*, has to be accounted for. As a consequence, collocational phrases should appear in as many related entries as possible, i.e. under various agents (e.g. *catalysts, catalytic converters*) or patients (e.g. *oxidation, conversion, reduction*) (as in *catalysts promote the oxidation of*), or, in case the verb is a lemma, even under the verb (e.g. *to promote*) (S4).

Collocations might also be extremely important in that they are the only category suitable for dealing with the many **general language words** which give rise to difficulties in technical texts. They include, in particular, a wide range of general language **verbs**, such as *to promote the reduction, stoichiometric A/F occurs at, operate in a narrow A/F band, hold the A/F in a narrow region near stoichiometry*, etc (S6, S7); but also **prepositions** like *to promote the reduction of NO to N₂ via reaction of NO with ...*

4.1.5 Compounds

Compounds turned out to be one of the most interesting phenomena in our user tests. They confront the user with the greatest problems regarding comprehension, retrieval and translation.

Net oxidizing conditions, for example, is one of the more specific terms for which there is no German equivalent, but only a suggested translation, and which was therefore not lemmatized as a compound in our Termbank. It occurred to none of our subjects that it could be found as a collocation of *to oxidize*. We therefore believe that the question as to which terminological expressions should be treated as lemmatized compounds, as opposed to collocations, is of paramount importance.

The following example goes to show that a great number not only of fixed compounds but also of arbitrary nominalized term combinations should be taken up as compound lemmata by the Termbank. For most subjects, it was very confusing to find that *emission control requirements* did not in this form appear in the Termbank. What they found instead was either *exhaust emission requirements* (if "wild card"+*requirements* was the query) or *emission control standards* (if *exhaust emission*+"wild card" was typed in), not knowing whether *emission control requirements* and *exhaust emission requirements* or *emission control requirements* and *emission control standards* were synonyms. Furthermore it did not become clear that *requirements* and *standards* are, although strictly speaking non-synonyms, often synonymously used. Indeed, in the text, *Grenzwerte* [*standards*] turned out to be the more appropriate translational solution for *requirements* [*Vorschriften*]. The various compounds

should, therefore, be listed as (near-)synonyms together with a clear usage comment stating the difference between them. If the terms are too far apart to be listed as synonyms the usage comment could provide cross-references to related but non-synonymous entry terms. Thus, in addition to the **exhaustive treatment of all compounds**, the **links between compounds** have to be explicated.

It was generally stressed that, while all compounds should be integrated, this was especially important for the semantically rather **vague compounds** such as *catalytic control*, *catalyst system* and *non-catalyst technology* (Subject 6 said: "I feel that *catalyst system* should have its own entry, coming up as often as it does, even though it may be clear what is meant by it. Other expressions are also treated even when they are quite clear" - "Ich finde *catalyst system* sollte einen eigenen Eintrag haben; es kommt ja oft genug vor, auch wenn vielleicht klar ist, was damit gemeint ist, aber es sind ja auch andere Sachen aufgeführt, die klar sind"). These compounds typically do not occur in dictionaries like the Schmitt dictionary, which mostly deals with the more concrete facts and objects. However, it is these vague compounds that lead to enormous translation problems (S8), especially where there is no equivalent, so that a suggested translation has to be given (e.g. for *non-catalyst technology*).

An even more complicated case is that of the many **compound clippings** typical of English texts. In the evaluation text, for example, *control system(s)* appeared twice as a short form of *emission control system*. In the Termbank, the user (S1, S6) could find *emission control (system)* [*Abgasreinigung(ssystem)*] and *control* [(1) *Regelung/Steuerung*, (2) *Verminderung, Abbau*]. In the Schmitt dictionary, however, *control systems* is given only in its basic meaning: *Steuer- und Regelsysteme*. Apart from our "expert" users (S8, S9), all our subjects had difficulty finding the exact meaning of *control system(s)*, since the information that *control system* could stand for *emission control system* was only given in the transfer comment under *emission control system*.

Another example belongs to the general field of automotive engineering: The term *exhaust system* [*Abgasanlage*] was not part of our Termbank on catalytic emission control. What the user found, however, was the term *exhaust purification system* [*Abgasreinigungssystem*], of which *exhaust system* was, therefore, wrongly taken to be a short form; this then led to an incorrect translation (S6). It therefore seems not only the whole range of compounds that should be integrated exhaustively into the Termbank (including those extending to other domains), but all their clippings, too. In addition, usage comments (cf. chap. 4.2.1) should be introduced to explain the usage conditions governing the textual occurrence of clipped compounds (cf. chap. 5.3.2).

4.2 Extended Termbank Information

4.2.1 Usage Comments

Usage information in general has been considered extremely important (S4). Users were delighted when they found usage differences between terms explained, as was the case with *Emissionen* and *Schadstoffe*.

However, it became clear that (short) usage notes, i.e. ISO-standardized information - such as *frequent*, *depricated*, *in-house* - does not suffice. Synonyms, variants and compounds need an additional **usage comment** specifying the differences between the listed terms and/or explicating the major peculiarities in connection with them. The following types of problems should be dealt with:

(a) peculiarities in the use of synoymys or variants

In the case of *feedback* and *closed-loop*, it is not enough to present these expressions as synonyms; there should also be a usage comment stating that both are often used tautologously (as in *closed-loop feed-back control system*) in order to avoid confusion (S1). The same applies to cases where the synonymous use is irritating if in contradiction with general knowledge of the world, as in the case of *Luft-/ Sauerstoffüberschuß*, *Luft-/Sauerstoffmangel*, where it took some users a very long time to become clear as to the synonymous use of *Luft* and *Sauerstoff*.

(b) distinguishing between synonyms or variants

Considerable variety in the usage of terms of synonym or variant clusters, often occasions great difficulty (S4) as, e.g., in the case of *Umwandlung* [*conversion*] with six German synonyms. Not only is there a difference in use between *Konvertierung* and *Konversion*; it is also important to know how *Umwandlung* is used in the chemical field, especially since it can stand for *reduction* in the sense of *Verringerung/Verminderung*. It is not always obvious whether a term can be used in the technical or the chemical field or both.

(c) compound clippings

A usage comment might be best to clarify problems with regard to compound clippings. A usage indicator could make clear in what way the textually intended meaning of clipped compounds can deviate from its denotational meaning (cf. chap. 4.1.5).

(d) textual and register conventions

LSP conventions were found to be of immense importance (S8). The whole "feel" of a technical text will change depending on whether a translator speaks of *Autoabgaskatalysator* or *Abgaskatalysator für Ottomotoren*. In order simply to distinguish between the two

synonymous expressions, a usage note such as [*pop./press*] and [*tech*] would suffice. Clarification of the usage concerning *Ottomotoren*, however, was found to be urgently required (S7, S8). For that purpose a usage comment would be necessary.

The above example (d) also makes clear the **interaction of usage comment and transfer comment**. *Automobile*, frequently used in various phrases in (American) English (such as *exhaust of automobiles*, *automobile catalytic converters*), is not translated by *Kraftfahrzeug*, but by *Ottomotoren* (*im Abgas von Ottomotoren, katalytische Abgasreinigungsverfahren für Ottomotoren*), so that in the transfer part it is the TC that brings home to the user this German convention, which has to be borne in mind in the translation process.

One should investigate to what extent long usage comments could be a category whose function includes the treatment of all the vagaries and inaccuracies of LSP texts.

4.2.2 Transfer Comments

Transfer comments (TCs) were the categories most frequently made use of by our subjects. In terms of translation-relevant information, they were considered to be extremely helpful (S7) and, indeed, indispensable (S4).

The TC on *stoichiometry* is one of the popular examples, as it very clearly states that the corresponding *Stöchiometrie* in German only exists in the chemical field, while in the technical field it has to be expressed by, for example, *stöchiometrischer Punkt*. This is the kind of information not usually found in any dictionary and therefore highly welcomed by all our subjects. If criticism was nevertheless levelled at the comments, this was due to non-compliance, in some cases, with the following **criteria in the elaboration of the TCs**:

(a) term consistency

Transfer comments need careful matching with respect to the transfer equivalents they comment upon; i.e. it has to be thoroughly checked as to whether all equivalents dealt with are also listed as equivalents.

(b) content analysis

All terminological aspects explained in the TCs should be examined with respect to their relevance, as they can otherwise be misleading.

(c) transparency

Users should not be confused by a combination of straightforward information with highly difficult technicalities. The relation between the two should be made explicit and the difficult

part should be dealt with in a separate sentence (with a cross-reference to the relevant unit explaining the context for the beginners).

(d) explicitness

The above aspect begs the question as to the **user-relevance** (which is also important for encyclopaedic units). Since TCs are by nature very concise and specific and, therefore, not always easy to understand, they have to be thoroughly elaborated upon, stating what they want to convey in a very explicit, unambiguous and clear way.

(f) cross-references

The needs of beginners can best be met by extensive cross-references to other categories, especially encyclopaedic units, so that they can receive additional information if need be. In fact users gladly made use of such a possibility.

(g) structure

More than anything else, it is the **structure** of the TCs that supports their usefulness. An extreme example was the TC giving a host of relevant information on the translation of *A/F*, e.g.:

- *A/F ratio* and *A/F mixture* [*Luft-Kraftstoff-Gemisch/Luft-Kraftstoff-Verhältnis*], although strictly speaking non-synonyms, are often synonymously use;
- in German it has to be *Luft-Kraftstoff-Verhältnis* rather than *Kraftstoff-Luft-Verhältnis*;
- *air ratio* must not be confounded with *air/fuel ratio*;
- *A/F ratio* and *A/F mixture* are often referred to only by the acronym (*A/F*), whereas in German the corresponding full compounds are used;
- there is no German equivalent for the English acronym *A/F*;
- the English acronym can be used in German if followed by figures as in *A/F > 14.7*.

All this information was presented in a rather long and not very well organized TC, which was consequently often not read in its entirety (S7), so that relevant information was sometimes simply overlooked. It is somewhat ironic to observe that lengthy and laborious searching was undertaken by some users although the required information actually was in the TC. In the post-test interview it was therefore stressed that the TCs needed very clear structuring.

(h) length

Lengthy TCs are certainly not very conducive to the aim and the kind of information they provide. The question still to be investigated is whether one TC should contain all the various pieces of transfer information on one entry term (and its equivalents) or whether these should

be dealt with in separate comments, which might contribute to an eventual classification of transfer comments.

4.2.3 Encyclopaedic Units

The tests revealed a discrepancy between the pronouncements made by subjects in the post-test interviews to the effect that encyclopaedic units were very interesting (S4) and in some cases indispensable (S1, S2), and the fact that they were not frequently consulted. The following observations were made with respect to the Encyclopaedia.

Positive points

(a) important supplier of background information and more helpful in this respect than dictionaries

Beginners in particular needed to have the facts, objects and factual background of a given domain explained in order to understand rather exotic terms such as *lambda*, a term which occurs very often, as well as the central terms such as *A/F*, *stoichiometric*, or *Lambdaregelung*. It was therefore thought extremely useful to have an encyclopaedic unit such as that on GEMISCHREGELUNG (S3, S4).

(b) means of orientation

Beginners above all are given an opportunity to gain initial insights and an idea of what the subject is all about without which a proper understanding of the (rather difficult) technical text would not have been possible (S1, S8), e.g. in the case of *the functioning of the exhaust oxygen sensor [Lambdasonde]*, *the broader context in which to place the A/F ratio [Luft-Kraftstoff-Verhältnis]* or *the stoichiometric mixture [stöchiometrisches Gemisch]*, etc.

(c) means of presenting the relations and interactions of components

Subject 1 felt that it was through the units that an understanding of the context in which to place the various aspects could be gained as well as a picture of how they all fitted together. The encyclopaedic unit AUFBAU DES KATALYSATORS, for example, gives a clear picture of the components that make up a catalyst. More importantly, encyclopaedic units are well fit to provide super- and subordinate but not strictly hierarchical relations as could not be presented by the category 'hierarchy' (e.g. the encyclopaedic unit ABGASREINIGUNGSVERFAHREN gives an overview on the following relations: *Emission control* comprises *non-catalyst technology* as well as *catalyst systems*; *non-catalyst technology* includes *engine-related measures [motorinterne Maßnahmen]* as well as *thermal afterburning [thermische Nachverbrennung]*; *thermal afterburning* as well as *catalytic afterburning [katalytische Nachverbrennung]* are

engine-external measures [Abgasnachbehandlung mit Zusatzeinrichtungen]; *engine-internal measures* do not only comprise measures such as *ignition timing* [Einstellung des Zündzeitpunktes], *fuel metering* [Kraftstoffdosierung], etc. but also *exhaust gas recirculation* [Abgasrückführung].

(d) supplement to meaning definitions

Given the complicated nature of the subject in question, definitions often did not suffice to create a proper understanding of a domain aspect, e.g. the role of *Kraftstoffdosiersystem* as part of the *control system* (cf. chap. 4.1.1). Additional factual information was obviously required, which only the encyclopaedic units could provide (S4, S8). By the same token, culture-specific points could not possibly be provided solely by meaning definitions, e.g.: *the fact that due to an initial introduction of automotive emission control in California, the more stringent emission control standards there differ from those in the other 49 states. This is the reason why standards are given with the indicator of 49 states in brackets, i.e. emission control standards of 1 gram per mile ($g\ mi^{-1}$) (49 states).*

(f) last resort

Units were often consulted when no satisfactory solution emerged from the Termbank (S8).

Negative points

(a) intricacy

In many cases, encyclopaedic units were found to be too specific and concise (S1, S4), full of complicated technical terms, which appeared rather exotic to beginners (e.g. *stoichiometric - S1*); they were often too difficult to read. Some users therefore had the impression that they would first have to look up words in order to understand the technical knowledge the units were designed to convey. The unit *REGELUNG/STEUERUNG DER GEMISCHBILDUNG BEI OTTOMOTOREN*, for example, was a very complex one: Subject 4 found it difficult to be confronted with the *ratio* $L:L_{th}$; and subject 4 imagined that L stood for *Luft*, but was unclear as to the meaning of L_{th} , although this was explained as referring to *theoretischer Luftbedarf*. Naturally, this applied even more to the English units which were felt to be particularly complicated, as in the case of *fuel metering system*, where subject 4 wanted additional information, but was then put off.

(b) information overload

In view of the concise and intricate nature of the encyclopaedic units, users sometimes felt that they had not the processing capacity to cope with the flood of new information they were faced with (S1, S4, "Da fühlt man sich völlig erschlagen", S6). As the units were felt to be aimed mainly at the beginners in a field (S8), subjects said that they should be written in a

more **redundant** way to help the user grasp their meaning (S1). In order to improve their intelligibility, units should be less complex and more illustrative ("weniger komplex und anschaulicher", S4).

(c) unfamiliarity

An additional factor contributing to the unease with which the encyclopaedic units were approached was that users were not fully acquainted with the very concept of the Encyclopaedia, so that it was sometimes only after a while that they understood what they had read in an unit earlier on (S4). Users would have preferred to have had some kind of **manual** specifying how the units are organized and how they should be used.

(d) time-consuming nature

For the above-mentioned reasons, encyclopaedic units were considered to be too time-consuming, and users were reluctant to make use of them. On the one hand, this was due to the fact that the sophisticated TWB-design of the Encyclopaedia could not be provided under MULTITERM. On the other hand, users turned out to be over-eager to have quick solutions, although recognizing the necessity for a thorough understanding of the subject matter. This interesting phenomenon will be discussed in chapter 5.

(e) terminological information

In line with the phenomena mentioned under (c) and (d) it soon became clear that users approached the Encyclopaedia with expectations that this was not designed to meet in the first place. They wanted the units to deliver customized solutions for translation-specific terminological problems. It is, of course, not the task of encyclopaedic units at all, to provide in all cases the idiomatic way of putting something, i.e. the texttype and context adequate expressions. As a result, encyclopaedic units were considered beneficial insofar as they helped confirm or reject decisions the translator had arrived at, but were not appealing enough as they were not sufficiently in line with the users' wish to **hit instantly on prefabricated terminological solutions** (cf. chap. 5). While this made it clear that users have to be told how to make good use of the Encyclopaedia (cf. (c)), yet another conclusion was arrived at. Instead of covering a domain in a rather global fashion, encyclopaedic units should be selectively geared towards particular text/translation problems for whose solution they are relevant. Therefore the starting point should be intricate terminological problems for which specific subject knowledge is needed.

Accessibility

Encyclopaedic units were accessed:

(a) to gain a general understanding of the domain

This obvious function of the Encyclopaedia, i.e. to make factual background knowledge approachable via the list of headers, was only rarely resorted to (S3). It was rather with respect to specific problems that users were willing to turn to the units.

(b) to supplement Termbank information

Users mostly consulted the Encyclopaedia, when the Termbank was found wanting, i.e. when the information it provided was felt to be incomplete, when terminological expressions could not be found, or when the retrieved information, e.g. meaning definitions, did not provide the necessary support (S1, S4, S7).

(c) via cross-references

Units were obviously also accessed in cases where references were provided from other information categories, such as the TCs (e.g. the TC on *closed loop air fuel ratio control* referring to the encyclopaedic unit *REGELUNG/STEUERUNG DER GEMISCHBILDUNG BEI OTTOMOTOREN*, or that on *catalyst*, which led to the encyclopaedic unit *AUFBAU DES KATALYSATORS*, S1, S4).

(d) on advice

When no cross-references were provided, the units were, more often than not, consulted only on advice given by the experimenter. This was especially the case with *non-catalyst technology*, a term not to be found in the Termbank at all. No cross-referencing was therefore possible. The relevant background information, i.e. the underlying conceptual hierarchy implying a possible translational solution, was to be found in the encyclopaedic unit *ABGASREINIGUNGSVERFAHREN*, which could, however, not be inferred by the user.

All this demonstrates that there is a strong need for reasonably transparent indications paving the users' way into the encyclopaedic units. Thus, in order to secure accessibility, not only the **grouping paths** via the headers have to be thoroughly elaborated by the terminologist, but also the **cross-references** from other termbank categories. Unless appropriate **navigation procedures** are mapped out, users will not make use of this otherwise valuable source of information. It is especially the transfer comments with their concise and intricate information as well as the rather short meaning definitions that need to be linked up with complementary encyclopaedic information.

Integration

Extensive networking, naturally, tallies in with the concept underlying the design of the Encyclopaedia. There are two aspects which need to be distinguished:

(a) internal integration

Users were often deterred by the English units and would have preferred to read the domain-specific technicalities in their mother-tongue. However, the English encyclopaedic terms (i.e. terms linked to encyclopaedic units) did not always have a German counterpart (e.g. *reducing agents*, *ignition timing*). A very important aspect soon became clear: Although it is certainly advantageous to write **English and German units** independently, corresponding as it were to their respective LSP systems, they afterwards have to be thoroughly **matched** and linked with one another to allow for free and flexible navigation between the Encyclopaedias of the various termbank languages. It has been explicitly stated in the interviews that reading the same subject-related information in both the **source and the target language** is often of greater help than bilingual dictionary information ("Ich hab' bisher immer die Erfahrung gemacht, daß ich weiter komme, wenn ich Fachbücher lese, zuerst mich im Deutschen kundig mache, um zu verstehen, um was es geht und dann das gleiche Fachbuch auf Englisch lese und dann die englischen Termini im Fachbuch finde; also ich arbeite sehr selten mit Wörterbüchern", S4). The Encyclopaedia is not only perfectly suited to cater for this requirement; it is also the only information category of the Termbank to fulfill this task.

(b) external integration

The Encyclopaedia also needs to be thoroughly **matched up with the rest of the Termbank**. Encyclopaedic terms establish the necessary connections. They need to be carefully selected on the basis of comprehension and translation problems requiring encyclopaedic information. In other words, (encyclopaedic) terms should only be linked to those encyclopaedic units which provide information which is relevant for these terms. Mere occurrence in a unit does not justify for a term to be given "encyclopaedic" status.

5

Analysis of Search Strategies

5.1 The Translator's Paradox

5.1.1 The Discrepancy between Theory and Practice

It is widely accepted, today, that knowledge of the world is indispensable in discourse processing. People do not understand and produce utterances solely on the basis of their linguistic knowledge; their world knowledge is equally important. In creating textual meaning - in the comprehension as well as in the production process - two types of processing can be distinguished: *bottom-up* and *top-down*. In bottom-up processing we identify lexical and grammatical clues to meaning provided by the text; top-down processing taps the general world knowledge and specialized subject knowledge of the readers/producers to create expectations and inferences. The successful creation of textual meaning always requires a close and complex interaction between bottom-up and top-down (Kohn 1990). Clearly, these processes are of particular relevance when translating a text from one language into another, and this is why we favoured the idea of including an Encyclopaedia in the Termbank (cf. Albl et al. 1990).

These theoretical reflections were also supported by our test subjects, who emphasized the importance of getting a general idea of a special area before starting to translate a LSP text. In practice, however, things turned out to be quite different: only one test subject (S3) of the four with no knowledge of catalytic converter technology decided to start off by reading some basic German encyclopaedic units (1. KATALYSATOR, 2. GEMISCHREGELUNG, 3. STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS) before trying to solve the quite specific problems the text presented. Equally, the subjects who had only little knowledge of the domain hardly ever had recourse to the meaning definitions and/or the Encyclopaedia. Being used to working with conventional dictionaries in their daily routine, our translators for the most part took the accustomed way of going directly to a specific lemma ("Ich nehm dann halt immer den Weg, daß ich wie im Wörterbuch ganz einfach den Begriff nachschaue.", S1).

The retrieval behaviour of users turned out to be determined by the aim of finding concise answers to their translation problems in as short a time as possible. Users are interested in

obtaining quick and customized translation solutions and clear-cut equivalents, rather than in gaining a detailed understanding of the area under consideration ("...stehen wieder viele Sachen drinnen, die ich im einzelnen gar nicht durchlese, weil ich im Grunde genommen nur an der Übersetzung interessiert bin.", S4 commenting upon *closed-loop air-fuel ratio control*; "...weil ja letztendlich meine Aufgabe schon ist, den Text zu übersetzen und nicht da Spezialistin für Katalysatortechnik zu werden.", S3; "*net oxidizing conditions* - das habe ich zwar schon mal gehört, aber ganz genau weiß ich nicht, was das ist und schon gar nicht, wie ich das übersetzen soll ... *A/F* muß wohl was mit dem Luft-Kraftstoff-Verhältnis zu tun haben, aber was es genau damit auf sich hat ... an sich weiß ich ja, was es damit auf sich hat, bloß wäre mir ganz wohl, wenn ich diese *net oxidizing conditions* finden würde, am besten gleich mit einem Äquivalent", S7).

As a result, we are faced with a somewhat paradoxical situation: Although our test subjects acknowledged the fact that world knowledge - and hence specific subject knowledge - is a prerequisite for arriving at appropriate translational solutions, they made hardly any use of the Encyclopaedia, even when they were quite aware of the fact that they were lacking certain technical information. Instead of trying to grasp the underlying meaning of a certain state of affairs, users often preferred to turn to equivalents, transfer comments and/or collocations for help, hoping to obtain relevant information more quickly that way ("jetzt hab ich bloß auch vergessen, wie das is, Lambda größer 1, kleiner 1, deshalb schau ich jetzt einfach mal bei *A/F* mit Stern nach und schau noch mal bei *air-fuel ratio* ... transfer comment, den könnt ich doch gebrauchen", S7). It is symptomatic of our subjects' search strategies that only one of them (S4) took the trouble to crosscheck the solutions found by consulting encyclopaedic units or meaning definitions. All users considered the Termbank very helpful whenever the search path was extremely short, i.e. lemma --> equivalent (possibly transfer comment), and/or whenever the Termbank presented a problem in exactly the same way as the evaluation text did, so that entire phrases could be conveniently copied ("daß ich einen ganzen Satz aus der Datenbank übernehmen kann, find ich ja ausgesprochen positiv", S5 commenting upon *by operating closer to the stoichiometric A/F*). For this reason, equivalents, transfer comments and collocations were by far the users' favourite categories. Encyclopaedic units were unpopular by comparison.

The discrepancy between the evident need to bridge knowledge gaps by consulting meaning definitions and, in particular, encyclopaedic units and, at the same time, the desire to obtain prefabricated translation solutions by consulting only equivalents and transfer comments might be referred to as the **Translator's Paradox**. In this context, two questions are of particular interest. First, why is it that our translators were so reluctant to look up things in the Encyclopaedia? And second, is it still possible for translators to arrive at appropriate

translation solutions without having grasped the meaning of the particular piece of text they are working on? These questions will be examined more closely in the following sections.

5.1.2 The Time Factor

The fact that the Encyclopaedia was not as popular with our subjects as we would have expected it to be - something which became increasingly clear in the course of our test - can be put down to two factors which are, in fact, interrelated. For one thing, encyclopaedic units were regarded as rather specific, presenting background information in a very condensed form so that those in particular who were unfamiliar with the field of automotive emission control had difficulty understanding them ("Actually, I felt completely overwhelmed [by the information given]", ["Da fühlte ich mich eigentlich total erschlagen"], S4; "Again, the explications are very specific", ["Die Erklärungen sind wieder sehr fachspezifisch"], S1). Consequently, reading encyclopaedic units requires quite a lot of time, and here we are faced with the second problem which is of rather a general nature: the fact is that, in practice, translators more often than not work under great pressure of time. They simply cannot afford to spend too much time on their work, if only because of the deadlines they have to meet and the amount of money they receive. Although this was not the case with respect to our test - the subjects had all the time they wanted - their main concern was still to find quick solutions rather than to gain a thorough understanding of what they were translating. Reading encyclopaedic units was simply considered to be too time-consuming. Even subject 3, who studied a few introductory units at the outset, repeatedly stressed the importance of time: "... this is a question of time", ["das ist halt wieder eine Zeitfrage"]; "If I had a lot of time, I would take a look at the corresponding unit as well, but that's too much for the moment", ["Wenn ich jetzt ganz viel Zeit hätte, würde ich mir auch noch die entsprechende unit ankucken, aber das ist mir jetzt zuviel"]. Concerning encyclopaedic units, subject 2 said: "This is really interesting; you can learn a lot about catalyst technology by reading those units. I'd still like to look up a few terms, though" ["Das ist echt interessant; da erfährt man ja einiges über Katalysatortechnik, wenn man die units liest. Trotzdem will ich noch einige Begriffe nachschauen"]. Later on, subject 2 felt that he had read a piece of information in a unit which might be useful to him. The problem was that he could not quite remember where it was. After thinking about it for a few moments, he decided to leave it at that and to look up a few more terms instead.

It therefore comes as no surprise that the Termbank was praised whenever it presented ready-made solutions such as the following:

closed-loop air-fuel ratio control -> equivalent: *Gemischregelung*

exhaust gas oxygen sensor -> equivalent: *Lambdasonde*

electronic control unit -> equivalent: *elektronisches Steuergerät*

air-fuel ratio -> equivalent: *Luft-Kraftstoff-Verhältnis*

Quite obviously, these terms do not in themselves present major difficulties, so that there is no evident need to verify things by taking a look at the corresponding encyclopaedic units. But even in cases where units were consulted for technical information, it was quite exceptional to find a translator taking the trouble to read more than one or, at the most, two units when trying to solve a particular textual difficulty. Very often, however, encyclopaedic units were not consulted for *factual* information, but for *linguistic* information, i.e. information concerning the conventions of LSP texts.

The remarkable thing in the overall process was that users tended to stop navigating within the Termbank even though they had not yet fully grasped the core of a given problem. They relied too heavily upon the information provided by the Termbank and, generally, gave too little personal thought to the difficulties they encountered in the text ("I don't understand that, but, well, I'll take it for granted", ["Das verstehe ich zwar nicht, aber ich nehm's mal hin"] S4; "Once you've started [working with the Termbank] you tend to look up more and more terms, and you don't realize that you could find the solution a lot quicker by thinking about it" ["Wenn man mal mit diesem Ding angefangen hat, daß man tausend Sachen versucht zu finden und eigentlich schneller drauf käme, wenn man mal ein bißchen überlegen würde"] S7). The question now is whether translators can come up with appropriate translation solutions without having fully understood the meaning of a particular passage of text.

5.1.3 Translating Along the Text Surface

A highly illustrative example of the Translator's Paradox is the phrase *under net oxidizing conditions*, which simply means that there is more oxygen in the combustion chamber than is theoretically needed for complete combustion of the fuel content. The whole complex relating to *stoichiometry*, *lean/rich* and thus to *net oxidizing conditions* is explained in both an English and a German encyclopaedic unit (THE AIR-FUEL-RATIO AND ITS INFLUENCE ON EMISSIONS; STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS). The phrase *under net oxidizing conditions*, however, was correctly translated (*bei Luftüberschuß*) by all subjects with the help of non-encyclopaedic information. The very same knowledge complex is referred to several times in the second and third paragraph (*at lean A/F; stoichiometric A/F; stoichiometric composition; stoichiometry*), with *net oxidizing condition* even being expressly repeated in parenthesis. It is here that the search strategies of certain test subjects clearly showed that they

had not fully understood what was actually meant when they translated *under net oxidizing conditions* with *bei Luftüberschuß*. In this respect, the strategies chosen by subject 5 and subject 6 respectively are particularly revealing.

Subject 6, for instance, decided to look up *engine operation* and *lean A/F* not realizing that the latter is basically the same as *net oxidizing conditions*. Furthermore, he looked up *operate closer to the stoichiometric A/F*, *the stoichiometric A/F occurs at*, *stoichiometric composition* and *near stoichiometry*, i.e. each item individually without trying to see how things related to one another. In doing so, he spent a considerable amount of time looking up things that are made clear in the encyclopaedic units mentioned above. Consulting these units would have shortened his search path enormously, but even when advised to take a look at the Encyclopaedia he responded: "...well, yes, I think I'll leave it at that", ["...hmm, na ja, lassen wir's vielleicht erst mal dabei"]. Interestingly, the solutions he came up with were all factually correct, but not always in accordance with the conventions of LSP texts. Generally speaking, the strategy described here, namely trying to do without additional encyclopaedic information, was typical of our participants' behaviour.

By contrast, subject 5 opted for quite a different strategy. In the first paragraph, she proceeded like the others and rendered *under net oxidizing conditions* correctly into German by looking up lemma information. In the second paragraph, however, she realized that she had not fully understood how *A/F*, *stoichiometry* etc. fit together and decided to read the unit STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS. After reading this unit, she commented: "This solves the translation problem" ["Damit wäre das Übersetzungsproblem gelöst"]. Interestingly, she was able to translate all the following text passages relating to this complex efficiently, correctly and appropriately.

On balance, the fact remains that, in general, our test subjects were rather reluctant to use encyclopaedic information. In spite of this, the solutions they came up with by consulting lemma information were factually correct, although, terminologically speaking, often not in accordance with LSP conventions. Reading encyclopaedic units and thus gaining a deeper understanding of how things fit together clearly enables translators to make more informed and efficient decisions, i.e. to arrive at the way in which an expert would put it.

5.2 Relevance of Subject Knowledge

In line with the translator's paradox, the data analysis clearly showed that, when searching for a solution to a particular textual difficulty, **users usually do not take the context sufficiently**

into account. In fact, the transcripts documented only three cases where the context served as a starting point for the translators' search strategies; e.g.

- * text difficulty: *under net oxidizing conditions (e.g. $A/F > 14.7$)* -> search: via *air-fuel ratio (A/F)* (S7)
- * text difficulty: *Catalyst systems (designed to reduce NOx are considerably more complex than the earlier control systems.)* -> search: via *control* (S8)
- * text difficulty: *(...emissions were met through) non-catalyst technology (primarily exhaust gas recirculation (EGR)...)* -> search: via *EGR* (S6).

Apart from these few exceptions, users tended to stick very closely to the words on the source text surface. For example, when coming to the last sentence of the evaluation text (*A closed-loop feed-back control system holds the A/F in a narrow region near stoichiometry.*), subject 6 decided to look up *control, closed-loop, feed-back, hold, region* and *stoichiometry*, even though the meaning of the sentence should have been clear by then, since the last sentence was nothing but a paraphrase of what had already been explained in the preceding lines (*Operation of the tree-way catalyst requires that the exhaust A/F be controlled close to the stoichiometric composition*). Unable to draw the appropriate analogies and inferences, subject 6 therefore did not recognize that both clauses referred to the same fact.

As is to be expected, this problem becomes even more acute in the case of the **beginners'** group. It could be observed that those doing a technical translation for the first time often kept **going round in circles**, looking up certain entries twice or even three times.

- * text difficulty: *catalytic converter vs. catalyst*
 search (S1): *catalytic converter* -> transfer comment -> unit AUFBAU DES KATALYSATORS -> unit THREE-WAY CATALYTIC CONVERTER -> unit 3-WEGE KATALYSATOR (-> advice given by the experimenter) -> *Katalysator* -> *catalytic converter* -> meaning definition 3 (-> advice) -> transfer comment -> *catalyst* (-> advice) -> meaning definition -> solution: *katalytischer Konverter*
- * text difficulty: *A/F > 14.7*
 search (S1): *air-fuel ratio* -> equivalent -> transfer comment (no German abbreviation for A/F) (-> advice) -> unit STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS -> *Mischungsverhältnis* -> solution: *Mischungsverhältnis von mehr als 14,7 zu 1* (-> advice) -> unit STÖCHIOMETRISCHES LUFT-KRAFTSTOFF-VERHÄLTNIS -> solution: $\lambda > 1$ (-> advice) -> *Lambda* (-> advice) -> meaning definition (-> advice) -> *Luft-Kraftstoff-Verhältnis* (collocations) (-> advice) -> *air-fuel ratio* (transfer comment) -> solution: A/F can be used in German, too.

Polysemous words such as *catalyst* were often found to raise problems of this sort, confusing users by their multitude of meaning definitions and producing **highly unsystematic and circular search strategies** (e.g. subject 3 looking up both *reduction* and *Katalysator* three times).

It soon became clear that beginners in the field often had **difficulty applying the information** they had read. Subject 1, for example, had problems translating *closed-loop feed-back control system*, although he had read the comment that *closed loop* and *feedback* are synonyms. It seemed that he would have required the more explicit information that the synonyms *feedback* and *closed-loop* may occur tautologously, meaning simply *geregelt*.

Another enormous translation problem was *these catalysts, contained in so-called catalytic converters*, although the TC stated that strictly speaking *catalysts* are part of *catalytic converters*, but that *Katalysator* is mostly used in German to cover both. It seemed that subject 1 would have needed the overexplicit mentioning of the fact that although *Katalysator* is the preferred term, the context might necessitate the use of different terms, in which case *katalytischer Konverter* is more adequate in catalytic converter technology, while *katalytischer Reaktor* is more appropriate in the chemical domain. Users (S1, S6, S7) did not trust the offered solution. Subject 7, when asked why she hesitated to accept the TC's recommendation to use *katalytischer Konverter* and why she still felt she could not take a translation decision, said, laughing: "Well, I don't quite know, if I relied on the Termbank I'd simply write *katalytischer Konverter*" - "Ja weiß ich nicht, wenn ich mich auf die Datenbank hier verlasse, dann schreib' ich einfach *katalytischer Konverter*". Other pieces of information which were put more explicitly in the Termbank, but were wrong in the given context of the text, were, on the other hand, readily accepted as a solution (e.g. *beschleunigen* for *to promote the oxidation* [*Oxidation herbeiführen*], when the collocation *to promote chemical reactions* [*chemische Reaktionen beschleunigen*] had been found (S7), or *Abgasreinigungssystem* for *exhaust system*, when only the entry *exhaust purification system* was retrieved (S6)).

Users accustomed to "one-to-one equations" seem unable to make good use of extended information when they lack a deep comprehension of the matter. This is due to the following reasons:

- (a) Insufficient knowledge often leads to a situation whereby users are easily put off by comprehensive and detailed information, as they do not know where to start and where to stop ("einer, der sich überhaupt nicht auskennt...weiß gar nicht, wo er anfangen und wo er aufhören soll", S6).

- (b) Confronted with unknown issues and strange-looking terms, users are easily confused, not trusting to their common sense and the contextual aids. Thus, our subjects did not dare a literal translation of *subsequent catalytic oxidation reactions*, which was not found in the Termbank, but which should simply have been rendered by *nachfolgende katalytische Oxidationsreaktionen*.

It seems that, depending on their background knowledge, users have all sorts of ideas in mind and come to all kinds of erroneous or misleading associations and interpretations. Often this applies even more to translators who have some domain knowledge than to those who have none at all. Thus, subjects with a knowledge of technical translation (S4, S5) but not of catalytic converter technology had less difficulty translating the text than those (S6, S7) who had some knowledge of catalytic converter technology, but were rather unfamiliar with technical translation. It is, therefore, not only the absolute beginners who need explicit and extensive information, but also the more advanced users.

Search strategies are then, naturally enough, influenced by users' **experience in the subject matter**. If it is not available, this not only brings about misinterpretations, wrong conclusions and uneconomic search paths, but also prevents the user from anticipating the Termbank's organization. Users, for example, who are totally unfamiliar with the field of automotive emission control simply do not know that, when faced with the problem of *ignition timing* or *non-catalyst technology*, they might find the information they are looking for (i.e. the equivalent) under the generic term *emission control* and in the corresponding German encyclopaedic unit ABGASREINIGUNGSVERFAHREN. Thus, except for cases where the link is apparent (e.g. *stoichiometric A/F* -> unit STÖCHIOMETRISCHES-LUFT-KRAFTSTOFF-VERHÄLTNIS), **a lack of subject knowledge prevents them from making full use of the Termbank's navigation possibilities**. This applies not only to navigating from lexical entries to encyclopaedic units, but also to navigating from one lexical entry to another. For example, not knowing that a sense relation exists between *requirements* and *standards*, many users did not look up the *standards* entry and thus did not find the equivalent *Grenzwerte*, which in our case would have been a more appropriate translation for *requirements*.

However, although background knowledge more than anything else helps to find the adequate translation solution, even with the best of (factual and bilingual terminological) knowledge it is not always clear how to put something in the target language. The major and sometimes insurmountable problem in technical translation is how to express something idiomatically, i.e. according to the LSP conventions, the texttype and the text's idiosyncracies. One of our "experts" (S8), for example, gave the following rather discouraging account of the source text

problem mentioned above (*Automobile catalytic converters: Catalysts have been widely used to lower the emissions of CO and HC...These catalysts, contained in so-called catalytic converters in the exhaust system of automobiles...*): "Actually catalyst should have been translated by *katalytisch aktive Schicht/Substanz* [active layer/substance] and catalytic converter by *Katalysator*, but since the text starts off with *catalyst* one should not speak of the *aktive Schicht* in *catalytic converters*. What one could do perhaps is to summarize *these catalysts contained in so-called catalytic converters* to say only *Katalysatoren eingebaut in die Abgasanlage* [built into the exhaust system] for the whole phrase; or else if one already said *Katalysatoren* for *catalysts* one needs a second term to bring in the *catalytic converters*, e.g. *gesamte Katalysatoranlage* [catalytic converter system] and then adapt it to the overall sentence. **In fact, it is quite clear what is meant here, the problem lies in how to put this in German.**" This goes to show that even the combination of encyclopaedic and terminological knowledge does not eliminate all trouble-spots emerging from texts and that it takes extremely detailed knowledge (in the mind or from the Termbank) together with all sorts of decision-taking strategies in order to translate technical texts adequately (e.g. the grasping of the particular technical style of an author: "after a while one realizes the idiosyncratic/individual vocabulary and understands the text better without having to look up too much", ["...hat auch jeder Text sein eigenes Vokabular, wenn du erstmal die Hälfte weg hast, dann weißt du ja schon um was es geht, dann brauchst du auch nicht mehr so oft in die units reinkucken, wenn du dich ein bißchen mit dem Thema auskennst"], S6).

This is mostly due to the arbitrariness and **imponderabilites found in LSP texts**. What our tests demonstrated most strikingly is that due to the very nature of technical communication, i.e. its lack of definiteness, it will always be hard for a termbank to compensate for an extremely clear understanding and a detailed in-depth knowledge of the domain. Nearly untranslatable, since altogether confusing for the lay translator, is a phrase such as *requires that the exhaust A/F be controlled close to the stoichiometric composition*, which would be much more logical the other way round, i.e. *exhaust composition and stoichiometric A/F*. This becomes clear from comparing the translations of subject 6 (*setzt voraus, daß die Motorabgaszusammensetzung nahe dem stöchiometrischen Luft-Kraftstoff-Verhältnis geregelt wird*) with that of our "expert" translator S9 (*setzt die Einhaltung eines engen Regelbereichs um den stöchiometrischen Punkt voraus*). The latter translation, even with the best of termbanks, would not be possible without detailed specialized knowledge.

Another striking example of the intricacies translators are confronted with in technical texts, unless they have at their disposal in-depth factual and terminological knowledge of the domain, is the following evaluation text paragraph, which presented a major obstacle: *The stoichiometric A/F occurs at about 14.7. Three-way catalysts operate in a narrow A/F band*

between 14 and 15. Operation of the three-way catalyst requires that the exhaust A/F be controlled close to the stoichiometric composition. A closed-loop feed-back control system holds the A/F in a narrow region near stoichiometry. Subject S7, although not fully unfamiliar with catalytic converter technology, confessed that she did not know what to do with 14.7, for normally this is expressed in German by $\lambda = 1$, which is a problem here, since the next sentence continues with 14 and 15. She ended up clumsily converting the figures into $\lambda = 0,95$ and $\lambda = 1,02$, an altogether uncommon way of putting it. The result was a significant discrepancy between her translation of the above-mentioned paragraph and that of our "expert" (S9). Subject 7 came up with the following translation: *Von einem stöchiometrischen Mischungsverhältnis spricht man bei einem Luft-Kraftstoff-Verhältnis von $\lambda = 1$. Drei-Wege-Katalysatoren arbeiten im engen Regelbereich zwischen $\lambda = 0,95$ und $\lambda = 1,02$. Voraussetzung für den Betrieb des Drei-Wege-Katalysators ist, daß das dem Motor zugeführte Gemisch im stöchiometrischen Verhältnis steht. Die Lambda-Regelung ermöglicht die Einhaltung des stöchiometrischen Punktes.* Our "expert" translator, on the contrary, knew about the fact that $A/F > 14.7$ is borrowed from the English and that the numbers are simply stated as a quotient. Knowing also what to do with the many collocations, he came up with the following very clear translation: *Das stöchiometrische Luft-Kraftstoff-Verhältnis (A/F) liegt bei 14,7. Dreiwege-Katalysatoren arbeiten in einem engen Streubereich mit $A/F = 14...15$. Der Einsatz von Dreiwege-Katalysatoren setzt die Einhaltung eines engen Regelbereiches (Katalysator-Fenster) um den stöchiometrischen Punkt voraus. Durch ein Regelsystem mit geregelter Gemischbildung ist die Einhaltung eines engen Streubereichs des Luft-Kraftstoff-Verhältnisses um den stöchiometrischen Punkt gewährleistet.*

The lack of in-depth domain knowledge affects not only an - in terms of LSP conventions - adequate translation of texts, as well as the use translators make of a termbank's retrieval possibilities, but also the very handling of a system's functionalities. As a result, problems sometimes even cropped up in specifying a query. Subject 6, for example, typed in *control unit* without wild card in order to retrieve *electronic control unit*, thinking that there might be all sorts of *control units*. The lemmatized compound *electronic control unit* was, therefore, not retrieved. **Manuals** on how to use wild cards and **user instructions** with respect to the overall functioning of the Termbank are indispensable here.

In order to allow even the non-specialist to apply termbank information adequately, it has to be elaborated, organized and represented in a highly explicit and comprehensive way. The repercussions this has for the design of the Termbank will be discussed in the following chapter.

5.3 Search-efficient Representation of Information

What was striking in the tests was not only that the search paths were often circular, but also that the translators very frequently needed some **advice** in order to be able to continue their search. Had it not been for the many hints and instructions given by those in charge of the experiment, the users would not have found many of the solutions they eventually arrived at. Search strategies were clearly biased by the interventions of the "experts", i.e. those who knew the Termbank's information and organization. Search strategies were, therefore, influenced not only by the lack of background knowledge but also by certain shortcomings in the content and design of the Termbank.

5.3.1 Termbank Design

In particular, advice was urgently needed to find certain **collocations** and **compounds**. For example, without help virtually none of the users would have found the collocation *to promote the oxidation* in the *catalyst* entry. The same is true of compounds. In fact, it was seeking solutions for compounds that was considered to be most tiresome. Text difficulties such as *catalyst system* and *subsequent catalytic oxidation reactions* produced rather long search paths, which, however, more often than not led to no results, as in

* *catalyst system* -> *system* -> *oxidation catalyst system* -> *control system* -> *system* -> *emission control system* -> *emission control* (2 of 2) -> transfer comment

The equivalent is not found, (S6).

* *subsequent catalytic oxidation reactions* -> *subsequent** -> *reaction -> *catalytic oxidation** -> *oxidation reaction* -> *nachbehandlung

Note that the equivalent is not found even though the translator (S6) already has a fairly clear idea of what the correct solution might be.

The examples show that after having tried out the various components of a compound, the translators tend to give up unless they hit upon the relevant piece of information **by pure chance**. This happened more frequently than one might have supposed:

* *engine operation* is found by chance in the *catalyst* entry (S6)

* *net oxidizing conditions* is found by chance in the transfer comment in the entry of *oxidizing atmosphere* ("Und dabei hatte ich ja schon *oxidizing atmosphere* zufälligerweise angewählt, im Fenster einfach mal wahllos genommen", S4).

It is extremely irritating and even misleading for the user to find that admittedly valuable information is not well organized, but represented in an untargeted way and, therefore, irretrievable or hit upon purely by chance. ("für Anfänger...muß die Datenbank gut aufgebaut sein...; es ist halt manchmal so wie bei dem *Lexikon der Katalysatortechnik*: man findet vieles, aber nicht immer da, wo man es sucht, sondern entdeckt Sachen eher häufig per Zufall; man sollte sich damit schnell vertraut machen können", S8).

Accessibility

In our Prototype Termbank the above-mentioned problem occurred in the case of the following examples:

- * Unlemmatized compounds, such as *net oxidizing conditions* and *model year cars* were found by accident under collocations.
- * The measurement *gram per mile* is mentioned in the TC of the entry *standard*, but since the text speaks of *requirements* and not of *standards*, this could not be retrieved.
- * The unlemmatized expression *stoichiometric composition* occurs under *exhaust gas composition*, which is not apparent to the user.
- * *(A/F) Region*, a synonym of *(A/F) band*, erroneously occurs under the collocations of *band*, instead of being mentioned as its synonym.
- * The collocation *to control close to the stoichiometric composition* is found in the TC of the entry *control* instead of being one of its collocations.

In order to ensure accessibility, it is important to differentiate between **compounds and collocations**. While all compounds (including their clipped forms) should be lemmatized so that they can be retrieved by typing in the term, collocations, which are sometimes extremely long phrases, need not be lemmatized. Exactly where to draw the line between compounds and collocations requires investigation. In some cases, expressions may have to be stored both in the form of a compound (*emission control - Abgasreinigung*) and in the form a collocation (*to control emissions - Abgase reinigen*, nicht **kontrollieren!*).

Lemmatization of **measurements**, such as *gram per mile*, should be taken into consideration.

Storing particular collocations in **transfer comments**, makes for inaccessibility and leads to long and disorganized TCs.

Structure

The **interdependence of structure and accessibility** will already have become clear from the above remarks. Indeed, the problem of inaccessibility stemmed in large part from the structure. In terms of structure, it was, on the one hand, the **intercategorical organization** that was found wanting. Definitions, for example, should have been placed immediately after the lemma. More importantly, however, the **intracategorical structure** lead to a situation whereby not very clearly structured information was often simply overlooked or else not fully taken in ("manchmal hat mir die Übersicht gefehlt, trotz der unterschiedlichen Farben, jetzt fühl' ich mich relativ erschlagen", S4).

Collocations and their equivalents were often rather chaotically listed in the text-field. Although this was also due to the restricted possibilities provided by MULTITERM, the tests revealed that clearer ways of structuring the collocation field, particularly in view of the great numbers of collocations, have to be examined, so as to present more helpfully the important and highly estimated information (S6) provided by collocations, e.g.

at stoichiometry --> bei $\lambda = 1$

(operate) near # --> (bei) nahezu stöchiometrischem Betrieb

closer to # --> in der Nähe von $\lambda = 1$

A/F ratio close to # --> Mischungsverhältnis in der Nähe von $\lambda = 1$

*(operate) (slightly) lean/ rich of # --> (bei) (leichtem) Sauerstoffüberschuß/
Sauerstoffmangel*

Furthermore, the central aspects (verb, adjective, preposition) of the collocations should be highlighted (S7).

A major drawback of our Prototype Termbank was the organization of the **transfer comments**. In their prototypical state they served as a kind of **hold-all** for all sorts of highly valuable but unstructured information, with the result that the information was not properly retrieved (cf. the example of the TC on A/F described in chap. 4.2.2). Furthermore, it should be seen to it that all information described in detail in a TC should also appear in any other relevant category. A collocation, for example, which together with its equivalent(s) needs treatment by a TC, should not be omitted under collocations.

Concerning the Encyclopaedia, users approved the fact that the **ETs** (encyclopaedic terms) were marked in a different colour so that they could **scan the units** and find a particular aspect without having to read through the sometimes rather long unit-texts (S4).

All in all, it became clear that all **text-field categories** - such as meaning definitions, usage comments, collocations, transfer comments and encyclopaedic units - have to be well organized and clearly structured in order to be given an illustrative layout.

Contents

Finally, drawbacks with regard to the contents of the Termbank sometimes led subjects to draw the wrong conclusions, e.g.

* *stoichiometric composition* -> *stoichio** -> **Zusammensetzung** -> **Abgas** -> no result
> conclusion: *Abgaszusammensetzung* does not exist in German. (S6)

Major impediments were the omission and the incomplete treatment of entries. Users were clearly more upset when they were confronted with the system's message *not found* than when they hit upon somewhat confusing information.

All our subjects, therefore, demanded an **exhaustive treatment** of all possible terms. Subject 8 interestingly said: "In my view, all manner of things should be in the Termbank - the finer points, too, otherwise it's of no use, for you know the central bits", ["Wenn's nach mir geht, müßte alles mögliche drinstehen, auch die Feinheiten, sonst bringt's nichts, wenn nur die groben Züge drin sind, die hat man ja drauf."].

The wish for a **comprehensive, detailed and exhaustive elaboration of terms** is in line with the translator's demand for a quick solution, but there is also another aspect to it. It involves the finer, less obvious aspects that dictionaries do not normally come up with but which give rise to enormous problems. For, unlike manuals, technical texts more often than not describe the general system or concept, with the result that very vague terms crop up in them. Dictionaries (e.g. the Schmitt dictionary), however, prefer to restrict themselves to the concrete bits, objects and components (as appear in manuals). What the tests revealed over and over again was that it is the rather **unspecific terms**, expressions and phrases that were most troublesome (e.g. the vague compounds such as *catalyst systems* or *non-catalyst technology*, cf. chap. 4.1.5.).

This applies in particular to what one might term **LSP fringe words**, i.e. expressions which are neither technical terms nor general language words. These turned out to be most difficult and their solution required an enormous cognitive effort. The statement given by subject 8 is symptomatic of the problem: "It may sound silly, but what I find most difficult to translate in the first sentence is *1975 models*, for it is not really a term. The standard terms are sometimes

very difficult, but terms like this are wholly unfamiliar, so that despite my knowledge of catalytic converter technology, I can't just translate away as I normally do with the more uniformly used standard expressions". Accordingly, these expressions are very unpopular. The same is true of phrases like *widely used* [*serienmäßig eingesetzt*], the idiomatic translation of which is not apparent, but which are not normally treated anywhere (unless taken up as part of some collocations).

The more difficult issue is that of **general language verbs** which in LSP have a different meaning. All our subjects (S1-7) had difficulty with *to foster/ promote/ favor* (*the reduction, oxidation*). The collocations of *reduction*, for example, contained *to catalyze a reduction* [*eine Reaktion beschleunigen*]. This led some subjects (S4) to assume that *to promote a reduction* should also be *beschleunigen*, as this comes close to the general meaning of *to promote*, and seemed to fit the context, which implied some catalyzing activity. Interestingly, however, in LSP such verbs lose their semantic specification. They then simply mean *to bring about*. It remains to be examined whether these verbs should be dealt with under collocations. What has become clear, however, is that these less specific, rather vague and general words, which are doubtless part and parcel of technical texts, frequently reoccurring as they do, constantly give rise to translation problems. They are prominent examples of the lack of precision in LSP texts.

There is also a great number of very **unspecified nouns**, such as *conditions, system, parameters, volume* as in the evaluation text as well as *arrangement, configuration, behaviour, performance, response*, etc. These words are combined rather vaguely. Thus, an English text may mention a *catalyst behaviour*, although what is really meant and has to be expressed in German is *das Konvertierungsverhalten*. Furthermore, such words often give rise to clippings, e.g. *conditions*, which really stands for *operating conditions* and indeed has to be translated by *Betriebsbedingungen*.

Other words which fall into the same category include typical components of technical speech, e.g. *operation/ to operate, to apply, to use* [*zum Einsatz bringen* rather than *verwenden*], or even phrases like *to vary as a function of*. All subjects strongly favoured the integration of all these terms into the Termbank. Our "expert" subjects (S8, S9) even suggested their treatment in the form of an extra alphabetic list somewhere as a supplement to the Termbank. The mode of representation of such **common, vague and rather unspecific semi-technical expressions** should certainly be the subject of further investigation.

By the same token, **adjectives** such as *automotive* and *automobile* are felt to be extremely difficult, yet hardly found in dictionaries. Although they are technical terms, they are not very

specific and occur in certain combinations to form compounds or collocations. Depending on the combination, they need to be differently specified when translated (e.g. as *Auto-*, *Automobil-*, *Kraftfahrzeug-*, or *bei Ottomotoren*). While beginners in the field have problems with both these expressions and the more specific terms, the former constitute the true difficulty for expert translators who are familiar with the standard terms.

All this makes for an exhaustive treatment of a domain in a termbank, which also implies that "borderline cases" should be integrated, i.e. terms that extend into other domains but are not irrelevant to the given field, such as *gasoline engines* [*Ottomotoren*], *exhaust system* [*Abgasanlage*] and other terms from automotive engineering.

The demand for an exhaustive treatment of terms also raises the question as to whether **colloquial expressions** (e.g. *Schadstoffumwandler*, *Abgasreiniger*, *Abgasentgifter* for the *catalytic converter*) should be treated in the Termbank, and, if so, where they should be stored. On the one hand, subjects are flabbergasted to find them as entries, but on the other, they might be required for non-native recipients and should not be left out. The long usage comment of the synonyms of *catalytic converter* could be a possible category to cater for them.

The call for completeness makes **compounds** a particular problem to deal with. One should investigate to what extent each and every compound ought to be made an individual entry. This is certainly the case with *ignition timing* and *gasoline engines* - not so obviously, however, with *subsequent catalytic oxidation reactions* and *exhaust emission*, cases in which it is up to the user to think of an equivalent. *Engine operation*, finally, might as well be taken up under collocations, e.g. in the form of *to operate an engine*.

Again, it is the **compound clippings** that go to show how important it is to have extremely exhaustive entries. If, as was the case in the Prototype Termbank, the compound *emission control* is mentioned under *control* but not the compound *control system*, the user does not know that *control system* may also mean *Abgasreinigungssystem* (clipped meaning). *Control system* was in fact listed under the entry *emission control*, but is, of course, difficult to find there; furthermore, no differentiation was made between *the process* -> *Vorgang der Verbrennung* and *the system* -> *Abgasreinigungssystem/-anlage/-konzept*. *Control* is a complicated concept so that all its synonyms, hyponyms and related terms such as *control system*, *emission control*, *closed loop air fuel ratio control* have to be thoroughly treated. It might be best to take them all up as individual entries, but at the same time to refer to them in all other related entries, possibly in the usage comments' extended usage category.

Last but not least, extremely valuable information was sometimes given, but was **incomplete**, and therefore not helpful, as in the case of the TC of *stoichiometric exhaust mixture*, which draws attention to the improper use of *exhaust A/F* [*stöchiometrische Abgaszusammensetzung*] but fails to make the distinction between *engine-out* and *tail-pipe emissions* [*Motor- und Auspuff-Abgase*].

Functionality

Search strategies were, of course, also affected by the system itself. Problems arose with respect to the "replacement-**software**" (MULTITERM), as the representation was not very appealing and was even, in some cases, disturbing on account of the many technical programming signs that appeared on the screen. Furthermore, the **hardware** was not very supportive as the 286 slowed down the handling of the system. Apart from these software- and hardware-related points, the following suggestions should be observed in order to increase user-friendliness.

Sometimes users had apparent **difficulty working with the Termbank**. What could frequently be observed was that they forgot to enter an asterisk and thus did not find the entries they were looking for, e.g. when *control unit* was typed in, so that *electronic control unit* was not found (S6). **Wild cards** were said to be "a truly supportive retrieval facility" ["eine ganz sinnvolle Suchmöglichkeit"]. Their extensive and flexible use needs to be ensured. (S7). For this purpose, users repeatedly expressed the wish to have **manuals** clarifying the handling of the Termbank as well as the organization and representation of its information, as they found that the Termbank was a "very useful tool, once one knows how to handle it" ["sehr nützliche Sache, mit der man sehr gut arbeiten kann, wenn man eingearbeitet ist", S4].

An important additional feature users demanded was **graphics**, for they felt that "sometimes drawings explained more than two pages of written explications" ["manchmal ist eine Zeichnung wirklich aufschlußreicher, als zwei Seiten Erklärungen", S4].

5.3.2 Linking up Extended Termbank Information

All subjects agreed that the Termbank with its extended information was much more valuable and helpful than the Schmitt dictionary. One of the translators advanced in the field stated: "The Termbank is better than the dictionary on automotive engineering inasmuch as it contains more information; for example, what I had needed to look up two or three terms ago,

I would have found in the Termbank, but not in the dictionary; Schmitt has, of course, a number of useful comments, but the Termbank is generally preferable" ["Datenbank ist besser, als das Katalysator-Wörterbuch, sind mehr Sachen drin, z.B. was ich vor zwei, drei Semestern noch hätte suchen müssen, hätte ich hier gefunden, nicht jedoch im Wörterbuch; ...die allgemeineren Sachen sind bei ihm nicht drin; das ist hier besser, auch einfach mehr Informationen, er hat zwar lobenswerterweise viel Erläuterungen, aber die Datenbank ist generell vorzuziehen", S8]).

Extended Termbank information was found to be indispensable for the following phenomena:

- (a) **bilingual trouble spots**, such as *A/F > 14.7*, including explanations of technical details such as *gram per mile* (S1, S4);
- (b) terms with a **cluster** of synonyms (e.g. *Umwandlung, Reduktion*) or several equivalents (e.g. *reducing agents, ignition timing*), where most users, given their incomplete background knowledge, would not know what choice to make;
- (c) the unspecific use of rather **vaguely defined terms** in texts, e.g. *automobile*, which can be expressed in a variety of ways, as well as *non-catalyst technology* and *net oxidizing conditions*, which lack an equivalent and need a suggested translation;
- (d) terms which seem rather like general language expressions but are sometimes used with a **modified meaning** or are difficult to translate in LSP, in particular semi-LSP verbs (e.g. *to promote*), adjectives (e.g. *lean*) and prepositions (e.g. *via*);
- (e) most **non-standardized**, i.e. undefined and imprecise **terminological expressions** and phrases, which were often found to be more problematic than the standardized technical terms (S7, S8).

However, in order to provide the support which only Extended Termbank information can offer, it must be thoroughly linked up. While no category can function solely without the additional information of the other, it is especially the Extended Termbank categories such as long usage, transfer comment and encyclopaedic units that elaborate the Core Termbank information, and explicitly supplement one another.

The tests revealed a clear **division of labour** between usage notes ("short usage"), usage comments ("long usage") and transfer comments. **Usage notes**, which are part of the Core Termbank, are to be seen in analogy to "short grammar"; they are language-specific and give ISO-standardized information on the usage of terms. **Usage comments** belong to the Extended Termbank and provide elaborated usage information needed to differentiate between the many variants and (quasi-)synonyms of a term, e.g. *standards vs. requirements*. They are also seen as an adequate instrument for providing information, in relation to a given term X, about various non-synonymous terminological expressions either for which X can be

conventionally used in texts as a co-referential substitute, or by which X can be substituted itself. Such information would be of particular importance in the case of compounds and their various "textual" clippings, e.g. *exhaust emission control*, *emission control*, *exhaust control*, *control*. **Transfer comments** are also part of the Extended Termbank, and refer to translational problems of usage regarding the (quasi-)synonymous transfer equivalents and their co-referential substitutes, related to a specific source term, e.g. *Emissionen*, *Abgasemissionen*, *Schadstoffe* and *Schadstoffemissionen*, which are possible transfer equivalents of *emissions*.

While usage notes and usage comments both belong to the unilingual part of the Termbank, transfer comments - together with transfer equivalents - make up the transfer part of the Termbank. (For the distinction between unilingual and transfer information see Albl et al. 1990). In the following discussion of some terminological problems of translation, the differences in function between usage notes, usage comments and transfer comments will be further illustrated.

The treatment of synonyms is clearly one of the major challenges for termbank development; this is especially true because of a common confusion between denotational synonymy as a sense relation and what is sometimes rather misleadingly called "textual synonymy". Non-synonymous terms which, for a variety of textual reasons, are conventionally used to substitute each other in texts to designate the same textual referent create serious translational problems of text comprehension and text production (e.g. *catalyst system* and *control system*), and they certainly have to be dealt with in a termbank. But to put them into the same category as denotational (quasi-)synonyms confuses the matter (cf. Lyons 1977). We propose to treat them separately from (quasi-)synonyms as **co-referential substitutes**. Considering that relations of co-referential substitution are of a textual nature, and that they themselves are based on semantic relations between terms, in particular various grades of hyponymy and meaning relatedness, **usage comments** seem to be an adequate instrument for dealing with them. Such an approach avoids the undesirable result of describing the same terms as (quasi-)synonyms and, say, hyponyms at the same time, and thereby considerably contributes to the consistency and transparency of the Termbank.

Especially the various clippings of compounds, due to their inherent hyponymous nature, are quite extensively used as co-referential substitutes, often resulting in a cohesive crisscross of terms on the textual surface, whose intended meaning is different from their underlying denotational meaning (e.g. *control system* with the textual meaning of *emission control system*). The translational problems involved are outside the scope of conventional termbank information; they can be made explicit, however, by means of appropriate usage comments.

A somewhat special case is *A/F*, which is an abbreviation, i.e. a variant, of *air/fuel* and, at the same time, a clipping of *A/F ratio* or *A/F mixture*. A problem arises in relation to the translation into German. There is a convention in German texts to borrow the acronym *A/F* in those cases where it collocates with numbers (e.g. *A/F*). The latent German collocation **L/K* > 14.7 (*L/K* short for *Luft/Kraftstoff*), however, is not used. An appropriate **transfer comment** is required to explain this aspect of usage since users often know what a phrase is all about without being altogether sure as to how to render it in the target language. This is especially true when numbers and acronyms are involved (S4); with respect to acronyms, there seems to be no consistent strategy of borrowing. In the case of *EGR* (short for *exhaust gas recirculation*), for instance, *EGR* can be borrowed by German while, at the same time, the two German acronyms *AGR* and *AFR* (short for *Abgasrückführung*) can be used.

As has been revealed in the examination of users search strategies (cf. chap. 5.2.), complications may also arise in connection with polysemous and seemingly polysemous phenomena. *Reduction* clearly is an example of **polysemy**; it is a term with two meanings and therefore requires two definitions: (a) *reduction* in the purely chemical sense [*Reduktion*] as an antonym of *oxidation*; (b) *reduction* in the sense of *the reducing of exhaust emissions in automotive emission control* [*Reduzierung / Verminderung / Verringerung*]. In most cases, the two meaning definitions together with short **usage notes** to indicate the domains - [*Chem*] for (a) and [*KatTech*] for (b) - will be sufficient for finding the appropriate translation into German. There are two peculiarities, however, which might also ask for a **transfer comment** explaining that in German texts *Reduktion* is sometimes incorrectly used with meaning (b), and that *Reduzierung* as part of *Abgasnachbehandlung* comprises *Reduktion* as well as *Oxidation*.

In the case of *reducing agents*, we are also dealing with distinct German equivalents: (a) *Reduktionsmittel*, which is only used in the chemical sense, and *reduzierende Komponenten* or *Reduktionsgase*, which in addition implies the state of matter, in the field of catalytic converter technology. This distinction in German does not, however, create polysemy in English. Rather, the English term *reducing agents* has a wider denotative scope than its German equivalents, whereby a relationship of **bilingual divergence** is established between the two languages. Short **usage notes** [*Chem*] and [*KatTech*] would, in this case, probably not be explicit enough; and the user may want to consult a **transfer comment** clarifying the difference. The same type of problem is exemplified by the English *ignition timing* and its two German equivalents *Zündzeitpunkt* and *Zündwinkel*.

In cases such as *control* both aspects, **polysemy** and **bilingual divergence**, combine to increase the difficulty for the translator. *Control* is polysemous in that it has two meanings: (a) *Regelung/Steuerung* [*Regeltechnik*], and (b) *Verringerung, Reinigung, Begrenzung von*

Schadstoffen [KatTech]); at the same time, because of the split between *Regelung* und *Steuerung* in German, the relationship between *control* and its German equivalents in the case of meaning (a) is one of bilingual divergence. Again, short **usage notes** and a **transfer comment** may be required to supplement the meaning definitions.

The English term *catalyst* exemplifies yet another type of terminological difficulty. In its denotational meaning *catalyst* denotes the *catalysing substance* or the *substrate or layer on which the catalysis takes place*. In texts, however, it can also be found to refer to *catalytic converter*, i.e. the *device* built into the exhaust system of a car, or to the overall *catalyst system*. The German term *Katalysator* is often used in very much the same intended textual meanings; it is even more of a general covering term than *catalyst* as it refers to *Konverter* more often than does *catalyst* to *catalytic converter*. Because of its textual nature, this phenomenon should not be subsumed under polysemy; it rather is a case of pseudo-polysemy, which is more adequately referred to as **textual semantic shifting**. The sometimes intricate interaction between denotational meaning and intended meaning in semantic shifting probably calls for interrelated means of explication: **meaning definitions** to explain the functions of *substance*, *substrate*, *device* and *system*; **usage notes** to indicate the chemical or technical character as well as the frequency of the great number of synonyms of these central terms; **usage comments** to explain conventional patterns of semantic shifting underlying the use of *catalyst* and *Katalysator* and the neighbouring terms involved; a **transfer comment** to clarify how this is put in German and how to solve the problems resulting in the process of translation.

5.3.3 User Orientation

In addition to other types of user requirements, it should be borne in mind that different users groups have different needs which clearly depend on the knowledge users already possess.

Our **advanced technical translators** (S8, S9) did not particularly rely on Core Termbank information, with which they felt to be sufficiently acquainted. What they needed instead was a kind of checking tool for "the many things one has heard over and over again, but keeps forgetting" ["man merkt sich halt nicht alles, auch solche Sachen nicht, die man schon tausendmal gehört hat", S8]. Above all, they were interested in rather detailed and specific information about (a) factual technicalities of the domain and (b) terminological subtleties.

(a) Concerning **factual knowledge**, our advanced translators wanted to be reminded of certain important factual aspects in a concise and specific way as a kind of **aide-memoire**; as subject 9 said, a translator familiar with a certain domain is not to be taken for a subject specialist who "when woken up in the middle of the night, has all the details at his fingertips". An

example of the kind of information they were looking for is given in the following passage (taken from Thome 1986, 296) concerning the catalyst's chemical reactions including their formulas: "... carbon monoxide (CO) and hydrocarbons (HC) are converted into the harmless chemical substances carbon dioxide (CO₂) and water (H₂O), and the nitrogen oxides (NO_x) into harmless nitrogen (N₂) and carbon dioxide (CO₂), amounting to the following reactions:

- 1) $2\text{CO} + \text{O}_2 \rightarrow 2\text{CO}_2$
- 2) $4\text{C}_m\text{H}_n + (4m+n)\text{O}_2 \rightarrow 4m\text{CO}_2 + 2n\text{H}_2\text{O}$
- 3) $2\text{NO} + 2\text{CO} \rightarrow 2\text{CO}_2 + \text{N}_2$ ".

Thus, the needs of advanced translators point to certain pieces of factual information which are useful only when assembled in a wider context. The only tool to cater for such a need is the Encyclopaedia.

(b) With respect to **terminological information**, they were interested in subtle terminological points especially in connection with the meaning and usage of vague and undefined terminological expressions (e.g. *model year cars*) and translational aspects of collocation (e.g. *fahren/betreiben* for *to operate a lean engine* or *einsetzen/zum Einsatz kommen/betreiben* [not *verwenden*] for *the use/application of catalysts*). All in all, they wanted to be reminded of "**how the expert would put a given aspect**", S8.

In the case of **beginners and less advanced translators**, the observations were somewhat paradoxical (cf. chap. 5.1). On the one hand, there can be no doubt that these users do need both Core and Extended Termbank information. On the other hand, they are reluctant to work their way through the necessarily extensive and complex network of information. What they wanted was a termbank catering for each and every translational problem by way of short ready-made solutions (cf. chap. 5.1). This clearly asks for a problem-oriented termbank approach: the Termbank should provide comprehensive, detailed and clearly structured information carefully geared towards specific translational problems; the information should be presented in a rather redundant and less complicated way, and it should be supported by transparent navigation procedures.

6

Conclusion

The needs of our translators were clearly determined by the **textual and translational difficulties** they encountered. In searching through the Termbank their attention was therefore entirely focused on the translational decisions they had to take. On the whole, they did not want the Termbank to compensate for their general lack of knowledge (terminological and/or encyclopaedic), but to provide them with problem-oriented information and, if possible, customized solutions.

The exhaustive treatment of "all" terms and terminological expressions, equivalents, synonyms and variants, compounds and collocations conventionally used in technical texts and translations was considered to be of paramount importance. More often than not, the terminological problems our users encountered, and for whose solution they were out to find helpful information, were primarily of a textual nature requiring text-specific explanations. Underlying the textual and translational surface is the whole lot of usage variation between synonyms and variants, co-referential substitution, semantic shifting, compound clipping, and collocation - all different again for the source and the target text. It was especially in connection with these textual terminological problems that usage comments, equivalents and transfer comments, and (to a somewhat lesser extent) encyclopaedic units proved to be helpful; they provided the type of information needed, and were therefore highly valued by our subjects.

The results of our investigation strongly speak in favour of a basically problem-oriented approach. In the context of translation, terminological information is needed whenever the translator is confronted with difficulties. These then have to be solved efficiently and economically, with no time to spare for detours. Terminological information should therefore be prepared with specific terminological problems in mind. The corpus-based elicitation of terms (cf. Ahmad et al. 1990, 12) is thus supplemented by a process of elaboration which takes as its point of departure the textual and/or translational difficulties encountered in LSP texts.

But even adequate, problem-oriented terminological information is only of little value, unless it is supported by a transparent termbank structure, easy access and suggestive easy-to-follow navigation procedures.

On the whole, our subjects liked to work with the TWB-Termbank (despite problems of layout and navigation caused by the MULTITERM design), and found it by far more valuable than conventional LSP dictionaries (including the Schmitt dictionary, with which they were all well acquainted). Moreover, all users acknowledged that without the information provided by the Termbank, in particular the **Extended Termbank information** (including the somewhat more difficult and time-consuming bits), it would not have been possible for them to translate the rather difficult evaluation text.

In our investigation, it became clear that the relevance and usefulness of a termbank has to be judged within the context of the translation attitudes and strategies characteristic of a particular user group. We found that the **search behaviour** of our users was **negatively influenced** by their

- * lack of subject/background knowledge
- * disregard of the context
- * all too faithful belief in ready-made solutions
- * unwillingness to do in-depth queries (despite the many intricate text-and/or domain-specific problems they encountered).

As a consequence, our users did not make optimum use of the information potential and navigation possibilities the Termbank provides. Specific **guidelines on good practice** are definitely needed, which do not only describe the design and organization of the Termbank, but also give advice on how to make best use of it in a translational environment.

REFERENCES

- Ahmad, K., Davies, A., Hughes, M. (1990). *A Methodology for Building Multilingual Termbases and Special-Purpose Lexica*. Esprit Project No. 2315: Translator's Workbench. University of Surrey.
- Albl, M., Kohn, K., Pooth, S., Zabel, R. (1990). *Specification of Terminological Knowledge for Translation Purposes*. Esprit Project No. 2315: Translator's Workbench. University of Heidelberg.
- Albl, M., Kohn, K., Mikasa, H., Patt, C., Zabel, R. (1991). *Conceptual Design of Termbanks for Translation Purposes*. Esprit Project No. 2315: Translator's Workbench. University of Heidelberg.
- Bosch GmbH (Hrsg.) (1985). *Abgastechnik für Ottomotoren - Technische Unterrichtung*. Ohne Verlagsangabe.
- Bosch GmbH (Hrsg.) (201987). *Kraftfahrtechnisches Taschenbuch*. Düsseldorf: VDI-Verlag.
- Kohn, K. (1988). «Fachsprache und Fachübersetzen. Psycholinguistische Dimensionen der Fachsprachenforschung». In: C. Gnutzmann (Hrsg.): *Neue Perspektiven des fachbezogenen Fremdsprachenunterrichts*. Tübingen: Narr, 39-64.
- Kohn, K. (1990). «Translation as Conflict». In: P.H. Nelde (ed.): *Confli(c)t. Proceedings of the International Symposium 'Contact & Confli(c)t', Brussels, 2-4 June 1988*. Association Belge de Linguistique Appliquée: ABLA Papers 14, 105-113.
- Krings, H. P. (1986). *Was in den Köpfen von Übersetzern vorgeht. Eine empirische Untersuchung zur Struktur des Übersetzungsprozesses an fortgeschrittenen Französischlernern*. Tübingen: Narr.
- Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
- Schmitt, P. A. (1986). *Lexikon der Katalysator-technik - Dictionary of Automotive Emission Control*. Wiesbaden: Brandstetter.
- Thome, K. (Hrsg.) (1986). *Wie funktioniert das - Die Technik im Leben von Heute*. Mannheim/Wien/Zürich: Meyers Lexikonverlag.
- Taylor, K.C. (1984). *Automobile Catalytic Converters*. Berlin: Springer.