THE CONTRIBUTION OF TERMINOLOGY TO TEXT ANALYSIS IN SPECIALISED TRANSLATION¹

Maria Teresa Musacchio and Marella Magris² SSLMIT, Università di Trieste

Traditionally, terminological products such as glossaries have been used to find suitable equivalents in the coding stage of translation, i.e. in the reformulation of the message in the target language. Developments in terminology have led to the creation of glossaries and terminological data banks which now often include background information on terms. On these grounds, an experiment has been conducted with fourth-year students of the translation course at the Scuola Superiore di Lingue Moderne per Interpreti e Traduttori of the University of Trieste to ascertain whether terminological products can also be of service in the preliminary stage of translation, i.e. in message decoding through text analysis. A model of terminological text analysis for the purposes of specialised translation was devised and a group of students was asked to adopt it for the experiment, while a control group of students was given a list of translation equivalents and was allowed to use other reference material. Although further tests are needed to confirm the authors' hypotheses, the results of the experiment suggest that terminological products including background information on terms may partially make up for lack of subject-specific knowledge when conducting text analysis in the preliminary stage of translation of highly specialised texts.

1. Introduction

In an entry on terminology theory in the *Encyclopedia of Translation Studies*, Sager (1998: 259) states that the difference between translation and terminology can be summarised as follows:

translators deal with instances of *parole* (i.e. language in use), whereas terminologists may use instances of parole but are essentially concerned with recording facts of *langue* (i.e. language as an abstract system).

¹ This paper was presented at the 2nd Conference on "Training and Career Development in Translating and Interpreting", Universidad Europea de Madrid, 17-20 February 1999.

² Sections 1, 2, 4 and 6 have been written by M.T. Musacchio, Sections 3 and 5 by M. Magris.

This means that terminologists/terminographers look for specialist and terminological information in texts, while translators encode specialist information and conceptual structures into a text (Nuopponen 1996: 1070-1). Therefore, though translators might be thought to need minimal knowledge of the theory and practice of terminology, they often have to "establish identity among concepts, to deal with instances where concepts are similar rather than identical, and to create target language terms for new concepts" (Sager 1998: 259). Moreover, they are frequently asked to deal with texts of varying length on restricted domains they have little or no prior knowledge of, though they might be well acquainted with the wider discipline. In other words, translators may be called upon to perform what Wright and Wright call "ad hoc terminology management", i.e. terminology work for one-off text production – often under considerable time pressure (Wright & Wright 1997: 147). This means that translators,

confronted with a limited perspective on a discipline, must reconstruct enough of the domain in question to understand the source text and produce a reliable target-language version. (Wright & Wright 1997: 149)

In the experiment outlined below, fourth-year students of the translation course at the Scuola Superiore di Lingue Moderne per Interpreti e Traduttori of the University of Trieste found themselves in a similar predicament. They were asked to translate a text on a highly specific legal domain, the protection of minors on the Internet. They had no prior knowledge of the domain in question, though they did have some notion of legal language, legal texts, and legal translation. The aim of the experiment was to ascertain whether and to what extent terminological products can be of service in the preliminary stage of translation, i.e. in message decoding. For the purposes of the experiment Nuopponen's model of extended terminological text analysis was used. In translating, text analysis actually involves

analysing characteristics of concepts, delimiting and defining concepts, determining relationships between the concepts and structuring concept systems, as well as establishing relationships between terms and concepts and thus solving term problems such as equivalence, synonymy and polysemy. (Nuopponen 1996: 1071)

The outcome of all these processes then forms the basis of the following stages in translating, i.e. text encoding in the TL and TT revision.

2. The experiment in translation: general features

For the purposes of the experiment students were asked to translate an approximately 300-word-long text on the protection of minors on the Internet

from English or German into Italian. Though both SL texts fell into the broad category of legal texts, their language set them apart from 'ordinary' legal texts. For obvious reasons, the protection of minors on the Internet is a legal domain which only began to develop in recent years. In this field, attempts have been made both at national and international level to establish whether measures enacted to protect minors in other areas, such as in publishing, mail and telephoning, could be extended to the Internet. Needless to say, discussion on the matter is far from over. As a consequence, the language of this domain is extremely hybrid as it includes terms drawn from Internet jargon, computing and telecommunications, terms designating new legal subjects, and consolidated legal terms which can be used in their traditional meaning or in special senses that only apply to the Internet. This was largely to be expected as protection of minors on the Internet has reached different stages of regulation in different countries, hence its language is far from being normalised or standardised.

For the purposes of the experiment both students translating from German into Italian and those translating from English into Italian were divided into three groups. Each group was assigned different reference materials to help them in text analysis. One group had access to the on-line version of a domain-specific glossary that is now part of TERMit, the term bank under development in the SSLMIT, while the other two groups per language pair acted as control groups and were supplied with reference material ranging from general and specialised dictionaries to encyclopaedic dictionaries and parallel texts.

While features of the traditional reference materials supplied can be easily imagined, and hence need no further explanation, the structure of the term bank TERMit at SSLMIT will be briefly outlined below. Our term bank is being developed using Multiterm '95 Professional Plus software by Trados. It currently includes glossaries on specific subject domains in two or three of the following languages: English, French, German, Italian, Spanish and Portuguese. The glossaries are created by students who decide to do terminographic work for their theses. As they work, so they add to the existing number of entries in the bank.

Apart from an index field indicating the language, each entry term may be illustrated in the bank with one or more of the following fields: Concept field, Context, Definition, Equivalence, Grammar, Lexica, Note, Phraseologism(s), Related Words, Source, Subfield, Synonyms and Variant(s) (text fields), and Category, Morphosyntax, Regional label, Style label, Subject, Synonymy, Type of relation and Usage label (attribute fields). Whenever a new entry is created,

³ Germany appears to be the most advanced European country in this field.

⁴ A certain degree of harmonisation is the outcome of debate at EU-level (see Green Paper on the Protection of Minors and Human Dignity in Audiovisual and Information Services, COM(96)0483).

the program automatically adds administrative data – Creation date, Created by and Entry number.⁵ Our term bank is concept- and translation-oriented. Terms together with their definitions and contexts are chosen from corpora of texts and are usually selected with the help of subject experts. The main criterion for selection of a term as main term is frequency. Lexicographic and terminographic collections are only used to record the presence of terms in authoritative dictionaries, glossaries, etc.

3. The experiment in translation: potential problems and limits

A decision to test the role played by conceptual information in the comprehension of a special language text might be expected to produce easily predictable results. Assuming that the process of comprehension is guided both by the text and by its reader's prior knowledge (cf. Baumann 1996: 96-97 and Weise 1994: 99) and that in given circumstances a translator does not have adequate knowledge, it might be highly predictable that a glossary providing definitions of terms, their positions within the relevant conceptual systems and other information will always be useful. However, the decision to carry out a test of this kind originates from an awareness of the potential limits to the usefulness of terminological information, which will be briefly analysed in this section.

A first limit may lie in a mismatch between the information provided by the glossary and that contained in a particular text. Gerzymisch-Arbogast (1996: 18-19) stressed that the ideal conditions of a one-to-one correspondence between terms and concepts and context-independence of terms are not always fulfilled in texts. On a conceptual level, she singled out the following deviations (Gerzymisch-Arbogast 1996: 172, 184, 196, 207): partial realisation of the meaning of a term in a given context (*Aspektivierung*), narrower meaning of a broad concept (*Terminologisierung*) and use of a term whose meaning explicitly or implicitly deviates from standard use (*autorenspezifische Anwendung*).

Although more than one definition for each term may be included in our glossaries, the terminographer cannot record each and every variation found in the corpus. The possibility of a discrepancy between a term's definition and its actual meaning in a text is particularly high in a field such as the protection of minors on the Internet, which poses new problems and is still rapidly evolving. Mismatches between glossary and text may be even more striking at the level of conceptual relationships, because our terminographers are instructed to include a single representation for each term. So, if several alternatives are available,

⁵ Considerations of space do not allow us to provide an exhaustive description of our term bank. For a more detailed description see Magris (1996).

they are asked to choose the one most experts agree on. A discrepancy between text and glossary, either at the level of definitions or of conceptual relationships, may baffle translators and adversely affect the comprehension process.

When a full match between the conceptual structures of the glossary and those of a particular text does exist, a further consideration is in order. In our glossaries, special attention is paid to logical and partitive concept systems, whereas all other types of conceptual relationships fall under one single category, i.e. 'general'. No explicit reference is therefore made to causal, spatial, temporal relationships or those identified by characteristics such as properties. However, these relationships can play an important role in the process of comprehension.

A further aspect to be taken into consideration is the difference between the information provided by conceptual systems and the more general knowledge related to schemes, frames, etc. According to Hoffmann (1990: 57-58), for instance, a systematic representation of the fundamental objects of a domain and of their relationships provides a static description of a text and does not give its whole meaning, but only its terminological core. On the other hand, frames, schemes and the like allow readers to shift their attention from the objects themselves to the situation they refer to. The role played by this kind of knowledge in the comprehension of specialised texts has been underlined by several authors, such as Jahr (1996) and Weise (1994). Although our glossaries are mainly oriented towards a traditional representation of conceptual systems, some information of the latter kind is at times provided in notes dealing with encyclopaedic knowledge related to a particular term. Through this experiment we wanted to test whether unsystematic notes of this kind may be of any help to the translator. In particular, we wanted to ascertain whether they can assist translators in the production of a coherent target text.

As previously mentioned, glossaries should partly make up for lack of relevant knowledge, thus helping the translator in the comprehension process. However, glossaries are also texts and as such they ought to meet the requirement specified by Nussbaumer (quoted in Jahr 1996: 22), namely that the author should structure his/her text so that the reader is at least able to draw the inferences which are necessary to reconstruct the author's presuppositions. This bears upon two aspects: on one hand, the information content should be adequate to the reader's level of expertise in order to function as a link between his/her prior knowledge and the text; on the other hand, the very structure of a data bank carries particular meanings which should be accessible to the user. Langefors (quoted in Capurro 1986: 50) mentioned two essential questions that have to be addressed prior to the creation of an information system: what kind of information has to be included in order to meet the user's needs, and how the system should be structured. Moreover, according to Capurro, an essential

condition for successful on-line search is that the user be able to master the system (1986: 77). The responsibility for that lies not only with the user, but also with the creator of a data bank, who has to anticipate the user's ability in order to develop a user-friendly system.

The above-mentioned condition was not completely satisfied during the test: the students were allowed to use the data bank, but received no detailed explanation on how to use it, though they had had the opportunity to attend a prior presentation of the data bank. For instance, students may not have remembered the exact meaning of the label "main term" in the data bank: a term chosen among a number of synonyms for its higher frequency. Furthermore, they might not have been able to recall the assumptions on which our representation of conceptual relationships is based. As a consequence, these aspects of the experiment may well have been a source of uncertainty for students.

Last but not least, potential limits do not only concern the conceptual level, but more generally every element in the glossaries, and are associated with the role two factors can play during access to a data bank: noise and silence. Noise refers to every piece of information that does not meet users' needs and makes their search more time-consuming. Our glossaries have a rather complex structure and provide users with information of different kinds, some of which will obviously be non-essential to solve a specific translation problem. This factor can adversely affect the search, at least until users learn to focus on specific parts of the terminological records and to skip unnecessary data. The other factor, silence, refers to every piece of information sought but not found by the translator. Silence obviously concerns every aspect of reference material, including dictionaries and parallel texts; nevertheless, its relevance could be greater in the case of glossaries. Inclusion of terms in a glossary is the result of complex choices made by the terminographer on the basis of term frequency, relevance and degree of specificity. A domain-specific glossary will never cover all the terms to be found in a particular text. Although the texts for the experiment were chosen so as to include a large number of the terms to be found in the glossary, some students were surprised – and probably disappointed – to find that part of their research yielded no results. Indeed, the students' remarks seem to point to the fact that silence interferes with research more than noise.

4. The experiment in translation from English into Italian

The text for the experiment was taken from Chapter 13 (slightly adapted) "Governance of Pornography and Child Pornography on the Global Internet: A Multi-Layered Approach" of *Law and the Internet: Regulating Cyberspace* (Akdeniz 1997). The passage chosen for the experiment was taken from the

Conclusions and provides a broad summary of problems regarding child pornography on the Internet: responsibility of content suppliers, access providers, parents and teachers, and also possible tools, such as rating systems, which can be used to protect children from harmful or illegal content.

Students were divided into three groups: group A consisted of three people. group B and C of two. Group A was given an on-line version of the glossary (Montagna 1997-98), control Group B received a list of terms to be found in the glossary complete with translation equivalents in Italian, a wide range of general and specialised dictionaries and reference material, Group C had the parallel text in Italian Internet e diritto (Torrani & Parise 1998), dictionaries and reference material, but had also access to the Internet. To speed up search on the Internet, Group C was also given a list of relevant sites.

Analysis of the translations supplied by the three groups has shown interesting results. First of all, though no group completed the translation, Group A was able to translate most of the text, Group B translated two paragraphs out of three and Group C only managed to finish the first paragraph. Though possible differences in translation speed must be taken into account, this result seems to indicate that background information contained in the glossary helped students of Group A in message decoding. Moreover, it also confirms that looking for relevant information on the Internet can be extremely timeconsuming even when one is pointed in the right direction, that is, has been given a list of relevant URLs.

Secondly, all groups had problems in decoding and translating the first sentence of the text:

There is a real problem of availability of child pornography on the Internet, as well as that of the availability of sexually explicit material to unsuitable audiences, such as children.

Students found it difficult to transfer sexually explicit material. Group B did not find an equivalent in the list of terms of the glossary and translated the noun group as materiale esplicitamente sessuale; Group C translated it literally as materiale sessualmente esplicito, while Group A came up with materiale con contenuto esplicitamente sessuale, an altogether more satisfactory solution. The equivalent chosen by Group A indicates that, by skimming through the glossary, students had recognised that material and content are technical and not general terms in this domain, as their choice of translation equivalents for harmful and illegal content further confirms.

A third problem arose in connection with the second sentence of the text:

The regulation of potentially harmful content such as pornography on the Internet and regulation of invariably illegal content such as child pornography are different in nature and should not be confused.

Here harmful content and illegal content are domain-specific terms. Moreover, potentially harmful content is a typical collocation. Group B did not identify the first as a specific term and was probably misled by the fact that it is a fairly common expression in English that is semantically redefined in this context, so they translated it as contenuti potenzialmente pericolosi. As to the second term in this sentence, they recognised it as a specific term, but translated it as contenuti assolutamente vietati, because they were only given a list of terms and their translations, from which they could not tell that contenuto vietato is a partial synonym of contenuto illegale or illecito, since it is less frequent and also has a direct equivalent in English (prohibited material). In the allotted time, Group C could not trace suitable equivalents and translated the terms as materia potenzialmente dannosa and materia invariabilmente illegale, which appear general descriptions in the TL rather than proper terms. Finally, Group A not only identified and translated the terms correctly, but also recognised the collocation and used contenuti potenzialmente nocivi and contenuti illegali a tutti gli effetti in their target version.

In general, the three translations from English into Italian show that students were at least partially misled by the apparently 'standard' language of the text that set it apart from traditional legal texts. In the following two sentences they all had problems distinguishing terms from standard language words:

Other types of offensive content, by contrast, are "victimless crimes" and have no proven ill-effects on other people. At the moment, bans or precensorship acts in relation to Internet pornography or sexual content would in any case be unworkable because of the diversity of pornographic sources.

All groups failed to see that the term *crimes*, though used in a collocation between inverted commas (*victimless crimes*), required an equivalent in legal Italian in the target text (*reati che non producono vittime* and not *crimini senza vittime*), that *pre-censorship acts* was a general, albeit concise description and not a term (they used *atti di pre-censura* or similar expressions instead of *provvedimenti di censura preventiva*). Finally, only Group A recognised that, though *unworkable* was not a term, in the context it required a somewhat 'legal' translation and opted for *inefficace*. It is indeed conceivable that the greatest difficulty for students is to find the information they need for their work in a parallel text or in reference material. One might therefore extend to translators what Nuopponen says about terminologists or terminographers:

An experienced terminologist or teminographer finds the information he is searching very quickly from special language texts, even if he is not very familiar with the subject field in question. He can distinguish between terms and non-terms, relating concepts, different classifying

criteria, etc. For a student or any other inexperienced reader of a special language text, it is often more time-consuming to find such information in a text. (Nuopponen 1996: 1069)

5. The experiment in translation from German into Italian

The text for the experiment was taken from the article "Strafrechtliche Verantwortlichkeit für den Datenverkehr in internationalen Computernetzen" by U. Sieber (1996) and slightly adapted. The article was published before the German law regulating the activity of Internet subjects was passed. The passage chosen for the experiment analyses the role of Internet providers with a view to identifying their liability for illegal acts committed through the Internet. Translators were divided into three groups of three persons each. The first group was supplied with an on-line version of the glossary (Montagna 1997-98); the second group received the parallel text in Italian *Internet e diritto* (Torrani & Parise 1998) together with a list of the terms to be found in the glossary complete with translation equivalents in Italian and a number of specialised dictionaries; the third group had a list of terms and their translations and a wide range of specialised dictionaries.

Analysis of the translation supplied by the three groups has shown interesting results. For instance, in the following sentence:

Der für strafbare Inhalte in Computernetzen primär Verantwortliche ist der Content-Provider, also der Urheber der jeweiligen Information. Dabei kann es sich z.B. um den Ersteller einer Datenbank, den Verfasser eines Beitrags oder den Verleger handeln, der Teile seiner Zeitschriftenausgabe im Internet anbieten.

the author uses *Urheber* as a synonym for *Content-Provider* in a way that differs from what can be inferred from the glossary. Here *Urheber* is indicated as a synonym for *Autor*, i.e. "a person who produces content available on the Internet". The glossary also specifies that *Autor* is not synonymous with the person introducing the content in the Net, that is the *Content-Provider*. However, it should be stressed that in German *Urheber* has a wider meaning than *Autor*, as it indicates the originator of a particular document. This allows Sieber to use the two terms as synonyms and ensure the coherence of the text. This is particularly the case in the subsequent sentence, where other terms referring to Internet subjects are listed, including that indicating the publisher. Coherence was not achieved in the translations by the two control groups as they both used *autore*. On the contrary, the group which had access to the glossary recognised that *Urheber* has a somewhat unusual meaning in this text

and got round the problem by paraphrasing the term (il fornitore di contenuto, cioè il soggetto che immette le informazioni in rete).

The following sentence was also a source of problems for students:

Als strafrechtlich (Mit-)Verantwortliche kommen darüber hinaus die Service-Provider in Betracht, die ihren Teilnehmern eine Zugangsmöglichkeit zum Internet verschaffen.

The two control groups interpreted both the verbal group and the use of the parenthesis incorrectly and translated strafrecthlich (Mit-)Verantwortliche as possono ritenersi corresponsabili or vengono considerati corresponsabili. In other words, they did not understand that the possibility to prosecute had not yet been laid down by law. As a result, in this sentence only hypothetical possibilities are discussed. Sieber explains that Internet providers could be liable (verantwortlich) or jointly-liable (mit-verantwortlich) on the basis of a further subdivision of Internet service providers into categories – access providers, hosting providers, etc. – to be found later on in the text. In this case too the group using the glossary was able to trace the necessary information to interpret the sentence correctly and translated accordingly (viene esaminata la responsabilità o corresponsabilità penale...). The information used by this group relates to a scheme rather than to a traditional conceptual structure. This confirms the importance of this type of knowledge when interpreting a specialised text.

In this case, the necessary information consisted in what Weise terms the setting of the problem underlying the article, namely a developing sector whose regulation is just beginning (Weise 1994: 101). No doubt students got an idea of the present conceptual and terminological fuzziness from the encyclopaedic notes and from the terminographer's observations on synonymy and equivalence to be found in the glossary. However, students in this group translated *Service-Provider* with *fornitore di accesso*, a choice which is not in line with the information given in the text. Sieber divides service providers into many categories, including hosting providers and subjects providing access to Internet. So, for Sieber an access provider is a kind of service provider, rather than a synonym.

The students' choice cannot be explained on the basis of the information provided by the glossary, where the two terms are kept clearly separate, although the terminographer mentions that the terms can indicate partially overlapping Internet subjects. In this case the students' prior, non-specialised knowledge may have played a role as they are at home with the computer, they surf the Net, join mailing lists, etc. Therefore, they may have an idea of what an access provider or a service provider is, though their underlying concepts do not necessarily coincide with the legal concepts. Indeed, the fact that access

provider and service provider are largely used as synonyms in the standard language could well have made students reluctant to take in new, i.e. legal, information. Jahr (1996: 40) stresses that existing cognitive structures may compromise acquisition of new knowledge as they may be too deeply ingrained to allow alteration or correction. Though Jahr mainly refers to the case in which a term takes on different meanings in different disciplines, thus misleading the receiver, it is conceivable that such a problem can also arise as a consequence of prior, non-specialised knowledge. In this case, as students were aware that conceptual overlapping between the glossary and the text was not perfect, they may have been unwilling to modify their prior knowledge.

6. Conclusions

As can be seen, students who took part in the experiment were confronted with the problem of matching the terminological information they could get from reference material – the level of *langue* in Sager's words – with the data they found in the English and German STs – the level of parole. To do this, they had to analyse the features of concepts, establish relationships between concepts, and solve problems of synonymy and equivalence. All these activities are part of - though by no means limited to - the preliminary stage of translation, i.e. text analysis. In this process students needed to identify advantages and drawbacks of reference material at their disposal. They also had to be aware of the intrinsically fuzzy boundaries of subject domains and of decoding problems that might arise as a consequence of interference between their prior knowledge and knowledge encoded in the STs. Though further research and more and wider tests are needed before final conclusions can be drawn, the results of the experiment described in this paper indicate that the background information supplied by a glossary in a term bank - if used wisely - can help translators at the text analysis stage especially if and when they have little or no domainspecific knowledge. Thus glossary information can partly make up for, but never completely replace, domain-specific knowledge, in particular, when translators work under time pressure and are forced to perform ad hoc terminology management.

References

Akdeniz Y. (1997) "Governance of pornography and child pornography on the global Internet: a multi-layered approach", in *Law and the Internet: Regulating Cyberspace*, L. Edwards & C. Waelde (eds), Leeds, Hart Publishing, pp. 223-241.

- Baumann K.-D. (1996) "Die Mehrdimensionalität des Fachtextverstehensprozesses. Fachtexsorten und Kognition", in *Multilingualism in Special Communication. Proceedings of the 10th European LSP Symposium* (Vienna 29 Aug.-1 Sept. 1995), G. Budin (ed.), Vienna, Intern. Network for Terminology, pp. 95-112.
- Capurro R. (1986) Hermeneutik der Fachinformation, Freiburg/München, Alber.
- Gerzymisch-Arbogast H. (1996) Termini im Kontext: Verfahren zur Erschliessung und Übersetzung der textspezifischen Bedeutung von fachlichen Ausdrücken, Tübingen, Narr.
- Hoffman L. (1990) "Thesaurus und Fachtext", in *Empfehlung, Standard, Norm.*Beiträge zur Rationalisierung in der Fachkommunikation, L.

 Hoffmann (Hg.), Leipzig, Verlag Enzyklopädie, pp. 56-69.
- Jahr S. (1996) Das Verstehen von Fachtexten: Rezeption, Kognition, Applikation, Tübingen, Narr.
- Magris M. (1996) "Costituzione di una banca dati terminologica presso la SSLMIT di Trieste", in *Rivista internazionale di tecnica della traduzione*, 2, Trieste, Lint, pp. 138-151.
- Montagna A. (1997-98) "La tutela dei minori su Internet: un'indagine terminografica in italiano, inglese e tedesco", Unpublished Thesis, SSLMIT, Università di Trieste.
- Nuopponen A. (1996) "Concept systems and analysis of special language texts. Towards a terminological text analysis method", in *Multilingualism in Special Communication. Proceedings of the 10th European LSP Symposium* (Vienna 29 Aug.-1 Sept. 1995), G. Budin (ed.), Vienna, Intern. Network for Terminology, pp. 1069-1078.
- Sager J. (1998) "Terminology theory" in *Encyclopedia of Translation Studies*, M. Baker (ed.), London, Routledge, pp. 258-262
- Sieber U. (1996) "Strafrechtliche Verantwortlichkeit für den Datenverkehr in internationalen Computernetzen", *Juristenzeitung*, 9, pp. 429ff.
- Torrani O. & Parise S. (1998) Internet e diritto, Milano, Il Sole 24 Ore Libri.
- Weise G. (1994) "Stages in the comprehension of scientific texts", *Fachsprache*, 16, 3-4, pp. 98-105.
- Wright S.E. & Wright L.D. (1997) "Terminology management for technical translation", in *Handbook of Terminology Management*, S.E. Wright and G. Budin (eds), Vol. 1, Amsterdam/Philadelphia, Benjamins, pp. 147-159.