

# THE DISTRIBUTION OF INFORMATION IN LSP TRANSLATION. A CORPUS STUDY OF ITALIAN\*

Maria Teresa Musacchio  
University of Padova/University of Trieste

## Abstract

*In LSP translating the production of naturally-sounding, idiomatic target texts is often taken to mean special attention in the use of terminology. However, research in the distribution of the elements of the message in different languages shows that the representation of the information structure is language-specific and that there are language-specific focus-attributing positions within the syntactic structure of sentences. In this paper, features of information structure and focus-attributing positions will be studied in a corpus of English-Italian translations of popular physics articles and compared with a corpus of original Italian articles on the same topic to see if the Italian information structure is used in translation or whether target texts reproduce the original English information structure. Results will show that translations use – at least to a certain extent – the original English information structure, therefore making target texts sound partly unnatural. A number of revision strategies will be proposed to reinstate a more typical Italian information structure in order to ensure that the informational value of sentence elements is preserved and higher-quality LSP translations are produced.*

## 1. Introduction

Research in LSP translating has shown that in science and technology target texts are expected to sound natural and idiomatic – that is, to leave readers under the impression that they are originals. On closer inspection, the assessment of how natural an LSP translation sounds often rests on an evaluation of quality and consistency of terminology – in particular “extended” or compound terms, subtechnical vocabulary and specialised phraseology. Yet terminology only

---

\* This paper was presented at the 14th European Symposium on Languages for Special Purposes “Communication, Culture, Knowledge” held at the University of Surrey on 18-22 August 2003. At present it features in the provisional on-line Proceedings of the Conference at [www.computing.surrey.ac.uk](http://www.computing.surrey.ac.uk), though it will be published as a brief summary in the final paper version of the Proceedings.

makes up a small part of a technical or scientific text: corpus analysis has shown that in these areas too the most frequent words are mainly closed-class, general-language words. It is clear, then, that the naturalness of LSP translations derives from a blend of different elements, lexical as well as syntactic – not to mention stylistic conventions as to text types. Moreover, if special language texts are considered that are not drafted in English – the international *lingua franca* of science and technology – they reproduce the information structure that is typical of their language. Based on the extent of adherence to target-language specific norms and conventions, researchers have variously concluded that translation gives rise to a “third code” (Frawley 1984); that it has to be studied as a genre of its own (Baker 1993), and that translating as a process is subject to source language interference (Toury 1995).

Pursuing the idea of linguistic features that are typical results of translating, Baker (1993: 243-245) has classified them in six categories – explicitation, simplification, normalisation of grammar, avoidance of ST repetitions, naturalisation (exaggerating features of the target language), and distinctive distribution of lexical items. Subsequent research has identified other possible categories such as natural linear order and structural weight of sentences, (possible) ambiguity, limits to sentence reordering, separating and linking of clauses (Doherty 1997a and 1997b). It can be argued that these classifications are overlapping to a greater or lesser extent. For example, simplification is a way to reduce structural weight, while separating or linking clauses can be regarded as a process of naturalisation and normalisation of grammar reflects attempts to re-create the natural linear order of TL sentences. In this paper, Doherty’s terminology will be preferred because it focuses more on the syntactic and textual aspects of the influence of a source language on a target language.

Considering the nature of the corpus analysed in this paper – popular science articles – another relevant perspective for this kind of investigation is the one offered by Myers (1990: 142) who points out that in scientific journal articles the material is organised hierarchically according to the scientist’s argument and to an arrangement of time in parallel series of events. This creates a *narrative of science*, whereas popular science articles construct a chronological narrative of scientific facts, focus on the observation of nature following the natural history tradition and therefore provide a *narrative of nature*. On a textual level, then, popular science articles – compared to scientific journal articles – show a distinctive organisation, syntax and vocabulary. In *Scientific American*, for example, editorial changes in syntax to fit the conventions of the genre take three forms “(1) rephrasing of introductory statements as questions and answers; (2) rephrasing of compound and complex sentences into several more simple

sentences; and (3) rephrasing passive and impersonal constructions in active voice” (Myers 1990: 175).

## 2. The distribution of information

Linguistic research in the distribution of the elements making up a message in different languages shows that the representation of the information is language-specific and that there are language-specific focus-attributing positions within the syntactic structure of sentences. With reference to the language pair that will be compared and contrasted in this paper – English and Italian – it should be pointed out that in English right-branching clauses are easiest to comprehend. Extensive clause embedding renders sentences awkward and indeed incomprehensible if the clauses are positioned initially in the sentence, where the length and complexity of the clauses contravenes the principle of end-weight (Quirk, Greenbaum, Leech & Svartvik 1985: 49-52). In Italian, the principle of end-focus also applies since the most important information is usually placed at the end of the sentence (Benincà, Salvi & Frison 1991: 121). However, Italian exhibits several structures where constituents are extracted and moved to the initial position of the sentence in a sort of “free zone” (Benincà 1993: 255) in order to reduce structural weight. Further, according to Halliday and Martin (1993: 16) languages have a distinctive way of constructing reality in science; English follows empiricist lines, French rationalist ones. If French is taken to be a typical example of a Romance language, it is even clearer that there is a language-specific distribution of information. In translating, then, awareness of the different distribution of information in SL and TL is important for successful communication as it is “a function of optimal processing conditions, varying according to language specific parameters” (Doherty 1997b: 72). In LSP translating this is even more important as adequate transfer of content is essential.

## 3. Method

This corpus-based investigation was carried out in two stages. In the first stage, features of information structure and focus-attributing positions were studied in a corpus of English-Italian translations of popular physics articles to identify examples where wording did not appear to reflect the natural Italian word order. In the second stage, these features were compared with those of a corpus of original Italian articles on the same topic to see if the Italian information structure was used in translation or whether target texts reproduced the original English information structure. Corpus analysis was partly computer-assisted

using dedicated software – *WordSmith Tools* (Scott & OUP 1998). The corpus was also compared with the contemporary Italian component of a larger corpus of nuclear physics jointly developed by the University of Surrey and the University of Trieste (Ahmad & Musacchio 2003). The aim of the analysis was to trace strategies that can be implemented in LSP translating and translation revision (1) to balance information structure or (2) reduce structural weight of sentences, (3) to improve cohesion and (4) re-instate the prevalent Italian construction of reality along logical lines in popular science. Given the relevance of content in science these aspects will also be studied with a view to detecting possible ambiguities.

#### 4. The corpus

The corpus developed for investigation consists of popular science articles on particle physics published over a 10-year period. Articles were taken from the American monthly *Scientific American* and from the Italian monthly *Le Scienze*. *Scientific American* has a general readership, though many of its readers have some kind of scientific or technical training (Myers 1990: 144). The articles are written by research scientists – by invitation – or by the magazine staff and are close in form to scientific articles. *Le Scienze* started off as the Italian translation of *Scientific American*, but in the last decade it has published both translations from *Scientific American* – which still make up most of the articles in the magazine – and articles written by Italian research scientists and scientific reporters. Its readership is very similar to that of *Scientific American*. There are two components to the corpus: (1) a parallel component made up of 9 English articles published by *Scientific American* from 1993 to 2003 and their Italian translations published by *Le Scienze* over the same period and (2) a comparable component consisting of 9 articles originally written in Italian and published by *Le Scienze* in the same 10-year period.

	<b>English parallel component, 1993-2003</b>	<b>Italian parallel component, 1993-2003</b>	<b>Italian comparable component, 1993-2003</b>
	<i>Scientific American</i>	<i>Le Scienze</i>	<i>Le Scienze</i>
Tokens	40,633	41,414	41,417
Sentences	1,765	1,802	1,516
Sentence length	21.57	22.52	26.21
Paragraphs	172	166	188
Paragraph length	236.24	251.89	220.54

Table 1. A parallel/comparable corpus of popular science articles

The number of articles selected for inclusion in the corpus was dictated by the translations published by *Le Science* on the topic and thus determined corpus size. Corpus size and composition are summarised in Table 1. A list of the articles in the corpus is given in the References.

## 5. Analysis

The corpus was first analysed using *WordSmith Tools*. Statistical information about the three components of the corpus given in Table 1 above reveals what follows:

1. Though the number of articles is the same, Italian translations and original Italian articles tend to be longer than the original English texts. In particular, translations are longer even if in some cases they are abridged versions of the source texts.
2. Sentence length is roughly comparable in the English and Italian parallel sections of the corpus, though sentences are slightly longer in the Italian comparable sections of the corpus. This may point to an influence of English on Italian translations, especially considering that sentence length is as high as 31.51 in the popular science and 35.82 in the secondary-school textbook component of the Italian corpus of nuclear physics developed by the University of Surrey and Trieste (Ahmad and Musacchio 2003).<sup>1</sup>
3. Translated Italian articles tend to reproduce the paragraphing of the English source texts. The shorter number of paragraphs in the parallel Italian component is a consequence of text abridging during the process of translating, revising or editing. On the contrary, original Italian articles alternate short and long paragraphs so that overall number of words per paragraph is lower – 220.54 compared to 236.24 of original English articles.

Myers' remarks (1990: 175) about editorial changes in textual organisation in popular science are also worth investigating in a preliminary stage. As to rephrasing of introductory statements, questions and answers appear in three of the 9 English articles, in two of their translations and in two of the comparable Italian articles. Most articles, though, start with some kind of historical background or background information on the topic. This confirms Myers' view (1990: 188) that research in particle physics is not so amenable to being turned into a narrative of nature as – say – biology. Rephrasing of passive and impersonal constructions in active voice can be gleaned at by running concordances of words designating people involved in research and mentioned in the articles to emphasise the activity of the scientists. In the parallel Italian component there are 104 occurrences of words such as *fisico* (physicist) = 5,

---

1 The total size of these components of the corpus is 185,985 tokens.

*fisici* (physicists) = 52, *teorico* (theorist) = 1, *teorici* (theorists) = 29, *sperimentatore* (experimenter) = 1 and *sperimentatori* (experimenters) = 16. In the comparable Italian component there are 35 occurrences of these words (*fisico*, 3; *fisici*, 29; *sperimentatore*, 2; *sperimentatori*, 1). Clearly, in this case Italian translations are influenced by the source text.

These preliminary data warrant further investigation into popular article organisation and structure in English and Italian.

### 5.1. Information structure

The typical, subject-verb-object (SVO) word order in English is to start with given information and then move on to new information. This order of given-new information or theme-rheme makes it easier for receivers to understand the message. New information is normally to be found at the end of the clause. When an initial element is the focus, it is prominent – this creates emphasis. As in English, the unmarked word order in Italian is SVO and the information structure is based on the information-flow principle of given and new and the principle of end-focus.

To promote an unmarked theme to marked theme, however, Italian has the same options as English – fronting, left and right hand dislocation, clefting, and inversion –, but may resort to them more or less frequently. A simple way to promote the verb to marked theme is subject-verb inversion. In English inversion is a relatively rare phenomenon, especially in academic prose – approximately 500-600 occurrences per million words (Biber, Johansson, Leech, Conrad & Finegan 1999: 926). As a consequence of greater morphological inflection, Italian has a freer sentence structure and inversion is more common. For example, verbs of happening such as *accadere*, *succedere*, *avvenire* usually cause the subject-verb order to be inverted. As can be seen in the following example, the natural Italian word order one expects when *accadere* is used was not re-created in the translation:<sup>2</sup>

Before the advent of the Standard Model, physicists had become used to experiments producing unexpected new particles or other signposts to a new theory almost before the chalk dust had settled on the old one. They have been waiting 30 years for that to happen with the Standard Model. (Kane 2003: 68)

Prima dell'avvento del modello standard i fisici si erano abituati al fatto che gli esperimenti producessero particelle inattese o altri indizi che puntavano verso una nuova teoria, quasi prima che la precedente fosse

---

2 Underlining indicates parts that are relevant to current discussion.

stata completata. Da 30 anni si aspetta che una cosa del genere accada con il modello standard. (Kane 2003: 38)

In other cases the reasons why information flow is impaired in Italian translations are more complex. In the example below the two sentences are joined by the coordinating conjunction *e* (and), but *very well tested* is translated using explicitation in the form of a very long relative clause which sits clumsily in the middle of the sentence and makes information focus more difficult to identify. This form of explicitation is reinforced by the adjunct *correttamente*:

The Standard Model is very well tested. It predicted the existence of the W and Z bosons, the gluon and two of the heavier quarks (the charm and the top quark). (Kane 2003: 73)

Il modello è stato sottoposto a innumerevoli verifiche, che ne hanno immancabilmente confermato la validità, e ha correttamente previsto l'esistenza dei bosoni W e Z, dei gluoni e di due dei quark più pesanti ("charm" e "top"). (Kane 2003: 42)

Considering that Italian tends to move peripheral information to the left, an implicit relative clause at the beginning of the sentence could balance information structure in Italian. Further, the information provided in the relative clause is redundant and can be more concisely expressed by the adjunct *correttamente*:

Sottoposto a innumerevoli verifiche, il modello ha correttamente previsto l'esistenza dei bosoni W e Z, dei gluoni e di due dei quark più pesanti ("charm" e "top").

Similarly, in the following example, the long noun group followed by two relative clauses and used as theme in the original English text forces the translator to resort to parataxis in Italian to simplify the sentence, but information focus – progress made in the construction of the accelerator – is lost as indeed is part of the information (*which will take the major responsibility for constructing the accelerator itself*), probably because it is thought to be redundant in Italian.

This vast and technologically challenging project, coordinated by CERN (the European laboratory for particle physics), which will take the major responsibility for constructing the accelerator itself, is already well under way. (Llewellyn Smith 2000: 71)

Questo progetto, di eccezionale livello tecnologico, è già in fase avanzata di realizzazione ed è coordinato dal CERN, il Laboratorio europeo per la fisica delle particelle. (Llewellyn Smith 2000: 63)

Again, fronting the relative clause would help restore information balance:

Coordinato dal CERN (il Laboratorio europeo per la fisica delle particelle), questo progetto di eccezionale livello tecnologico è già in avanzata fase di realizzazione.

## 5.2. Structural weight

In English, the preferred distribution of elements follows the principle of end-weight: long and complex elements are placed towards the end of the clause so that receivers can decode the message more easily. In Italian, new information comes at the end of the sentence while more peripheral information is moved to the “free zone” on the left. Somewhat conflicting strategies are used in English and Italian and should be handled carefully in translating from one language to the other.

In the following English sentence, the specification *so far to no avail* is added at the end of the sentence according to the principle of end-weight. In Italian the English word order is preserved contravening the idea of end-focus. The problem is compounded by keeping the long adjunct *at their highest-energy colliders* in mid- rather than left-peripheral position and what the comparable corpus confirms to be two unusual collocations in popular physics, *massime energie* and *cercare la presenza di qc.* The comparable corpus further indicates that active forms with subjects such as *experimenters* are not common in Italian popular science articles, where an impersonal form introduced by *si* or a passive is preferred.

Experimenters, however, have searched at their highest-energy colliders for particles predicted by supersymmetry, so far to no avail. (Jolie 2002: 71)

I fisici sperimentali, però, hanno cercato alle massime energie raggiungibili negli acceleratori la presenza di particelle previste dalla supersimmetria, finora senza successo. (Jolie 2002: 47)

*Senza successo* is a literal translation of *to no avail* which is used here instead of the standard equivalent adverb *invano* as in the proposed revision of the translation below:

Pur sfruttando le più elevate energie raggiungibili dagli acceleratori, finora si sono purtroppo cercate invano tracce delle particelle previste dalla supersimmetria.



Another problem is created by clauses between brackets or dashes, which do not take the same position in Italian as they do in English. Preservation of the English structure violates the natural word order of Italian:

Third, CP symmetry – essentially, the symmetry between matter and antimatter – must be violated. (Quinn & Witherell 1998: 79)

Infine, la simmetria CP – essenzialmente la simmetria tra materia e antimateria – deve essere violata. (Quinn & Witherell 1999: 69)

Explication of the link between the main clause and the clause between dashes and subject-verb inversion would help to keep structural weight under control:

Infine deve essere violata la simmetria CP, cioè – in sostanza – la simmetria tra materia e antimateria.

In the following example structural weight in the Italian translation is increased by positioning peripheral information – introduced in English by *although* – in mid-sentence and by the implicit concessive clause where the gerund *rappresentando* refers cataphorically to the subject of the *mentre*-clause, *coppie di quark*. After a concessive clause introduced by *pur* one expects to find the subject – either explicit or implied. In this case, however, the subject is further removed by subject-verb inversion.

Whereas quarks do not exist freely in nature, mesons do – although they are often unstable. (Cline 1994: 45)

Va ricordato che i singoli quark non esistono liberi in natura mentre – pur rappresentando stati spesso instabili – esistono coppie di quark, i mesoni appunto. (Cline 1994: 51)

In Italian the position of the clause between dashes makes the contrast *quarks do not exist – mesons do* less sharp. To reduce structural weight and keep information focus on the contrast, subject-verb inversion can be used:

In natura non esistono quark liberi, mentre sono presenti coppie di quark, cioè i mesoni, anche se spesso rappresentano stati instabili.

Similar problems originate from the position of adverbs or adverbials. In the following Italian translation the adjunct *individualmente* is placed after the operator and before the main verb as *individually* is in English. The Italian equivalent of the adverb, *individualmente*, is ambiguous and *così* followed by a *sebbene*-clause makes for heavy reading of the sentence:

So although charge and parity symmetry are individually broken by neutrinos, in combination their dictates would seem to be obeyed. (Quinn & Witherell 1998: 77-78)

Così, sebbene le simmetrie di parità e di carica siano individualmente violate dai neutrini, sembra che vengano rispettate quando sono applicate in combinazione. (Quinn & Witherell 1999: 68)

In the proposed revision of the translation below *così* at the beginning of the sentence is replaced by *dunque*, a more frequent cohesive device in the comparable Italian component of the corpus (see 5.3. below); *individualmente* is substituted by the unambiguous adjunct *in singoli casi* and *se* takes the place of *quando* as the more frequent subordinating conjunction used in the comparable Italian articles to introduce what are actually hypothetical and not temporal clauses:

Per quanto in singoli casi siano violate dai neutrini, le simmetrie di parità e carica paiono dunque rispettate se si applicano in combinazione.

### 5.3. Cohesion

Compared to English, Italian is known to prefer longer, more complex sentences where complexity is often the result of hypotaxis. As can be seen from Table 1, corpus analysis of sentence length confirms that sentences are longer in the comparable, original Italian component. As to sentence complexity, a comparison of subordinate clauses in the parallel and comparable components of the corpus by looking at occurrences of subordinating conjunctions – such as *perché*, *poiché*, *affinché*, *sebbene*, *anche se*, *se*, *quando*, *mentre*, *senza*, etc. – shows that frequencies are broadly similar. There is, however, a difference in the distribution of some coordinating conjunctions as shown in Table 2.

	Parallel Italian component	Comparable Italian component
dunque	8	27
quindi	27	49
ma	140	97
cioè	10	35
infatti	6	25

Table 2. Occurrences of some coordinating conjunctives

The higher frequencies of *dunque*, *quindi*, *cioè* and *infatti* in the comparable component suggest that in popular physics translations the number of conclusive and explicative cohesive links should be increased. On the contrary, the higher

frequency of *ma* in the parallel component may point to a preference for a different text organisation based on contrast as highlighted by adversative conjunctions and reflected in Italian translations.

According to Halliday and Hasan (1976) cohesion is created by reference, substitution, ellipsis, conjunction and lexical cohesion by collocation or reiteration – repetition, synonymy or near-synonymy, superordination or use of general words. In English science cohesion is often created by reiteration. In Italian reiteration by means of repetition is avoided for stylistic reasons unless non-repetition is a source of ambiguity. In technical and scientific texts it is particularly important to avoid involuntary ambiguities as this impairs adequate text decoding. In the following example, cohesion is achieved in English by noun repetition. In this case repetition is not necessary in Italian and can be replaced by ellipsis:

Little ripples in these fields carry energy and momentum from place to place, and quantum mechanics tells us that these ripples come in bundles, or quanta, that are recognized in the laboratory as elementary particles. For instance, the quantum of the electromagnetic field is a particle known as the photon. (Weinberg 1999)

Piccole increspature in questi campi trasportano energia e impulso da un luogo a un altro, e la meccanica quantistica ci dice che queste increspature arrivano a pacchetti, o quanti, identificati in laboratorio come particelle elementari. Per esempio, il quanto del campo elettromagnetico è una particella ormai ben nota: il fotone. (Weinberg 1999: 50).

In other cases English repetition at the end of a sentence and at the beginning of the following sentence is reproduced in Italian. This does not improve cohesion in Italian while it increases structural weight.

Up and down, strange and charm, and bottom and top are closely related to each other and are paired into “families”. Up and down, for instance, are the two lightest quarks and belong to the first family. In each family one quark has an electric charge of  $2/3$  (up, charm and top), and the other has an electric charge of  $-1/3$  (down, strange and bottom). (Cline 1994: 41)

Le coppie su e giù, strano e incantato, basso e alto, che sono costituite da quark strettamente correlati, rappresentano tre “famiglie”. Su e giù, per esempio, sono i quark più leggeri e appartengono alla prima famiglia. In ogni famiglia uno dei due quark ha carica elettrica pari a  $2/3$  (su, incantato e alto), mentre l'altro ha carica  $-1/3$  (giù, strano e basso). (Cline 1994: 47)

At least one repetition could be avoided in Italian by text re-organisation and ellipsis:

Le coppie su e giù, strano e incantato, basso e alto, che sono costituite da quark strettamente correlati, rappresentano tre “famiglie”: alla prima appartengono per esempio su e giù, i quark più leggeri. In ogni famiglia uno dei due quark ha carica elettrica pari a  $2/3$  (su, incantato e alto), mentre l’altro ha carica  $-1/3$  (giù, strano e basso).

As can be seen, fronting of the anaphoric *alla prima ...* makes the paratactic *sono ... e appartengono* unnecessary.

Another frequent type of cohesion is given by substitution. In the following example, *such a decay* in the English original refers back to a previously mentioned *decay*. Cohesion is further created by conjunction as sentences 2 and 3 of the original text are connected by the conjunction *e* in Italian. However, in these sentences the English SVO structure is closely rendered in Italian:

In late 1993 such a decay was seen at the Cornell electron-positron storage ring. Only a few such events have been detected so far. Calculating the likelihood of this process is quite difficult. (Cline 1994: 46)

Alla fine del 1993 un simile decadimento è stato osservato nell’anello di accumulazione elettroni-positroni di Cornell. Pochi di questi eventi sono stati finora rivelati e calcolare la probabilità di un simile processo è molto difficile. (Cline 1994: 52)

In the example, a slight ambiguity is created by *simile* as a translation of *such*. When *simile* is used as a premodifier it is a synonym of *tale*, another equivalent of *such* in Italian. In the comparable component of the corpus, however, it is always used as a postmodifier meaning *similar to*. Further, in this component of the corpus *simile* has 9 occurrences and *tale* 46, whereas in the parallel component occurrences are 26 and 7 respectively. A concordance of *simile* from the comparable component of the corpus shows how it is used in popular physics:

e adottando un dispositivo sperimentale concettualmente	<b>simile</b>	a quello sopra menzionato, furono in grado di	Realis~1.txt 89
miliardi di kelvin. Si ipotizza che una transizione di fase	<b>simile,</b>	in senso inverso, sia possibile all'interno di alcune	Ilplas~1.txt 80
in alto nella pagina a fronte, la QCD (o qualunque teoria	<b>simile)</b>	permette di scrivere la massa di ciascuno come	Modell~1.txt 48
ad esempio, le proprietà del protone si calcolano in modo	<b>simile</b>	a quelle di un nucleo di $^3\text{He}$ fatto di quark (due u e	Modell~1.txt 45
ha sfidato i fisici nella sua comprensione, ha una intensità	<b>simile</b>	a quella elettromagnetica ed è oggi unificata con	Laviol~1.txt 32

Table 3. A concordance of *simile* from the comparable component of the corpus

Taking all these aspects into consideration leads us to formulate the following alternative translation where cohesion is increased by making the link between sentence 2 and 3 more explicit using *quindi* rather than *e* and placing *finora* in its standard position according to the comparable component of the corpus, i.e. in clause-initial position. Finally, demonstrative reference in the English original (*this process*) is rendered by the periphrasis *un processo del genere* to avoid repetition of *tale* (*tale decadimento/tale processo*):

Alla fine del 1993 si è osservato un tale decadimento nell'anello di accumulazione elettroni-positroni di Cornell. Finora sono stati rivelati pochi di questi eventi, quindi è molto difficile calcolare la probabilità di un processo del genere.

#### 5.4. Construction of reality in Italian popular science

Halliday and Martin (1993: 16) maintain that different languages construct reality in science in distinctive ways. Further investigation in this area will be needed to prove the validity of the idea. If we consider the following sentence taken from the parallel Italian component of the corpus:

Uno dei principali scopi della fisica, forse il primo in assoluto, è quello di comprendere l'affascinante varietà della natura in modo unificato. (Weinberg 1999: 48)

it is clear that the problem is not so much the close rendering of the English SVO word order:

One of the primary goals of physics is to understand the wonderful variety of nature in a unified way. (Weinberg 1999)

The sentence does not sound quite natural in Italian, though it is not an example of a “third code” in Frawley’s sense. The following sentence from the comparable Italian component of the corpus shows how reality is constructed in Italian popular physics and may help to identify what went “wrong” in the translating process:

Uno degli sforzi della fisica moderna è quello di cercare di descrivere queste forze in modo unificato, come aspetti diversi di un'unica interazione. Ma una teoria unificata non c'è ancora (...). (Gruppo Athena 2002: 58)

In Italian then, one does not *comprendere* (understand), but does *descrivere* (describe) in a unified way. Another important clue is provided by the second sentence in the comparable Italian text where *una teoria unificata* (a unified

theory) sums up the concept expressed in the preceding sentence: by the time Italian readers get to the verbal group (*è quello di comprendere*) they have formed expectations as to how the sentence will unfold and an adjunct such as *in modo unificato* is not very likely to take clause-ending position. A more readily understandable Italian version would be as follows:

Uno dei principali scopi della fisica, forse il primo in assoluto, è quello di ricondurre a una teoria unificata/descrivere tramite un'unica teoria l'affascinante varietà della natura.

As can be seen, the problem does not only lie in the choice of words and collocations, but also in the position of the adjunct *in modo unificato*. To sum up, English presents the process as the discovery of unity behind the variety of nature. Italian prefers to look at it as the formulation of a theory that can account for different phenomena in nature.

Another example of the interplay between syntax and semantics in the construction of reality in Italian popular science is the following:

In the 1980s nuclear theorists proposed that superviolent collisions were not necessarily the only way to see supersymmetry; they predicted that a different form of supersymmetry could exist in certain atomic nuclei. (Jolie 2002: 71)

Negli anni ottanta alcuni teorici ipotizzarono che le collisioni ad altissime energie non fossero necessariamente l'unico modo per rivelare la supersimmetria; secondo le loro previsioni, una forma di supersimmetria differente poteva esistere in certi nuclei atomici. (Jolie 2002: 48)

The comparable component of the corpus shows that Italian prefers an impersonal form or a passive to active sentences such as *alcuni teorici ipotizzarono* unless the difference between the work of theoretical and experimental physicists is foregrounded. In other words, in this case Italian takes for granted that in physics a theory can only be put forward by a physicist and the explicitation is considered redundant. Moreover, the anaphoric *secondo le loro previsioni* is ambiguous because in the preceding sentence theorists did not *predict* but *ipotizzarono* (proposed or hypothesised). Finally, the rather clumsy, close rendering of the English SVO suggests that fronting of the adjunct *in certi nuclei atomici* improves information balance and reduces structural weight:

Negli anni ottanta fu avanzata l'ipotesi che le collisioni ad altissime energie non fossero necessariamente l'unico modo per rivelare la supersimmetria; secondo tale ipotesi, in certi nuclei atomici poteva esistere una diversa forma di supersimmetria.

## 6. Conclusions

Analysis of the parallel and comparable components of the corpus suggests that in translation reproducing the natural linear order of the target language implies an awareness and a feeling for the following: information structures; focus-attributing positions of elements in the sentence; structural weight, and knowledge of cohesive strategies and typical ways of constructing reality in Italian in the field and genre.

Illustration of translation problems by contrasting sentences from translated articles with similar sentences in the original Italian articles shows how a comparable corpus can help in the process of translating and revising. In short, results indicate that translating and revising popular physics articles are not just a question of care in the selection of terminology, subtechnical vocabulary and specialised phraseology. They also require consideration of the interplay between syntax and lexis and make it necessary to take into account factors affecting the ways in which clauses are adapted to context or varied for adequate focus and emphasis in the target language.

## References

- Ahmad K. & Musacchio M.T. (2003) "Enrico Fermi and the making of the language of nuclear physics", *Fachsprache* 3-4, pp. 120-140.
- Baker M. (1993) "Corpus linguistics and translation studies – implications and applications", in *Text and Technology. In Honour of John Sinclair*. Ed by M. Baker, G. Francis & E. Tognini-Bonelli, Amsterdam/Philadelphia, Benjamins, pp. 233-250.
- Benincà P. (1993) "Sintassi", in *Introduzione all'italiano contemporaneo*, Vol I. *Le strutture*. A cura di A.A. Sobrero, Roma/Bari, Laterza, pp. 247-290.
- Benincà P., Salvi G.P. & Frison L. (1991) "L'ordine degli elementi della frase e le costruzioni marcate", in *Grande grammatica italiana di consultazione*, vol. I. A cura di L. Renzi & G.P. Salvi, II ed., Bologna, Il Mulino, pp. 115-225.
- Biber D., Johansson S., Leech G., Conrad S. & Finegan E. (1999) *Longman Grammar of Spoken and Written English*, Harlow, Pearson.
- Doherty M. (1997a) "'Acceptability' and language-specific preference in the distribution of information", *Target* 9:1, pp. 1-24.
- Doherty M. (1997b) "Textual garden paths – parametrized obstacles to target language adequate translations", in *Machine Translation and Translation Theory*. Ed. by C. Hauenschild & S. Heizmann, Berlin/New York, Mouton de Gruyter, pp. 69-89.

- Frawley W. (1984) "Prolegomena to a theory of translation", in *Translation: Literary, Linguistic, and Philosophical Perspectives*. Ed. by W. Frawley, London and Toronto, Associated University Presses, pp. 159-175.
- Halliday M.A.K. & Hasan R. (1976) *Cohesion in English*, Harlow, Longman.
- Halliday M.A.K. & Martin J.R. (1993) *Writing Science: Literacy and Discursive Power*, London/Washington, D.C., The Falmer Press.
- Myers G. (1990) "The social construction of popular science: The narrative of science and the narrative of nature", in *Writing Biology. Texts in the Social Construction of Scientific Knowledge*, Madison, Wi./London, The University of Wisconsin Press, pp. 141-192.
- Quirk R., Greenbaum S., Leech G. & Svartvik J. (1985) *A Comprehensive Grammar of the English Language*, Harlow, Longman.
- Scott M. & Oxford University Press (1998) *WordSmith Tools Manual*. Available: <http://www.liv.a.c.uk/~ms2928/wordsmith.htm> [accessed: 2001-2003].
- Toury G. (1995) "Beyond descriptive studies. Towards laws of translational behaviour", in *Descriptive Translation Studies and Beyond*, Amsterdam/Philadelphia, Benjamins, pp. 259-279.

#### The Corpus

##### *Scientific American* articles

- Cline D.B. (1994) "Low-energy ways to observe high-energy phenomena", September, pp. 40-47.
- Close F.E. & Page P.R. (1998) "Glueballs", November, pp. 80-85.
- Duff M.J. (1998) "The theory formerly known as strings", February, pp. 64-69.
- Jolie J. (2002) "Uncovering supersymmetry", July, pp. 70-77.
- Kane G. (2003) "Physics beyond the Standard Model", June, pp. 68-75.
- Llewellyn Smith C. (2000) "The large hadron collider", July, pp. 71-77.
- Nauenberg M., Stroud C. & Yeazell J. (1994) "The classical limit of an atom", June, pp. 44-49.
- Quinn H.R. & Witherell M.S. (1998) "The asymmetry between matter and antimatter", October, pp. 76-81.
- Weinberg S. (1999) "A unified physics by 2050?", December. Available: <http://www.sciam.org> [accessed: July-August 2003].



Italian translations published by *Le Scienze*

- Cline D.B. (1994) “Nuove tecniche per la fisica delle alte energie”, novembre, pp. 46-54.
- Close F.E. & Page P.R. (1999) “Le glueball”, febbraio, pp. 48-53.
- Duff M.J. (1998) “La teoria un tempo chiamata ‘delle corde’ ”, giugno, pp. 90-95.
- Jolie J. (2002) “Rivelare la supersimmetria”, settembre, pp. 46-53.
- Kane G. (2003) “La fisica, oltre il modello standard”, luglio, pp. 38-45.
- Llewellyn Smith, C. (2000) “Il Large Hadron Collider”, settembre, pp. 62-71.
- Nauenberg M., Stroud C. & Yeazell J. (1994) “Il limite classico di un atomo”, agosto, pp. 74-79.
- Quinn H.R. & Witherell M.S. (1999) “L’asimmetria tra materia e antimateria”, febbraio, pp. 66-71.
- Weinberg S. (1999) “Una fisica unificata entro il 2050?”, dicembre, pp. 48-56.

Italian comparable articles published by *Le Scienze*

- Antinori F. & Quercigh E. (2002) “Il plasma di quark e gluoni”, *Le Scienze Quaderni*, n. 125, pp. 90-94.
- Bellettini G. (1997) “La scoperta del quark *top*”, settembre, pp. 32-39.
- Beltrametti E. (1997) “Realismo classico e realtà quantistica”, settembre, pp. 4-11.
- Franzini P. (1997) “La violazione della simmetria *CP*”, *Le Scienze Quaderni*, n. 103, pp. 68-70.
- Franzini P., Franzini P. & Lee-Franzini J. (1999) “La fisica di DAΦNE”, giugno, pp. 60-66.
- Gruppo ATHENA (2002) “La fabbrica degli atomi”, novembre, pp. 56-63.
- Lodi Rizzini E., Venturelli L. & Zenoni A. (1998) “Atomi esotici e antiatomi”, luglio, pp. 40-46.
- Morpurgo G. (1996) “Modelli a quark e QCD”, settembre, pp. 74-80.
- Peruzzi G. (2000) “La costante di Planck compie cento anni”, novembre, pp. 4-9.