A Lexicogrammar approach to checking quality: looking at one or two cases of comparative translation

Christopher Gledhill

1. Introduction

In this paper I take a tried-and-tested methodology in linguistic analysis (the “lexicogrammar approach”) and apply it to a particular problem of translation (a comparison of two equivalent phrases in an English translation of a French text). My purpose in doing this is to raise a number of research questions which I believe should be of importance to anyone in the translation business.

My first question is very general: between two potentially equivalent translations, is it possible to identify which one is best? The assessment of any translation can often be highly subjective, but there appear to be some areas which are even more difficult to ascertain than others. In particular in this paper I examine the traditionally murky category of phraseology. However, I shall attempt to show here that it is possible to evaluate the phraseology of a particular translation in a scientific, almost forensic way, in particular by using corpus-based evidence. By “corpus-based”, I am referring here to the use of computer-held electronic archives of texts, whether texts found on the internet or more specifically texts prepared for linguistic analysis by “tagging”, or marking-up the corpus. In fact, it has now become the standard position of many empirical linguists (Sinclair 1991, Coulthard 1995, Hunston and Francis 2000, Tucker 2006) that no scientific statements about the linguistic features of a text can be based on introspection or gut-feeling alone, but should rather be supported by the meticulous observation and comparison of contextualised examples from a representative corpus of texts. To many, this might sound impractical and time-consuming, but the methodology of corpus linguistics has become fairly widely-accepted in the field of translation studies (Pearson 1996, Sinclair, Payne and Pérez Hernandez 1996, Bowker 1998, Xiao and Yue 2007). Furthermore, in this paper I show that it is feasible to conduct a systematic corpus-based analysis relatively quickly, especially if the
focus is on or one or two local phenomena. It seems to me that the potential benefits of corpus analysis are so great that professionals in the translation industry should at least be familiar with the observational methods and analytical skills involved in this field.

The second, more specific question I would like to raise here follows on from the first: is there a systematic way of assessing the “phraseology” of a particular translation? Traditionally, the study of phraseology has been concerned with identifying and classifying idiomatic expressions, proverbs and other formulaic phrases such as it’s raining cats and dogs, to take a rain check etc. as well as multi-word terms such as acid rain, heavy rain, torrential rain etc. (Moon 1994, Pavel 1994, Howarth 1996). There is now a considerable amount of corpus-based research on phraseology in translation studies and these analysts focus on much more varied phenomena than traditional idioms (Corpas Pastor 2000, Gledhill & Frath 2005a, 2005b, Pecman 2006, Siepmann 2008).

However, the notion of phraseology is still rather vague in the minds of many translators. On the one hand, many analysts use the terms phraseology and style interchangeably. On the other hand, the definitions proposed by many specialists for the terms “phraseological unit” and “collocation” are often so restricted as to make these notions inaccessible or unusable for the purposes of the translator or the student of translation. In this paper, I argue the case for a more general unit of analysis, the “lexicogrammatical pattern”. This term has its origins in Systemic Functional Linguistics (SFL, Halliday 1961, 1992 following Firth 1957). In the following sections I demonstrate how the analysis of lexicogrammatical patterns can be used to settle questions of phraseology in a sample translation. But before looking at some specific examples of this, I set out the basic premises of the lexicogrammar approach in the following section.

2. Lexicogrammar and lexicogrammatical patterns

The term lexicogrammar refers to two distinct but related notions: (1) the typical lexical and grammatical environment of a sign as it is habitually used in naturally occurring texts or “discourse”, and (2) the core stratum of “wording” which, in Michael Halliday’s model of language, mediates between a lower level of “sounding” (graphology / phonology) and a higher level of “meaning” (semantics / discourse). In this paper, I shall be particularly concerned with the first, formal notion of the “lexicogrammatical (LG) pattern”. However, it is worth setting out in this section some of
the basic theoretical assumptions that underlie the Systemic Functional
approach, so that a link can be made between LG patterns on the one
hand and the idea of the lexicogrammar as a fundamental feature of
language.

One of the central tenets of SFL is that lexis (the structured system of
signs that serves to organise the vocabulary of a language) and grammar
(the structured system of choices that serves to organise sequences of signs
into texts) are not different in nature, but rather form a unified stratum
in the language: the lexicogrammar. A further central assumption of
SFL, following Firth (1957), is that no aspect of lexis or grammar can
be properly discussed without reference to its typical context of use (or
“co-text”) that is to say in actual stretches of texts or discourse. It follows
from this that SFL rejects the structuralist view that the abstract system of
language (langue) is independent from language in use or discourse
(parole). Rather, the language system is constantly interacting with and
being shaped by different types of speech event (the “context of situation”) within a community of speakers (the “context of culture”).

Just as lexis and grammar are considered to form a single stratum,
Halliday considers that the lexicogrammar is not a separate system or
“module” apart from semantics, but is rather an underlying component
of the meaning-making system of language. The stratum of semantics is
thus not thought of as an abstract or logical structure, but rather as the
medium through which humans use language to interact in their social
and cultural context. A consequence of this is that the language, and in
particular the lexicogrammar, is structured by the expressive and com-
municative functions it has evolved to convey. Another way of putting
this, following Martin (2001), is to say that everything in language, from
lexical items and grammatical constructions to whole texts, has evolved to
express very specific discourse functions, in the form of situational “regis-
ters” (the lexicogrammatical resources associated with a specific speech
activity, such as impersonal expressions, nominal style, taxonomies of
terms, etc.), as well as “genres” (goal-oriented, culturally specific speech
activities, such as conversation on a scientific topic, exposition in popular
science, narration in a research article, etc.).

I mentioned above that the lexicogrammar approach insists on the
analysis of signs in their co-text, that is to say in actual stretches of text.
This is what Firth originally meant by “collocation”, which refers to the
degree to which the meaning and use of a sign depends on the presence
of other signs in the same stretch of text:
Words must not be treated as if they had isolate meaning and occurred and could be used in free distribution. (Firth 1968b: 18).

The collocation of a word or a ‘piece’ is not to be regarded as mere juxta-position, it is an order of mutual expectancy. The words are mutually expectant and mutually prehended. (Firth 1957: 181).

It follows from this that the main objects of study from an SFL perspective are not individual signs, phraseological units or grammatical constructions, but rather lexicogrammatical (LG) patterns (Stubbs 1995, Hunston & Francis 1998, 2000, Tucker 1998, Legallois & François 2006). I have argued elsewhere (Gledhill 1999, 2008 and Gledhill & Frath 2005a, 2005b) that LG patterns do not correspond to constituents or phrases in traditional grammar, nor do they correspond to idioms in the traditional phraseological sense. Rather, an LG pattern may include as its permanent elements not only lexical items, but also grammatical signs such as functional words (for example, the pronoun and the particle in the pattern it’s {bucketing, chucking, pelting, piddling, throwing} it down, as well as more abstract grammatical morphemes and inflections (the progressive in it BE VERB-ing it down). In the following sections, I shall assume that lexicogrammatical patterns have the following general properties:

– a LG pattern is a predictable but also productive sequence of signs, which as a whole shares a stable, coherent frame of reference,
– a LG pattern can be composed of lexical signs, or more abstract signs, including grammatical morphemes and constructions,
– a LG pattern is composed of permanent “pivotal” signs and a more productive “paradigm”, a feature which allows the pattern to be reformulated and integrated into other patterns and thus into on-going discourse,
– a LG pattern may extend over a long stretch of text, it may be discontinuous and it may or may not be a syntactic constituent or phrase.

This is a formal definition of LG patterns. However, as I shall demonstrate in the following sections, perhaps the most important feature of LG patterns is that, when they are studied in their habitual textual environment, they usually have a very specific discourse function. This can be seen in the degree to which certain LG patterns often only occur in very specialised contexts. I should perhaps also add here that it is the notion of discourse function that makes the LG pattern distinct from similar ideas that have emerged in contemporary linguistics, such as the “construction” (Goldberg 1995). Thus, it is the preoccupation with context (in
the narrow, textual sense, or in the broader situational or social sense) that makes the SFL approach so different to other models of language.

Finally, it should be pointed out here that there has been a considerable amount of recent work in SFL on the notions of evaluation and intervention in translation (for example, Munday 2010). This work traces its origins to the discourse analysis tradition in SFL (Bloor & Bloor 2007) which emphasises a critical approach to authorial stance, ideology, hedging and the other interpersonal features of text. Although the “critical points” of a translation are relevant to the data examined below, I shall restrict myself in this paper to establishing the notions of “phraseology” and “lexicogrammatical patterns” using the descriptive apparatus of systemic functional grammar.

3. Comparing the linguistic features of equivalent texts

In the following sections, I examine a specific problem of comparative translation from the point of view of the lexicogrammar perspective. The particular translation problem encountered here arose as part of a technical translation exercise for first year students on a Master’s course in specialised translation at the University of Lille. The source text (ST) involves four pages in French from the website of Dassault aviation. The original is too long to be reproduced here (the website is given in appendix 1), but I have set out the first seven lines of the ST in Table 1, in addition to two target translations (TT1 and TT2). One of the TTs is the official translation on Dassault aviation’s website, and the other is the one which emerged as the best attempt by my (mostly French-speaking) students. This trainee translation was attempted with no help from dictionaries or internet.

The experienced translator, proofreader or linguist may not need much help in deciding which TT in Table 1 is the trainee translation (TT1), and which is the professional published text (TT2). However, in the following discussion I am going to assume that both TTs are potentially equivalent candidates, and my question shall be: “In the absence of any background information, how can we decide which translation is better?”

Before sizing up the relative quality of TT1 and TT2, let us examine some of the basic differences between each text from a traditional linguistic standpoint. In Table 2 below, I have lined up some examples of the main types of formal difference that can be seen in the first seven lines of the translation. The sequence I have adopted here implies that the differ-
Table 1. Two translations of the Dassault Aviation text

<table>
<thead>
<tr>
<th>Source Text</th>
<th>Target Text 1</th>
<th>Target Text 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. La conception numérique, qui préside à la production du Falcon 7X, s'appuie sur la modélisation de toutes les pièces d'un avion en 3D.</td>
<td>The digital design which is at the heart of the production of the Falcon 7X, is based on modelling all the aircraft parts in 3D.</td>
<td>Digital design, which controls production of the Falcon 7X, is based on 3D modelling of all aircraft parts.</td>
</tr>
<tr>
<td>2. Cette représentation virtuelle a pu voir le jour grâce au logiciel de conception CATIA, développé par Dassault Systèmes.</td>
<td>This virtual representation has been made possible thanks to the CATIA design software, a development of Dassault systems.</td>
<td>This virtual representation was made possible thanks to the CATIA design software developed by Dassault Systèmes.</td>
</tr>
<tr>
<td>3. Grâce à CATIA, la maquette physique disparaît au profit d’une maquette numérique.</td>
<td>Thanks to CATIA, the physical model has been replaced by a digital one.</td>
<td>With CATIA, the [sic] physical mockup is replaced by a digital mockup.</td>
</tr>
<tr>
<td>4. Inaugurée sur le Rafale et le Falcon 2000, la maquette numérique présente une définition 3D complète de l’avion ainsi qu’une gestion de l’appareil pièce à pièce.</td>
<td>First used on the Rafale and the Falcon 2000, the digital model gives a complete 3D definition of the aircraft as well as the part-by-part management of the plane.</td>
<td>First used for the Rafale and Falcon 2000, the digital mockup presents a complete 3D definition of the aircraft and the management of each individual part.</td>
</tr>
<tr>
<td>5. Elle est le référentiel unique du produit dans l’entreprise.</td>
<td>This has become the single reference point for the product in the company.</td>
<td>It is the sole reference for the product within the company.</td>
</tr>
<tr>
<td>6. Cette méthode de conception numérique est aujourd’hui utilisée dans le monde entier par les industries aéronautiques, automobiles, navales etc.</td>
<td>This method of digital design is used today worldwide by industries such as aeronautics, car-manufacturing and ship-building.</td>
<td>Today this digital design method is used across the globe in the aerospace, automobile, ship-building and other industries.</td>
</tr>
<tr>
<td>7. Le développement du programme Falcon 7X a repoussé les limites de la conception numérique en mettant en œuvre un ensemble d’outils informatiques novateurs : le Product Lifecycle Management.</td>
<td>The development of the Falcon 7X program has pushed back the limits of digital design by establishing a suite of innovative computer-based tools: the Product Lifecycle Management.</td>
<td>The development of the Falcon 7X programme pushed the envelope of digital design by using a set of innovative software tools: Product Lifecycle Management.</td>
</tr>
</tbody>
</table>
Table 2. Differences between the two translations, sorted by rank

<table>
<thead>
<tr>
<th>Rank of analysis</th>
<th>Examples and comments</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Text</strong></td>
<td><strong>Examples and comments</strong></td>
</tr>
<tr>
<td>**ST(4) avion . . . appareil &gt; TT1 aircraft . . . plane</td>
<td>TT2 aircraft . . . (0)**</td>
</tr>
<tr>
<td></td>
<td>Issue of textual cohesion: TT1 reformulates a semi-technical term with a non-technical item (<em>plane</em>), whereas TT2 avoids the repetition.</td>
</tr>
<tr>
<td>**ST(3) disparaît &gt; TT1 has been replaced</td>
<td>TT2 was replaced**</td>
</tr>
<tr>
<td></td>
<td>Issue of textual cohesion: the present perfect in TT1 reads like an announcement, while the past in TT2 reads like a narration.</td>
</tr>
<tr>
<td><strong>Syntax</strong></td>
<td><strong>Examples and comments</strong></td>
</tr>
<tr>
<td>**ST(6) est aujourd'hui utilisé dans le monde entire &gt; TT1 is used today worldwide</td>
<td>TT2 is used today . . . across the globe**</td>
</tr>
<tr>
<td></td>
<td>Issue of style or syntax: the two adverbs in TT1 clash in scope, which leads to clumsy style and potentially misleading syntax.</td>
</tr>
<tr>
<td><strong>Lexis</strong></td>
<td><strong>Examples and comments</strong></td>
</tr>
<tr>
<td>**ST(3) maquette &gt; TT1 model</td>
<td>TT2 mockup**</td>
</tr>
<tr>
<td></td>
<td>Issue of terminology: TT1 gives the standard dictionary translation of <em>maquette</em>, while TT2 uses the accepted industry term.</td>
</tr>
<tr>
<td>**ST(6) industries aéronautiques &gt; TT1 aeronautics</td>
<td>TT2 aerospace industries**</td>
</tr>
<tr>
<td></td>
<td>Issue of equivalence: does TT2 over or under-translate the ST?</td>
</tr>
<tr>
<td>**ST(6) industries . . . automobiles, navales etc. &gt; TT1 car-manufacturing and ship-building</td>
<td>TT2 automobile, ship-building and other industries.**</td>
</tr>
<tr>
<td></td>
<td>Issue of omission: the inclusive <em>etc</em> in ST is not translated in TT1.</td>
</tr>
<tr>
<td><strong>Spelling</strong></td>
<td><strong>Examples and comments</strong></td>
</tr>
<tr>
<td>**ST(1) Dassault Systèmes &gt; TT1 Dassault systems</td>
<td>TT2 Dassault Systèmes**</td>
</tr>
<tr>
<td></td>
<td>Issue of translation coherence: with which spelling and in which language is the company to be referred to in the TL?</td>
</tr>
</tbody>
</table>
ences between the texts rise from small localised points (Spelling), to large
global differences that can only be gauged at the level of the whole text
(Text).

Table 2 is clearly not an exhaustive listing of all the possible differences
between ST and TT. However, it does set out a reasonable range of cate-
gories, which are essentially concerned with one particular problem at
each linguistic rank. Thus at each level, the problem can be stated as
follows: how does a particular sign (form or word) fit in with the general
spelling conventions of the text (UK or US orthography?), the grammati-
cal patterning of tense and aspect in the text (preterite or perfect mor-
phology?), the terminological preferences of the text (general or specialised
lexis?), and so on. To a certain extent, many of these differences range
across several levels at once. For example, whether to translate into the
past or the perfect (line 3). In this case, the problem is at once grammati-
cal (morphological) and textual (discourse / register). From a general
linguistic point of view, it would be hard to judge which of these texts
is the better translation. From my comments in Table 2 above, it might
appear that TT1 is worse than TT2 in some of these categories. But, there
are some instances such as grammar (the use of articles, for example) and,
as I demonstrate below, lexicogrammar, in which it could be argued that
the professional translation fares no better than that of my students. How-
ever, before looking at this notion more closely, it is necessary to examine
some examples which do not fit very well into the categories presented in
Table 2.

4. Comparing the phraseology of equivalent texts

I shall begin this section by examining phraseology from a fairly tradi-
tional perspective, and shall go on to examine lexicogrammatical patterns
in the next section. The reason for starting with traditional phraseology is
that for most analysts, the term is fairly intuitive: it refers to figures of
speech and “dead metaphors” which are fairly easy to spot in context.
For example, most analysts would agree that in ST(2) the expression *voir
le jour*, literally “see the day”, is a formulaic phrase, which stands in
contrast to the more prosaic translations in TT1 and TT2 (*has been made
possible / was made possible*). In Table 3 below, I have set out some of the
more striking examples of phraseological differences between the ST, TT1
and TT2. It is important to state here that, by definition, problems of
phraseology always involve various levels of analysis at the same time.
Thus, they often involve a choice of lexical items in combination (Lexis), choice of lexical and grammatical constructions (Syntax) and rhetorical contrasts such as “formal / informal” (Register).

I think that most analysts would agree that the issues described in Table 3 are much more complex than in Table 2. For example, there are asymmetries of expression between the ST and the TTs, which make it

**Table 3. A sample of phraseological differences between the ST and target texts.**

<table>
<thead>
<tr>
<th>Phraseology + Register</th>
<th>ST(3) grâce à CATIA &gt; TT1 thanks to CATIA / TT2 with CATIA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The choice of complex preposition depends on the degree of</td>
</tr>
<tr>
<td></td>
<td>register variation which would be acceptable in the TT:</td>
</tr>
<tr>
<td></td>
<td>TT1 is informal, TT2 is neutral.</td>
</tr>
<tr>
<td></td>
<td>ST(1) prêside à &gt; TT1 is at the heart of / TT2 controls</td>
</tr>
<tr>
<td></td>
<td>The choice of expression in TT depends on the degree of</td>
</tr>
<tr>
<td></td>
<td>metaphorical reformulation acceptable in TT. TT1 is more</td>
</tr>
<tr>
<td></td>
<td>elaborate, TT2 is more prosaic.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phraseology + Syntax</th>
<th>ST(3) méthode de conception numérique &gt; TT1 method of digital design / TT2 digital design method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The choice in TT is dependent on style (limits on the degree of post-modification by of in English), as well as existing terminology.</td>
</tr>
<tr>
<td></td>
<td>ST(4) présente une définition &gt; TT1 gives a definition / TT2 presents a definition</td>
</tr>
<tr>
<td></td>
<td>(This example is analysed in detail below).</td>
</tr>
<tr>
<td></td>
<td>ST(7) repoussé les limites de la conception numérique &gt; TT1 pushed back the limits of digital design / TT2 pushed the envelope of digital design</td>
</tr>
<tr>
<td></td>
<td>(This example is analysed in detail below).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phraseology + Lexis</th>
<th>ST(5) référentiel unique &gt; TT1 single reference point / TT2 sole reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The choice of term is determined in TT partly by the ambient phraseology of English in TT1, or by the specific of the company phraseology in in TT2.</td>
</tr>
<tr>
<td></td>
<td>ST(7) mettre en œuvre &gt; TT1 establishing a set / TT2 using a set</td>
</tr>
<tr>
<td></td>
<td>The complex verb in ST cannot be translated directly: the choice of verb in TT is determined by the general lexico-grammar of English.</td>
</tr>
</tbody>
</table>
difficult not only to categorise but also to identify the phraseological units in the first place. It seems to me that the best candidate for a phraseological unit in the first seven lines of the translation is TT2 *pushed the envelope*. Most phraseologists would agree that this is a clear case of an idiomatic expression. Technically speaking, the equivalents in ST(7) *repoussé les limites* and TT1 *pushed back the limits* are not idioms but relatively productive lexicogrammatical patterns in French and English. However the fact that a construction in the ST is translated by an expression in one TT is not evidence of better quality. Without any further evidence, who could say which is the better translation here, TT1 (*pushed back the limits*) or TT2 (*pushed the envelope*)? A similar point can be made with the translation of ST(4) *présenter une définition*. Without evidence, even a native speaker would be hard put to say whether TT1 *gives a definition* is better or worse than TT2 *presents a definition*.

This point leads me to the main analysis to be conducted here: in the following sections, I examine the lexicogrammatical differences between these two constructions in order to see which (if any) of the target translations is “better” (or more natural, more appropriate, etc.). As an aside, I should point out here that it is no accident that I have chosen the sequences *push + envelope / push + limits* and *give + definition / present + definition*: these are all good examples of what I have previously called Verb-Noun expressions (Gledhill 2008). They constitute a well-known category of lexicogrammatical pattern, which involves a verb with a very general sense (*do, make, give, present*) and noun which specifies the semantic “range” of the process expressed in the whole construction (*do good, make sense, give a summary, present an argument* etc.).

5. Comparing the lexicogrammar of equivalent texts

As pointed out above, the notion of lexicogrammar as defined by Halliday (1961, 1991) is much broader in scope than the traditional notion of phraseology. The term refers to an integrated linguistic level of wording in which there is no distinction in principle between grammar on the one hand, and lexis on the other. However, it is also noticeable that when Halliday and others talk about the lexicogrammar of a particular text (as opposed to the language system as a whole), they are referring to the linguistic properties of a particular text type or register. For example, Banks (1994) uses statistics to show that, in comparison with other texts in English, texts written in scientific English have a higher relative per-
percentage of (1) modal verbs (associated with “academic hedging”), (2) the passive (associated with impersonal discourse), and (3) embedded clauses (associated with expository, definitional discourse). If we look at the professional translation TT2, a similar set of properties can be observed (all of the following examples are from TT2):

1. Active verbs in ST are translated by TT passives (note the presence of modals and the mention of necessity / possibility in the ST and TT, here emphasised in bold):

   la maquette numérique *se visualise* en 3D > The digital mockup *can be* viewed in 3D. Les outils du PLM *permettent* la constitution d’une base de données > PLM tools *can be* used to create a configuration management database. il *convient* d’optimiser l’industrialisation du produit > [the] industrialisation process *must be* optimised, . . .

2. Nominalisations in ST are translated by TT clauses (note that extra lexical material is used in the TTs, here in bold):

   Outre l’intégration des métiers, > As well as *integrating* the different skill areas... Avec le PLM, des *optimisations de plus en plus avancées* sont possibles > PLM makes it possible to *achieve ever greater degrees of optimisation*. . . . Le dessin industriel facilite la compréhension d’un concept technique ou d’un produit en normalisant *sa représentation*. > Industrial drawings facilitate the understanding of technical concepts or products by standardising *the way in which they are represented*.

3. Thematic refocussing in ST is ignored or downplayed in TT:

   *Elle* offre une visualisation et donc une conception très précises > The smallest details can be visualised and taken into account in the design. *En effet*, un avion, *c’est* 100 000 pièces mais aussi 25 km de câbles > An aircraft is made up of 100,000 parts, as well as 25 km of cable, *En effet*, la maquette numérique en 3D améliore la détection des erreurs > The 3D digital mockup improves error detection. . . .

   It would be interesting to examine to what extent these are regular and recurring features of French and English in general, or differences that emerge in technical writing of this type. However, for the purposes of this paper, it is enough to suggest that if a non-professional translation does not display these properties, it may be less likely to be convincing. In other words, a successful technical translation from French into English: (1) uses
passives to translate active verbs or nouns (especially when a material process and modality is involved), (2) uses clauses and gerunds to translate lexical verbs and nominals (this may allow for lexical expansions and thus some degree of compensation for elements not translated elsewhere), (3) does not use the thematic devices that are commonly used in French expository texts (this may be compensated by other cohesive devices, such as lexical repetition).

In the preceding comments I have examined a sample of the lexico-grammatical properties of one of the translation texts (TT2). In the following sections, I look more closely at the lexicogrammar of two particular phrases. The point of this analysis is to show that constructions such as push back the limits / push the envelope or give a definition / present a definition are not exact equivalents, and are not used in random distribution. In addition, I shall demonstrate below that it is in fact possible to show that one formulation is “better than” or at least more appropriate to the general lexicogrammar of the English language.

5.1. push back the limits or push the envelope?

Which is the better translation: push back the limits or push the envelope?

We have seen that both of these sequences were proposed for the French ST(7) Le développement du programme Falcon 7X a repoussé les limites de la conception numérique en mettant en oeuvre un ensemble d’outils informatiques novateurs... From a traditional, phraseological point of view, push back the limits is not an idiomatic expression, but simply a productive construction that exists alongside other formulations such as push back the boundaries / push back the frontiers, etc. As we shall see below, the lexicogrammar of this construction in English usually involves a more extended stretch of text than simply the words push (back) the limits: typically the noun in the construction is post-modified (such as push back the limits of knowledge). This is the form used in TT1: ... program has pushed back the limits of digital design. In contrast, the sequence push the envelope used in TT2 is an idiom. Idiomatic expressions are usually defined as being relatively “unproductive” from a structural point of view (they cannot be reformulated by analogy with other constructions, e.g. *push the package, *push the envelopes, *this is the envelope that has been pushed, etc.) and “unpredictable” in terms of sense (they cannot be interpreted by analogy with other constructions). Thus, without any other contextual information, it would generally be difficult for a language learner to guess that push the envelope means “go beyond established
limits, innovate, pioneer” (Shorter Oxford English Dictionary, SOED.) In
fact, the origins of push the envelope are highly relevant to our translation:
the expression emerges in the middle of the 20th century in American
English, in particular in the discourse of aerospace and aeronautical engi-
neering as a contraction of a longer sequence push the flight envelope
(SOED: “originally aviation slang, relating to graphs of aerodynamic per-
formance”). From a phraseological perspective, then, the sequence push
the envelope would also appear to be an appropriate translation, and this
is why we find it in TT2: The development of the Falcon 7X programme
pushed the envelope of digital design.

So far I have discussed the phraseology of these two sequences out
of context (or rather, out of “co-text”). However, a lexicogrammatical
analysis of these sequences requires empirical data. The first step in this
type of analysis necessarily involves close attention to the lexical and
grammatical environment of a sign or sequence across a range of exam-
pies. This can be seen in the concordances set out below. In order to better
visualise these examples, I have limited the presentation to five clear and
typical occurrences of each of the main patterns that can be observed for
each form. In the following lists, all the examples are authentic extracts
from the web: the sources for each numbered extract (1–45) are supplied
at the end of this paper, in appendix 2. It is perhaps worth adding at
this point that, strictly speaking, this type of analysis should be based on
a representative corpus of texts, such as the British National Corpus. This
is the method I have used in the past (see Gledhill and Frath 2005a, 2005b
for example). However, the BNC is not always the best resource, espe-
cially for the analysis of very specialised varieties of language. As we shall
see below, the types of sequence that we are looking for are so specific to
technical English that the numbers involved may in fact be rather insigni-
ficant. For example, the expression push the envelope does occur in the
BNC, but only twice, and the construction push (back) the limits occurs
only nine times in the BNC. Such low frequencies are not enough to study
the typical LG patterns associated with these constructions. Thus it seems
to me that a careful preliminary analysis of the Internet using an on-line
concordancer such as Webconc or Webascorpus is useful and at times
necessary to examine the behaviour of very specific sequences.1

1. “Webcorp” is an on-line search engine available at:
   http://www.webcorp.org.uk/. “Web as corpus” has recently changed to
My first set of observations concern the sequence push back the limits. This is a very productive lexicogrammatical pattern in English, and it is possible to find over one million hits even for the specific sequences such as pushed back the limits of the possible on the Web. The pattern push(ed, es, ing) (back) the limits often involves the particle back (but not obligatorily, as we see below), and typically involves post-modification of the complement (limits + of + N) as well as, in some cases, an expansion clause (of the form by + V+ing). In terms of meaning, the pattern typically refers to the abstract “reduction” or “exploration” of an abstract noun expressing knowledge as an explorable space (the known, the unknown), or a “do-able” action (the possible, the impossible) as in the following examples:

(1) Tawhid Abdullah, Managing Director of Damas commented: “The new Harmony designs have pushed back the limits of the possible and have explored new venues.

(2) His show has pushed back the limits of what is acceptable on TV. American talk shows encourage their contestants to get violent.

(3) He explains the development of the sun-centered model of the universe in Renaissance Europe. He then tells how the development of the telescope, photography, and spectroscopy pushed back the limits of the observable universe and eventually brought astronomy into the twentieth century.

(4) At each major premiere, the adventurers of the last century pushed back the limits of the impossible.

(5) “In all these dreams of the politicians and merchants, sailors and geographers, who pushed back the limits of the unknown world, there is the same glitter of gold and precious stones, the same odour of far-fetched spices.”

A particularity of push back the limits is that it is not generally modified by an expansion clause in by V-ing. However, the sequence push the limits, without the particle, is generally used with by V-ing, as can be seen in the following examples (6–10). One reason for this difference may be that the semantic “extent” or “scope” of the spatial metaphor in push back the limits is expressed by back, whereas this particle may be felt to be redundant in the presence of a by-clause. In terms of meaning, most of these examples refer to “breaking” a physical limit (7, 10) or some form of social code (6, 8, 9):
(6) I knew I had pushed the limits by being out late, alone and female.

(7) On the fourth and final day Nuna again pushed the limits by driving at a top speed of 110 km per hour, finally setting a new world record.

(8) He pushed the limits by getting the contestants to bend to his every whim.

(9) Frank Abagnale Jr., notorious con-artist, forger and impersonator, pushed the limits by manipulating other people’s money into his own pockets.

(10) TechArt also decided that the 530 hp and 505 lb-ft of torque would not do it for them and have pushed the limits by reaching 700 hp!

The fact that we do not usually find expansion clauses with push back the limits (but we do find such clauses in 6–10) suggests that the formulation in the translation TT1 may be rather “heavy” or unnatural (pushed back the limits of digital design by establishing a suite of innovative computer-based tools: the Product Lifestyle Management).

Let us now examine the LG patterns associated with the sequence pushed the envelope. The complement envelope is not always post-modified, but in those examples in which is is, the qualifying phrase expresses a new creative genre or cultural activity:

(11) It was not the grandfather of all first person shooters, but when the first Doom was released by Id in 1993 it pushed the envelope of this brand spankin’ new genre.

(12) After he pushed the envelope of computer-generated special effects in The Abyss, director James Cameron turned this hotly anticipated sequel to Terminator into a well-written, action-packed showcase for advanced special effects and for one of the most invincible villains ever imagined.

(13) Radiohead pushed the envelope of modern rock, and this album was their strongest melding of convention and adventurousness.

(14) Each creation pushed the envelope of design and displayed thorough attention toward wear ability and technique.

(15) Artists such as Prince, Sting and Bjork, have pushed the envelope of creativity for years. But artists of their caliber who possess such sublime talent and . . .

As with push the limits (but not push back the limits), push the envelope can also be modified by an expansion clause in by V-ing. In these exam-
ple, the complement typically expresses a new technical process. This is very close to our French ST and its translation in TT2 (The development of the Falcon 7X programme pushed the envelope of digital design by using a set of innovative software tools). It is also significant that the most frequent verb involved in this construction is using, which is the same as the one used in TT2 (to have some perspective on this, it is possible to find over 3600 examples of pushed the envelope by using on the Web, but only 9 examples of pushed the envelope by being):

(16) This team has pushed the envelope by using existing commercial materials combined with cutting edge technology.

(17) In recent years, laparoscopic surgeons have pushed the envelope by using minimally invasive approaches for increasingly sophisticated procedures.

(18) To become “Internet Famous, according to Brett is to be an initiator, not a follower. For example, when social media first came on to the scene, those that became the most successful in blogging and developing apps were those that pushed the envelope, by using new tools and searching out new terrain to maneuver.

(19) In example after example, the Obama administration has pushed the envelope by using the Internet to give citizens a view of the inner workings of the government.

(20) Hence, creative effectiveness was of paramount importance, and we pushed the envelope by using the strategic combination of new-age marketing tools. And there is no better feeling than to get awarded a Global Effie for it.”

The analysis I have set out above appears to be fairly conclusive. Both push (back) the limits and push the envelope refer to a similar idea. Nevertheless, they also have very different contexts of use. The construction push (back) the limits tends to be used in reference to overcoming or breaking intellectual, physical or social boundaries. The expression push the envelope is used in the context of artistic and technical innovation. When push the limits is used with an expansion clause in by + V-ing, it tends not to be used with a particle (back) and the construction refers to a specific event, but not a method. Only the sequence push the envelope is regularly used with an expansion clause that expresses the means by which a process of innovation is accomplished.

The problem faced by the translators in this case is that the ST phrase involves a post-modified N which specifies the domain of the innovation
les limites de la conception numérique) and an expansion clause which specifies the means by which the innovation is achieved (en mettant en œuvre...). It would appear that only TT2 addresses both of these features of the ST phrase in a way that is close to the ambient phraseology of English in this particular domain.

5.2. **Give a definition or present a definition?**

In the previous section, it might have been predicted that an idiomatic expression with its origins in aeronautical discourse appears to be more appropriate to the translation of our ST. However, as I have shown above, in the absence of background information of this type, it is necessary to examine the lexicogrammar of both candidate translations in order to make sure. I would argue that this method is even more relevant to the pair of sequences TT1 give a definition and TT2 present a definition. As highly productive constructions in English, it is virtually impossible to say, intuitively and out of context, which one best corresponds to idiomatic English, or best corresponds to the French original, ST(4) Inaugurée sur le Rafale et le Falcon 2000, la maquette numérique présente une définition 3D complète de l’avion ainsi qu’une gestion de l’appareil pièce à pièce. Unfortunately, space precludes me from examining all of the different grammatical permutations that might be observed for give a definition / present a definition. For example, one construction may be used more frequently in the passive, or in non-finite clauses, and so on. For demonstration purposes, I shall simply limit my observations to those sequences which are closest to the active form we find in both TTs and which involve the adjective complete. This restriction makes the amount of data much more manageable. In any event, as can be seen in the following discussion, it turns out that only a handful of constructions that can be found on the Web refer to definition in the sense of computer-based visual processing in 3D.

My first set of data concerns give (a complete) definition. In fact, using a Web browser, I can only find one or two marginal examples of definition as a complement in this sequence. Instead, examples (21–25) below show that give a complete + N more typically refers to the results of an experiment, encapsulated metaphorically as a diagram, picture or table. None of these refer to a definition in the “visual processing” sense, however:

(21) ... this hierarchy of tests gives a complete and also computationally and theoretically appealing characterization of mixed bipartite entangled states.
The face plate gives a complete description of all signals with their pin numbers and explains the DB25/9 conversion.

This simulation gives a complete overview of strategic leadership approaches.

The kinetic NMR method described in this work gives a complete picture of the reconstitution process in water, close to industrial conditions.

The above URI gives a complete snapshot of the Indian Stock Market.

The construction present a complete + N also very rarely has definition as complement. What I find instead are subjects with metatextual references to books, documents, tables and complements referring to the kinds of argumentation or sub-section that one typically finds as parts of these texts (a similar set of contexts can be seen for other ditransitive verbs such as show). Although present a complete + N is a recurrent and regular lexicogrammatical pattern, it does not generally correspond to the wording we have in our translation, as the following examples show:

Table 5 presents a complete list of N-body scalar operators for N less than five.

Now in its fourth edition, this book presents a complete course in electrocardiography (ECG) for students and a reference for advanced trainees and . . .

The outcome of the Evergreen T4 Meetings, this book presents a complete overview of T4 research, from its earliest history to its latest developments.

This document presents a complete numerical example for the document “Proposal for the risk model (using standard modelling and equations)”.

This paper presents a complete integrated NLG system which uses a Description logic for the content determination module.

On the basis of this evidence, give a complete + N appears to be more appropriate than present a complete + N. Although the wording of TT1 give a complete + N is different to that of our ST, the typical subjects used in this construction do refer to processes which are relevant to our translation. However, the (textual) frame of reference of TT2 present a complete + N appears to be much less relevant. Given no other choice, we would have to conclude that in this case the translation TT1 has made a more appropriate “lexicogrammatical choice” than TT2.
However, for the sake of completeness, I would like to consider some other candidates, especially the very many examples of verbs which regularly introduce a complement of the form a complete 3D + N. For example, if we look for create a complete 3D + N using a Web browser, it is possible to find many thousands of examples in which the subject is a company or some technical innovator, and in which the complement involves a 3D environment or 3D model:

(31) **VR context’s Walkingside product automatically creates a complete 3D virtual model of facilities, merging files from diverse data sources and proprieta**

(32) **DELMIA QUEST creates a complete 3D digital factory environment for process flow simulation and analysis, accuracy, and profitability**

(33) **With these types of design programs, the architect or designer creates a complete 3D model that contains information about all the components and structures**

(34) **Parish and Müller [15] present a system that creates a complete 3D city model using a small set of statistical and geographical input . . .**

(35) **Here, the computer creates a complete 3D environment, and the user, represented by their own avatar, can move around the 3D space, meet and interact with**

Although none of these examples (31–35) involve definition, the items that emerge as complements correspond more or less exactly to what is referred to in the ST: the creation of a 3D virtual working environment. It is also worth noting that example (32) comes from a technical report about the company DELMIA, who happen to supply Dassault Systems (reported on Business Wire: the URL is referenced at the end of this paper). It occurs to me that this is an entirely predictable result. By a process of exploring increasingly specific patterns of speech in this way, it is possible to discover the emergent “style” (or should I say “phraseology”?) of a particular specialism or industry.

Nevertheless, although create a complete + N is quite a good candidate translation, it would be interesting to examine some further possibilities. Also, by examining some other candidate verbs, it is easier to understand how specific the LG patterns for each of these examples are, as compared with other constructions. For example, it is possible to find many hundreds of examples of the sequence produce a complete 3D + N, but these do not refer to a definition. But they do refer to the production of visual 3D data and 3D objects:
After one full revolution, which takes 30 to 60 min, the system produces a complete 3D data sheet (see Fig. 1).

Within a couple of seconds, the system produces a complete 3D digital map of the tile surface and then automatically finds and quantifies defects.

Interpretation begins with the density map and the (provided) amino-acid sequence(s) of the protein forming the crystal, and produces a complete 3D molecular model of the protein.

Finally, a simple mathematical inverse Fourier-transform produces a complete 3D reconstructed object.

Laser profiling equipment, also utilized at GME, produces a complete 3D survey of the rock face in front of the boreholes.

Finally, we come to the sequence provide (a complete) (3D) + definition. This pattern turns out to be much closer to our ST, in that it refers in many cases to 3D visual processing, and often takes definition as a complement (42–45). It is also interesting to note that provide happens to be the only verb that has definition as a regular complement in the BNC (there are nine occurrences, although none of these involve computing or definition in the sense of “visual processing”). In addition, the typical subjects in this pattern refer specifically to computational processes, as can be seen in the following examples:

Starting from the engine data, our propeller design software enables us to define the aerodynamic and structural characteristics of the propeller, and provides a 3D definition of the propeller compatible with most CAD software.

The trace information preferably provides a complete definition of the place in and/or contribution to one or more ECMPs associated with the identified.

It is obvious in all this that the Burgers vector provides a complete definition of the dislocation.

The data model provides a complete definition of the ISO zonal and nodal configuration.

Photo Backup provides a complete solution for handling photographs from digital cameras.

The data I have examined in this section are more messy and less conclusive than for the patterns push the envelope / push (back) the limits.
The reason for this is simple: we are looking at two very productive constructions. However, even the brief lexicogrammatical survey conducted here still leads to a fairly firm conclusion: the construction in TT1 give a definition is marginally better than TT2 present a definition. Further examination, however, shows that neither of these patterns share the exact same lexicogrammatical forms or the same frame of reference as the French ST. Instead, several other candidate patterns exist in English, although each has its own very specific context of use: create a complete 3D environment / model (in a commercial / computational context), produce a complete 3D model (in a mathematical / image-processing context) and provide a complete definition (in a computational / image-processing context). These results suggest that neither TT1 nor TT2 has successfully captured the ambient lexicogrammar of technical English in the domain of aviation, although it has to be said that TT1 provides (if that the right verb) the closest approximation that could have been achieved in an on-the-spot translation exercise.

6. Conclusion

In this paper I have shown how a lexicogrammar approach can be used in order to check the quality of a translation. At the heart of this approach are a number of basic assumptions about how language works and how linguistic analysis ought to proceed. These can be summarised as follows:

1. The lexicogrammar approach assumes that each sign in the language has its own particular lexical and grammatical niche in the language system, expressed in discourse as a lexicogrammatical (LG) pattern.
2. The lexicogrammar approach uses empirical data, in particular text corpora, as well as concordances and contextualised examples, in order to analyse, in a forensic way, the habitual use of signs in discourse.

Most of the terms used here have been defined or exemplified in the preceding text. However, I would like to end this paper by focussing on the notion of “forensic” analysis. I have used this term here because I believe that rather than analysing every feature of a text (which is possible, but highly time-consuming), it is usually only necessary to focus on one or two key problems or areas of doubt. A forensic approach is by its very nature meticulous and data-hungry, and thus requires the comparison not only of the particular text in question together with a more representative corpus of texts, but also a triangulation of data, that is to say the com-
parison of different types of data sets and examples (Coulthard 1995, Coulthard and Johnson 2007). The analysis I have set out above does not in fact use representative and comparable corpora of French texts, or representative corpora of texts in English in the same domain as the French original. However, few corpus linguists (indeed, few forensic scientists!) have access to such data, and even fewer professionals in the business of translation would have the time to build databases and to tag them properly (because for most linguists, a “real” corpus is in fact not just a computer-held archive of texts: it is a planned collection of texts that have been prepared or “tagged” for linguistic analysis).

Having said all this, I believe that even the analysis of “LG patterns” I have set out above is more systematic and empirically rewarding than an analysis carried out by simply comparing the two translations “manually”, that is to say by relying on intuition or gut-feeling. In addition, the analysis carried out here is certainly within the reach of most professionals. The analysis of LG patterns does not require the sophisticated classification involved in the analysis of “phraseological units”, nor does it require the statistical analysis that is often required in the analysis of “collocations”. Rather, the LG approach simply requires the use of an on-line or corpus-based concordancer and the ability to observe and compare the typical patterns of use of signs in different textual contexts.

So, as I pointed out in the introduction to this paper, I believe that specialists in the field of translation, in particular those of us interested in quality assessment, should at least be familiar with this type of approach, as well as its implications. In particular they should be aware of the notion of lexicogrammar, especially since this is considered by many linguists to be a central aspect of how language works. And although the analysis of LG patterns may not be able to answer every question relating to a translation, this type of analysis at least reminds us why the act of translation is at the same time so difficult, mysterious and ultimately, so fascinating.

References

Banks, David
1994 Writ in water, aspects of the scientific journal article. Brest: ERLA, Université de Bretagne Occidentale.

Bloor, Meriel and Tom Bloor
A Lexicogrammar approach to checking quality 93

Bowker, Lynne

Corpas Pastor, Gloria

Coulthard, Malcolm

Coulthard, Malcolm and Alison Johnson

Firth, John Rupert

Firth, John Rupert

Gledhill, Christopher

Gledhill, Christopher
2005a Free-range clusters or frozen chunks? Reference as a defining criterion for linguistic units. Recherches anglaises et Nord-américaines 38: 25–43.

Gledhill, Christopher and Pierre Frath

Goldberg, Adele

Halliday, Michael. A.K.


Pavel, Silvia 1994 Guide de la recherche phraséologique en langue de spécialité. Division Montréal-Québec Direction de la terminologie et de la documentation Bureau de la traduction.


A Lexicogrammar approach to checking quality

95

Siepmann, Dirk

Sinclair, John McH.

Sinclair, John McH., Jonathan Payne, and Chantal Pérez Hernandez

Stubbs, Michael

Tucker, Gordon

Tucker, Gordon

Xiao, Richard Z. and Ming Yue

Appendix 1. Source Texts


The reference to DELMIA and Dassault on “Business Wire”: http://www.thefreelibrary.com/ArvinMeritor+Visualizes+Real+Savings+with+DELMIA+Digital...a0135432657 (accessed on 10 February 2010)

Appendix 2. URLs of cited examples

The Web examples cited in this paper (numbered 1–45) were last accessed on 22 March 2010.
7. http://www.esa.int/esaCP/SEMSPU1P4HD_FeatureWeek_0.html
11. http://modernheirloomfinejewelry.blogspot.com/
17. http://www.protocolanalyser.com/rs232lm.htm
18. http://www.esa.int/esaCP/SEMSPU1P4HD_FeatureWeek_0.html
directories.headline&searchtext=v&searchmethod=starts&submit=
submit