Core collocation colligation preference prosody

In discussing the phrase *true feelings,* he illustrates this tendency more explicitly. The core or invariable part (by which we recognise the item in the linguistic trace) is *true feelings*. It collocates with items like *their, our, your, his, of, reveal, express, deny*. In colligational terms, these items fall into two grammatical categories: firstly, possessives (adjectives and *of* structures) and verbs (which normally appear to the left of the core showing how *true feelings* is usually in object position, e.g. '...to mask his true feelings'). These verbs also reveal a preference for the semantic area of expression and non-expression. The overall semantic prosody is unfavourable: a broad sense of people's difficulty in or reluctance towards expressing their true feelings.

Sinclair, in general, argues strongly against what we might call the 'mail-order' view of meaning in which 'each word carries or creates meaning by itself' and a speaker/writer only 'chooses the word, and the meaning arrives in the text' (p. 129). But these tasks show quite incontrovertibly not only how one word produces different meanings on different occasions, but also how word choice both depends on, and dictates, syntactic and semantic features of the surrounding context. Words come with their own unique baggage and their own requirements, many of which are not evident to unsupported introspection.

The coming of electronic corpora has brought about a sea-change in the way we study language. The key skill in this new environment is to be able to interrogate the corpus efficiently. *Reading Concordances* is both the first step in acquiring this innovative skill and an introduction to the field of corpus-based lexicology, as well as a model of inductive learning. The book presupposes no previous knowledge of corpus linguistics research and methods and the studies presented in it can easily be replicated. It is written in a clear and accessible style and significant technical terms are explained in the glossary. All this makes it an ideal textbook for introductory corpus linguistics courses and ELT teacher training programmes.

Alan Partington
Departimento di Lingue e Letterature Straniere
Università di Bologna
Bologna, Italy
E-mail address: partington@lingue.unibo.it

doi:10.1016/j.system.2004.06.003

Computers and translation: a translator's guide
Somers, H. (ed.); Amsterdam and Philadelphia, John Benjamins, 2003, xvi + 349 pp.

A professional translator can nowadays hardly dispense with the use of a computer. However, the history of machine translation (MT), with its misplaced claim to be able to replace the human translator 'soon', has caused many translators to be

cautious or even suspicious of any use of the computer beyond word processing. Meanwhile, the field of machine translation has changed quite radically and produced a variety of tools for use in collaboration with human translation. Much of the professional translators' reluctance to take full advantage of these tools is due to ignorance. The book under review sets out to give an accessible overview of the issues surrounding computer support in translation from the perspective of the professional translator.

The book consists of 17 chapters, most of which are around 20 pages in length. After the introduction, which includes a historical overview of the field, six chapters are devoted to the use of computers by translators, followed by 10 chapters on MT.

Chapters 2–4 concentrate on tools to be used by the translator. Chapter 2 'The translator's workstation' briefly presents a number of commercially available applications which are useful to the translator, ranging from word count and dictation tools to online dictionaries and concordances. Translation memory systems, treated in Chapter 3, match (fragments of) an existing source language text with earlier translated texts and propose translations on the basis of what they find. Chapter 4 deals with terminology tools. As professional translators spend an important part of their time on terminological research, various tools for the storage and retrieval of the results of this research have been developed.

Chapters 5–7 deal with other issues of practical concern to the translation. Localization, presented in Chapter 5, is the adaptation of products and manuals to a local context of use. It concerns especially software packages. Chapter 6 treats the problems related to working with 'minority languages', i.e. languages for which not much is available in terms of computational resources and products. Chapter 7 summarizes how different types of corpora can be used in translation research, training and practice.

The chapters on MT start with two chapters devoted to classical approaches in MT research. Chapter 8 presents some of the problems in analysis, transfer and synthesis modules as they arise in transfer and interlingua architectures. These dominated the field in the 1970s and 1980s and still remain important. It also sketches more data-oriented approaches to MT which emerged in the 1990s. Chapter 9 explains how linguistic theory can be used in transfer and interlingua models of MT.

The next three chapters cover commercial MT systems. Chapter 10 gives a historically oriented overview of commercial MT. Chapter 11 presents the issues of development and maintenance of a commercial MT system from the provider's point of view. It also addresses the question of how to use such a system to best effect. From a similar perspective, Chapter 12 discusses online translation services. Analysis of the translations entered shows that many are one-word translations and sentences apparently intended to trap the system. Such analyses, as well as bug reports, play an important role in system maintenance.

Chapter 13 is devoted to the evaluation of MT systems. It explains the problems involved in evaluation and sketches a number of methods that have been or may be used.

Next, three chapters describe different ways of using the limited possibilities of available MT systems in a real-life context. A controlled language, as treated in Chapter 14, is a set of constraints on the use of a natural language such that authors write texts which are easier to translate. The idea is illustrated by an extensive discussion of Caterpillar Technical English and the tools developed to support the authoring process. Chapter 15 considers sublanguages. A sublanguage is a naturally occurring variety of language use, which may exhibit constraints facilitating MT. The most famous example is the language used in Canadian weather bulletins, for which a fully automatic MT system has been in operation since 1977. A different approach to using MT in a practical context is to correct the output in a post-editing phase. Chapter 16 gives some examples of post-editing work as well as a characterization of different types of post-editing.

Finally, Chapter 17 suggests different contexts in which MT could be used in teaching, ranging from language teaching to translator training and job-related seminars on effective communication.

As this is an edited volume, the chapters are written by specialists in the specific areas. Thus, chapters 8 and 9 are written by Doug Arnold and Paul Bennett, who have been active in the field of MT at least since the days of the Eurotra project. Among the authors are not only academic specialists, however. Chapter 5, for instance, is written by Bert Esselink of L10nbridge, one of the major localization companies. Given the topics and aims of the book, this is entirely appropriate. In some cases, however, it tends to lead to rather uncritical coverage of the topic. In Chapter 12, for instance, two representatives of Systran present an account of online MT which only narrowly steers clear of outright publicity. Apart from the external contributions, there are also six chapters written by the editor.

An editorial choice which will strike any reader who reads the entire volume is that each Chapter is apparently designed to be independently readable. This leads to a quite significant amount of redundancy. Chapter 10, introducing commercial systems, for instance, recapitulates the history of MT (from Chapter 1), the translator's workstation (Chapter 2), localization (Chapter 5), and minority languages (chapter 6), while introducing chapters 11 and 12. In several chapters where the issue of minority languages is referred to, the same statement about the small range of languages for which computational tools are available, including the same list of languages, is made. Presumably, this redundancy is an advantage for readers who consult individual chapters only.

Another editorial decision concerns the style of the contributions. Although there is a range of variation typical of an edited volume, a clear effort has been made to make the style accessible rather than academic. The concern for accessibility is quite striking in the section on statistical approaches to MT in Chapter 8, which manages to avoid formulae and formal definitions completely. In some cases, the apparent concern for brevity in the presentation of pictures detracts from understandability. Thus, without knowing *DéjàVu*, it is difficult to make sense of Fig. 5 on p. 41. Similarly, the functionality of *TransIt-*TIGER is at most suggested by Fig. 2 on p. 330. At least the references are given to find a more detailed explanation.

While the large majority of the chapters are well-written, in three cases, I find the organization and presentation of information not up to the generally high standard of this book. Chapter 13, on evaluation, is full of sloppy formulations. The information is not properly structured and some sections are little more than enumerations without proper explanation. Similar defects are found in Chapter 16, on post-editing. Ironically, some post-editing of this Chapter might have helped to remove spelling errors, ungrammatical sentences and stylistically awkward phrasing. The section on post-editing guidelines and criteria consists of an enumeration of lists used in the classification of errors corrected in MT output. As these lists are neither explained nor exemplified, they do not contribute to the understanding of the issue. It would have been better to choose one such list and explain the nature and relevance of the classes.

Problems of a different kind affect the readability of Chapter 14. Section 5 of this long Chapter (twice as long as average) is devoted to a case study of controlled language, which gives the impression of a slightly edited conference presentation. More than half of the 17 pages of this section are presented in the form of lists, which seem to be taken directly from the slides and interspersed with some introductory text. The text also includes statements such as 'the average number of syntactic analyses dropped from 27% to about 1%' (p. 273), which are somewhat difficult to interpret.

In a book written for professional translators, errors in the examples with foreign languages are embarrassing. When the plural of German *Hund* ('dog') is given as *Hünde* instead of *Hunde* (p. 3), Dutch *men* is replaced by *man* (p. 129), German *adoptieren* is analysed as ambiguous in the same way as its English and French cognates (p. 150), or when *même* is translated as 'even' in 'sur la notion même' (p. 310), this is likely to diminish the credibility of the text for specialists in these languages. In each case, however, the substance of the argument is not affected by the errors.

The book as a whole is a welcome addition to the literature on computational linguistics. As an overview of the field aimed at the professional translator it is basically without competition. The problems noted above are really minor ones in view of the highly useful overview this book provides. For translators who have no background in computational linguistics this is a one-stop reference work. It can also be used very well in translator training. Its further reading sections make it also a good starting point for exploring the field in more detail. As the book is not written by translators, it does not presuppose translation experience and is equally accessible to others interested in recent developments in MT and computer-aided translation.

Pius ten Hacken
School of European Languages
University of Wales Swansea
Singleton Park
SA2 8PP, Swansea, UK
E-mail address: p.ten-hacken@swansea.ac.uk

doi:10.1016/j.system.2004.06.004