PUBLISHER'S DESCRIPTIONS OF LINGUISTIC BOOKS: COMPARISON OF CZECH AND ENGLISH DISCOURSE STRATEGIES

Aleš Klégr

Abstract

The paper investigates a specific type of genre, descriptions of linguistic books in Internet catalogues of three Czech and three Anglophone publishers. It strives to capture their structural organization and some other characteristics that could bring into relief differences between the Czech and the English discourse strategies of their composition. Using a two-level analytical framework to identify the structural parts of the descriptions and the meaning components that realize them, the study finds certain tendencies that to some extent distinguish between them. Still, the English descriptions and even less so the Czech ones are far from homogeneous and striking dissimilarities can be found both between and within the Anglophone and the Czech publishers.

Key words

English-Czech, book descriptions, on-line catalogues, discourse structure, contrastive study

1 Introduction

Publisher's descriptions of books in (on-line) catalogues are an important means of promoting these books for commercial, but also for academic purposes by providing essential information about them. The fact that large Anglophone publishing houses and Czech publishers considerably differ in the size of their output and markets, and in their respective histories and traditions raises the question as to what extent their descriptions will differ in format, internal organization and style. As, moreover, the Czech publishers often need to have Czech descriptions subsequently translated into English, the contrastive view of original English and original Czech book descriptions is both of theoretical and practical interest. Although the study is a tentative first step, the analysis and comparison of the English and Czech book descriptions hope to reveal some of the essential features specific to this particular genre in each language.

2 Discourse description

In order to place the study within a context, it can be described as falling, broadly speaking, under the heading of corpus-based discourse analysis.

Discourse analysis approaches divide into several categories according to Schiffrin, Tannen and Hamilton (2001). Basically, the authors define discourse as dealing with (1) anything beyond the sentence, (2) language use, and (3) a broader range of social practice that includes non-linguistic and nonspecific instances of language. Accordingly, the present study is of the "beyond the sentence" type and its focus on the internal structure and organization of book descriptions makes it an analysis of a specific genre as a rhetorical category (cf. Biber 1995, Martin & Rose 2008, Tardy 2011).

There are many studies of a similar kind reported in the literature. This particular type of discourse analysis has been extensively explored in Biber, Connor and Upton (2007). Using corpus-based methods to describe discourse organization, they adopt two approaches, 'top-down' and 'bottom-up'. In the top-down approach, the starting point of analysis is the a priori development of an analytical framework and its functional components. The framework is then applied to texts which are analysed in terms of these components. The bottom-up approach begins with the automatic segmentation of corpus texts into linguistically-defined discourse units whose systematic functional characteristics are then investigated in the texts. While the bottom-up approach is relatively new as it requires advanced computational techniques, most of the previous discourse analysis studies have been of the top-down kind. Biber et al. (2007) briefly review the top-down discourse structure theories that have cropped up since the 1980s, among them the Rhetorical Structure Theory presented in Mann and Thompson (1992) and applied to fund-raising texts. However, Biber et al. (2007) concentrate on and illustrate two particular top-down approaches to discourse – the so-called move analysis (Swales 1990), identifying the preferred sequences of move types and steps in the text, and rhetorical persuasive appeals analysis, dividing texts into sections using the three Aristotelian appeals, *ethos, pathos,* and *logos*. They are applied in the book to fund-raising letters, biology/biochemistry research articles, and university classroom teaching.

3   Book descriptions, reviews and abstracts

Before addressing publishers' descriptions, it needs to be said that although a special kind of genre, they have some affinities with at least two others: book reviews and abstracts. Both book descriptions and reviews inform about books, yet each has a slightly different agenda: book descriptions promote and sell the book by providing information about it and at the same time stressing its merits. Their aim is to advertise, so they focus on the aspects that make the book worth buying and disregard the potential shortcomings. Reviews, on the other hand, inform and evaluate. In addition, they are much longer than descriptions which

are typically about the same length as article abstracts. The function of abstracts, however, is only to summarize what the text is about, but neither to advertise nor assess it. The common goal of descriptions, reviews and abstracts – to present the contents in a systematic way – argues for similarities in internal organization; also, all three can be expected to conform to a particular style whether dictated by usage or determined by the publisher. The assumed similarities invite comparison which could highlight the specific features of each one of them. However, unlike reviews and abstracts, descriptions have apparently not been explored as yet. Book, or even film, reviews, on the other hand, have received a relatively great deal of attention, cf. for example Hyland (2004), dealing with interactions in book reviews, Lindholm-Romantschuk (1998), covering scholarly book reviews in the social sciences and humanities, or Taboada (2011), describing on-line movie reviews. Even abstracts appear to have been subject to investigation, by, e.g. Doró (2013), analysing rhetorical structure of research article abstracts, or Stotesbury (2003), describing evaluation in research article abstracts. The absence of research on book descriptions means starting from scratch, deciding on a method, etc. And so, although reference to book reviews and even abstracts could be useful to some extent, it seemed better to avoid comparison and more detailed demarcation of these three genres at this stage and address analysis of book descriptions independently.

4   Book descriptions study

Given the size of the samples, the adopted approach is inevitably top-down. The first step was to collect the samples whose examination was the starting point for the development of an analytical framework. The framework and its features were devised specifically to fit the descriptions under study and reference to concepts such as 'move' or 'appeal' was purposely avoided.

4.1 Sample collection

Both samples are comprised of descriptions of linguistics books. The total of 120 descriptions were downloaded from the web pages of the selected publishing houses: 20 from three Czech and 20 from three Anglophone publishers, forming a Czech language and an English language sample of 60 descriptions each. The described books come from different fields of linguistics, they are mostly monographs, but also academic textbooks and even dictionaries by single authors or editors, or teams. The descriptions were randomly selected regardless of format or length. The publishers were chosen on account of their reputation. The three Czech leading publishers of academic books are Karolinum (under Charles University, Prague), Nakladatelství lidové noviny (NLN, s.r.o., a private commercial publisher) and

Academia (originally under the Czechoslovak Academy of Sciences, now a private commercial subject). The Anglophone publishers are Cambridge University Press, John Benjamins Publishing Co. and Oxford University Press (USA). Although all six publish linguistic books, their publishing records are hardly comparable as far as the range and number of these books are concerned (and accordingly greater attention to house style was expected in the Anglophone publishers). The output of the Czech publishers in this field is minuscule by comparison and there are other differences. Descriptions in English Internet catalogues are often accompanied by further separate text sections (such as Table of Contents, Resources, About the Authors, Reviews or Subjects). In the Czech catalogues only Karolinum complements descriptions (called 'anotace') with a contents section. Only descriptions proper are taken into account.

4.2 Analytical framework

The starting assumption was that descriptions come in a finite number of patterns, and that the patterns involve a finite number of parts. Examination of the descriptions suggested that their structure is best captured at two levels. The first clue to the structure of descriptions is their division into paragraphs: although 71.6 per cent of the 120 descriptions have only one paragraph, 21.6 per cent have two and 6.7 per cent three. While a single paragraph text could hide all kinds of internal organization, descriptions with two and especially three paragraphs are indicative of the author's intended structuring of the text. Indeed, content analysis revealed that not only three-paragraph but also single-paragraph descriptions are internally structured into as many as three parts: introduction – core (nucleus) – closure. Although a tripartite pattern was taken as the basic frame of reference, content analysis also showed that almost half of the 120 descriptions (55; 45.8%) display a two-part pattern. Introduction and core are the key parts in both three- and two-part patterns. It appears that these two key parts are typically (but not exclusively) realized by a specific type of information. The introductory part mostly indicates the area the book covers, a type of information that can be called 'topic'. The core part informs about which specific aspects or problems of the area the book explores (in fact, outlining the book's contents), and can be called the 'focus' (analogy with information structure seems inevitable and logical).

However, although the two key parts of the book description are typically associated with topic and focus respectively, it is not always the case as they can be sometimes replaced by a different type of information. Moreover, in the third part of the tripartite pattern several different types of information may alternate. After assessing the types of information found in the descriptions (trying to make things as simple as possible, but not simpler), five types of information (or

meaning components) were identified: in addition to 'topic' and 'focus', it was 'addressee' (a component specifying the intended readership), 'goal' (a component specifying explicitly the purpose of the book) and finally 'differentia' (the most broadly defined component specifying any kind of unique feature or circumstance which distinguishes the book from other books). These three additional components may, in fact, appear in any of the three structural parts and combine with the key components or the other additional components within one part (paragraph). Hence, descriptions are analysed in terms of patterns (consisting of one, but mostly of two or three parts) and components which realize the parts of the patterns. Descriptions with two-part patterns typically consist of introduction and core. Descriptions with only one part are rare (two Czech language ones) and were found to be realized by only one component (focus or differentia). Whatever their structure and components, descriptions are of course never incomplete, they are always self-contained texts.

Jean Aitchison: Language Change. Progress or Decay?

How and why do languages change? Where does the evidence of language change come from? How do languages begin and end? This introduction to language change explores these and other questions, considering changes through time. The central theme of this book is whether language change is a symptom of progress or decay. This book will show you why it is neither, and that understanding the factors surrounding how language change occurs is essential to understanding why it happens. This updated edition remains non-technical and accessible to readers with no previous knowledge of linguistics.

Figure 1: An example of a three-part organization of a single paragraph description

Brian Roark and Richard Sproat
Computational Approaches to Morphology and Syntax

The book will appeal to scholars and advanced students of morphology, syntax, computational linguistics and natural language processing (NLP). It provides a critical and practical guide to computational techniques for handling morphological and syntactic phenomena, showing how these techniques have been used and modified in practice.

The authors discuss the nature and uses of syntactic parsers and examine the problems and opportunities of parsing algorithms for finite-state, context-free and various context-sensitive grammars. They relate approaches for describing syntax and morphology to formal mechanisms and algorithms, and present well-motivated approaches for augmenting grammars with weights or probabilities.

Figure 2: An example of a two-part organization of a two-paragraph description

The analytical framework includes only one linguistic characteristic: the syntactic subject introducing the two key parts. The reason for including it was that the type of subject used in descriptions was found to set the general stylistic tenor of the description (the presentation style), and may to some extent differentiate between the descriptions (and the two languages). Two groups of subjects were distinguished: subjects of the type *(Tato) Kniha/This book* (and its synonyms or co-hyponyms) and other types of subject (*author, goal*, subject of a question, descriptive or topic-related subjects). Thus, to take an English model example, the description of Aitchison's book in Figure 1 is an instance of a tripartite structure, with the introduction presenting topic, the core part containing focus and the closure indicating readership. While the key subject in the core part is *This book*, the introduction uses other types of subject (*language, evidence*). In Figure 2 the bipartite internal organization of the description and its two-paragraph structure coincide. The introduction is realized by two components, topic and readership, and the subject is of the type *This book*. The core is comprised of focus only and uses a different type of subject (*author*).

All in all, the descriptions and the samples were examined and compared in the following five characteristics: length (word count), number of paragraphs, distribution of patterns, distribution of components and their combinations, and the presentation style of introduction and core. Needless to say that the

boundaries between parts and between components are sometimes blurred, and so their identification is to some extent subjective. Also, the number of the examined descriptions is relatively small and so the data has to be treated with caution. It gives trends rather than accurate figures, different sets of descriptions could yield somewhat different results.

5 Sample analysis

Using the analytical framework and elected characteristics, the total sample of 120 descriptions consists of 14,961 words. The average length of a description is 124.7 words, 86 descriptions are formed by one paragraph (71.6%), 26 have two paragraphs (21.6%), and just eight descriptions have three paragraphs (6.7%). A tripartite internal pattern was found in 63 descriptions (52.5%), a two-part pattern in 55 descriptions (45.8%), and two descriptions were analysed as having only one part. The structural parts are mostly realized by just one component, still there are 26 cases when the part (especially introduction) is realized by two or even three components (introduction – 24 combinations, core – one combination, closure – one combination). However, the aggregate numbers for all the 120 descriptions are less interesting (and meaningful) than the situation in the original Czech language and English language descriptions. The respective findings for the Czech and the English sample document the tendencies within them and allow comparison of the composition strategies.

5.1 The Czech language sample analysis

The results of the analysis of the Czech language sample are summarized in Table 1. The most striking feature of the sample is the great differences between descriptions of the three publishers in the observed characteristics. So, NLN and Academia differ the most in the quantitative aspects: word count (1,479 vs. 2,559) and average length in words per description (74 vs. 128). Karolinum and NLN markedly contrast in the distribution of paragraphs and even patterns. They also differ in the way the introductory part is realized (10 vs. 3 combinations of components respectively). Academia and Karolinum differ, for instance, in the type of component most frequently realizing the closure (8 cases of readership in Academia, 5 cases of differentia in Karolinum). In both introduction and core all three subsamples show a distinct tendency to presentation by means of the subject "*(Tato) Kniha/Publikace/ Práce,* etc." (= This book) (68.1%) rather than by other types of subject (31.9%).

publisher feature	total		Karolinum	NLN	Academia
word count	5,954		1,916	1,479	2,559
average length in words	99 2-part patterns 3-part patterns		96 84 108	74 77 94	128 134 (! longest des.) 124 (! shortest des.)
number of paragraphs	1: 2: 3:	39 19 2	20 - -	2 17 1	17 2 1
patterns	1-part: 2-part: 3-part:	2 31 27	- 10 10	2 13 5	- 8 12
introduction: components	1 comp. 2-3 comp.	37 21	9T, 1D 6TD, 2TA, 2TG	15T 2TD, 1TG	12T 4TD,2TG,1TA,1TAG
core: components	1 comp. 2 comp.	57 1	20F	17F, 1G	19F 1GF
closure: components	1 comp. 2 comp.	26 1	5D, 2A, 2G 1GA	3A, 2D	8A, 2D, 2G
closure: sum of components		27	A 13 (48.1%), D 9 (33.3%), G 4 (14.8%), GA 1 (3.7%)		
introduction: style of subject	*This book* other S	44 14	17 2d, 1g	10 5d, 1a, 1?, 1g	17 2d, 1g
core: style of subject	*The book* other S	35 23	14 3a, 3p	8 8p, 2a	13 1a, 1g, 5p

Abbreviations: components: T – topic; F – focus/contents; A – addressee; D – differentia; G – goal; types of other subject: a – author; d – descriptive; g – goal; ? – S of question

Table 1: The Czech language sample analysis results

5.2 The English language sample analysis

The findings about the English descriptions presented in Table 2 show that they are far more uniform, for instance in the overall word count (the greatest difference is between Cambridge UP and Oxford UP, 436 words compared to 1,080 words between NLN and Academia) and the average number of words per description (cf. the difference of 32 words between Oxford UP and Cambridge UP

in contrast to 54 words between NLN and Academia). All three subsamples also share the tendency to avoid combining different components in one structural part and to prefer the specification of readership in the closing part to other types of information. As regards the presentation, although the subject *This book* (54.2%) is more frequent than other types (45.8%), the difference is relatively small. Of the three, it is the Oxford UP descriptions that show the greatest divergence. In general, however, it seems that the three Anglophone publishers are more united in their approach to descriptions which suggests a greater awareness and observance of 'house style' than in the Czech publishers.

publisher feature	total	Cambridge UP	Benjamins	Oxford UP
word count	9,007	2,767	3,037	3,203
average length in words	150 2-part patterns 3-part patterns	128	152	160
number of paragraphs	1: 47 2: 7 3: 6	20 - -	18 2 0	9 5 6
patterns	1-part - 2-part 24 3-part 36	- 5 15	- 11 9	- 8 12
introduction: components	1 comp. 57 2-3 comp. 3	19T 1TD	20T	18T 1TA,1TG
core: components	1 comp. 60 2-3 comp. 0	20F	20F	20F
closure: components	1 comp. 36 2-3 comp. 0	11A. 4D	6A, 3D	8A,3D,1G
closure: sum of components	36	A 25 (69.4%), D 10 (27.8%), G 1 (2.8%)		
introduction: style of subject	*This book* 33 other S: 27	4 12d, 4?	15 3d, 2a	12+2td 2a, 4d
core: style of subject	*This book* 32 other S 28	12 8a	11 5d, 4a	9 9a,1ad,1d

Abbreviations: components: T – topic; F – focus/contents; A – addressee; D – differentia; G – goal; types of other subject: a – author; d – descriptive; g – goal; ? – S of question; t – this book

Table 2: The English language sample analysis results

6 Comparison of the Czech and English sample findings

The ultimate goal of the study was to compare the Czech and English descriptions. So how do the English descriptions differ from the Czech ones? The differences are as follows:

- In terms of length the English descriptions with 150 words on average are longer by one third, compared to just 99 words of the Czech descriptions (which reflects the difference in the size or total word count of the samples: 9,007 and 5,954 respectively).
- In terms of paragraphs, the English descriptions more frequently form a single paragraph (47; 78.3%) than the Czech descriptions (39; 65.0%), and they notably less frequently have two paragraphs: seven instances compared to 19 in Czech.
- In terms of structure, English descriptions show a greater tendency to tripartite patterns, 60.0 per cent (36), in Czech only 45.0 per cent (27); by contrast, 52.0 per cent (31) of Czech descriptions have two-part structures (in English it is only 40.0%; 24) and two Czech descriptions have only one part (NLN, with single components: focus/contents, differentia).
- In terms of the composition, the two essential structural parts, introduction and core, are formed by topic and focus respectively in both samples. However, while in English the topic in the introduction and the focus in the core are never replaced by anything else, and only in three cases the topic combines with another component (differentiating feature, addressee and goal), the situation in the Czech sample is more complicated: leaving aside one-part descriptions, there are 37 introductions realized by a single component (topic in 36 cases and once by differentia), and 21 introductions realized by a combination of topic and another component (12 differentia, 3 addressee, 5 goals) in 20 cases and once by a combination of topic and two other components (topic-addressee-goal); the 58 core parts of Czech descriptions are realized by focus in 56 cases, once by goal and once by a combination of components (focus – goal). In sum, the realization of the structural parts in Czech descriptions is more heterogeneous, and especially the introduction includes a combination of meaning components (21 compared to 3 in English).
- The third part, closure, consists of the specification of the addressee (readership) in 25 of the 36 three-part English descriptions (69.4%), and although addressee comes first as a closure even in Czech descriptions (13 out of 17; 48.1%), it is by more than 20 per cent less frequent. The second commonest closing component is differentia in both English (27.8%) and Czech (33.3%). The next component, goal as a closure, forms 14.8 per cent in Czech, in English it is marginal (2.8%).

- In terms of style, the English introductions are only slightly more often presented by the subject *This book* (33; 55.0%) than by a different kind of subject (27; 45.0%); in the Czech introductions the use of *This book* was distinctly prevalent (44; 75.9%), other subject (14) forming only 24.1%.

- the presentation style of the core part shows a mild decline of the subject *The book* in English descriptions (32; 53.3%) compared to other types of subject (28; 46.7%); by contrast, the use of *The book* in the core of Czech descriptions is again more common, 35 cases (60.3%), while other types of subject appeared in 23 cases (39.7%). In other words, the presentation of the introduction and core in English tends to be more inventive and varied, while in Czech it is more repetitive.

Given the size of the samples, these quantitative differences, however small in some cases, are all that the interpretation of the tendencies can be statistically based on at this moment. There are, of course, other aspects of style that could have been taken into account. As a matter of interest, in two Czech language descriptions (Karolinum) the academic degree of professor was found with the name of the author, not an uncommon deferential usage in Czech but proscribed in English style guides (and generally regarded as typical of Czech English when appearing in translations). However, and contrary to expectation, the academic degree (Prof. and Dr.) appeared before the names of authors in four English descriptions (Oxford University Press) as well.

To summarize, the typical English description is longer than the Czech one, consists of one paragraph organized into three parts, each composed of a single functional component. The closing part specifies the readership. The introduction and core are often 'literary' in style, avoiding the pedestrian *This book*. The typical Czech description is shorter by one third; it has a two-part structure often corresponding to two paragraphs, with the introduction often combining two functional components. It closes either with readership or a special feature. It usually presents its key parts by *(Tato) Publikace/Kniha* (= *This/The book*).

7   Conclusion

When the idea to analyse book descriptions was first conceived, several things were envisaged. It was expected that while the Anglophone, highly professional publishers with a truly gigantic output would produce strictly regimented book descriptions according to house style templates, the Czech publishers' book descriptions would be somewhat loose and relaxed. It was assumed that it would be possible to present an elegant picture of features, both structural and others, clearly distinguishing between English and Czech descriptions. The results were expected to have not just theoretical significance, but also practical application. In other words, by

throwing light on the internal organization and linguistic properties of the original English and original Czech book descriptions, it was hoped that discourse analysis might contribute to the standardization of the Czech language descriptions, and provide help when they are translated into English.

As the results suggest these expectations were met only partially. There is not much that we find in one sample and not in the other, and so the differences are mostly of degree rather than kind. Also, these differences in structure and style are both between and within the Czech and the Anglophone publishers. Although the English descriptions are undoubtedly more homogeneous than the Czech ones, only two of the Anglophone publishers, Cambridge and Benjamins, appear to follow a relatively consistent policy of book description writing. In this sense, the effect of 'house style' in the English publishers is much weaker than expected (as far as book descriptions are concerned). All in all, there is little evidence of the influence of English on Czech book description style and the similarities with English book descriptions follow from the inner logic of the genre more than anything else. Even the dubious use of academic degrees before the names of authors occurring in Czech (and often found in English translations from Czech) did not differentiate between the two samples. All this makes the findings somewhat inconclusive – they suggest tendencies rather than taboos – and the well-worn but truthful statement that more data and research are needed applies even here. Still, having a rough idea of what a typical English book description looks like may provide at least some guidance when the need to translate a Czech book description into English arises.

References

Biber, D. (1995) *Dimensions of Register Variation. A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.
Biber, D., Connor, U. and Upton, T. A. (2007) *Discourse on the Move. Using Corpus Analysis to Describe Discourse Structure*. Amsterdam and Philadelphia: John Benjamins.
Bruce, I. (2008) *Academic Writing and Genre. A Systematic Analysis*. London and New York: Continuum.
Doró, K. (2013) 'The rhetoric structure of research article abstracts in English studies journals.' *Prague Journal of English Studies 2/1*, 119-139.
Hyland, K. (2004) 'Praise and criticism. Interactions in book reviews.' In: Hyland, K. *Disciplinary Discourses. Social Interactions in Academic Writing*. Ann Arbor: The University of Michigan Press. 41-62.
Lindholm-Romantschuk, Y. (1998) *Scholarly Book Reviewing in the Social Sciences and Humanities*. Westport, Connecticut and London: Greenwood Press.
Mann, W. and Thompson, S. (eds) (1992) *Discourse Description. Diverse Linguistic Analyses of a Fund-Raising Text*. Amsterdam and Philadelphia: John Benjamins.
Martin, J. R. and Rose, D. (2008) *Genre Relations. Mapping Culture*. London: Equinox.

Schiffrin, D., Tannen, D. and Hamilton, H. (eds) (2001) *The Handbook of Discourse Analysis*. Oxford: Blackwell Publishers.
Stotesbury, H. (2003) 'Evaluation in research article abstracts in the narrative and hard sciences.' *Journal of English for Academic Purposes 2/4*, 327-341.
Swales, J. (1990) *Genre Analysis. English for Academic and Research Settings.* Cambridge: Cambridge University Press.
Taboada, M. (2011) 'Stages in an online movie review.' *Text & Talk 31/2*, 247-269.
Tardy, C. M. (2011) 'Genre analysis.' In: Hyland, K. and Paltridge, B. (eds) *Continuum Companion to Discourse Analysis.* London and New York: Continuum. 54-68.

Internet sources of the Czech sample:

Karolinum, nakladatelství Univerzity Karlovy v Praze:
    http://www.cupress.cuni.cz/ink2_stat/, retrieved in 2014.
Nakladatelství lidové noviny: NLN, s.r.o:
    https://www.nln.cz/knihy/, retrieved in 2014.
Nakladatelství Academia:
    http://www.academia.cz/edice/lingvistika.html, retrieved in 2014.

Internet sources of the Anglophone sample:

Cambridge University Press:
    http://www.cambridge.org/cz/academic/subjects/languages-linguistics/, retrieved in 2014.
John Benjamins Publishing Company:
    https://benjamins.com/catalog/subjects retrieved in 2014.
Oxford University Press, USA:
    http://global.oup.com/academic/category/arts-and-humanities/?cc=cz&lang=en&, retrieved in 2014.