

Maria Balaskó

## What Does the *Figure* Show?

### Patterns of Translationese in a Hungarian Comparable Corpus

#### 1 Introduction

In the process of translation the translator is confronted with the task of re-expressing ideas formulated in another language by someone else using a new linguistic form. The constraints imposed on him or her leave traces on the newly formulated text and distinguish it in several ways from text originally created in the same language. Differences in the frequency or in the distribution of the linguistic items for example, are typical features which result in what has been known as *translationese* or the *third code*. Furthermore, it has been noticed that the frequency and the distribution of *recurring patterns* in translated and non-translated text, as well as their linguistic make-up, present dissimilarities equally. By examining and comparing the patterning of linguistic items which carry metatextual meaning in non-translated and translated academic texts the following paper aims to identify instances of *translationese*.

#### 2 On Patterns

The present investigation sets out from the view accepted by more and more linguists that "our language does not expect us to build everything starting with lumber, nails, and blueprint, but provides us with an incredibly large number of prefabs" (Bolinger 1976: 2). The validity of this statement has been attested by studies in second language acquisition (Weinert 1995) and in psycholinguistics (Pawley/Syder 1983, Kuiper/Haggo 1984), while findings in corpus linguistics (Sinclair 1991, Kjellmer 1994, Altenberg/Eeg-Olofsson 1990, Moon 1998) highlight the degree to which our language production leans on conventional patterns. The term *pattern* is used here in the sense of "a repeated configuration of grammatical and lexical items around a node lexical word" (Mauranen 2000: 131). In this context, Sinclair's idiom principle postulates that "words do not occur at random in a text" (Sinclair 1991: 110) and that "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments" (Sinclair 1991: 110). The semi-preconstructed phrases Sinclair is referring to are not those fairly rare fixed combinations of words which build up phraseological units but looser word combinations showing frequent co-occurrence. These frequent combinations of words – or multiword lexical units – while common and typical, are not

striking from a psychological point of view, and so they pass almost unnoticed in everyday speech situations. As Sinclair and Renouf (1988) put it:

The human being, contrary to popular belief, is not well organised for isolating consciously what is central and typical in the language; anything unusual is sharply perceived, but the humdrum of everyday events are appreciated subliminally. (Sinclair/Renouf 1988: 151)

So inconspicuous as they might be, these multiword units seem to be nevertheless the typical, common and preferred building blocks of written and spoken language. Hanks (1996) takes up the point showing that corpus data are a clear evidence of the fact that language relies heavily on conventional, common, usual and typical patterns:

The creative potential of language is undeniable, but the concordances to a corpus remind us forcibly that in most of our utterances we are creatures of habit, immensely predictable, rehearsing the same old platitudes and the same old clichés in almost everything we say. If it were not so, language would become unworkable. Humankind cannot bear much creativity. (Hanks 1996: 85)

Research on corpus data in descriptive translation studies has backed the idea that the context of its production makes translated text differ from authentic text. If we accept that the language as a whole is made up in greater part of "prefabs" or typical patterns, and also the idea that the text is shaped by the aim for which language is used and by the context in which it is used, then both non-translated and translated language must have their own typical patterning which will have to differ in several respects.

Approaching the problem from a different angle, the argumentation leads to a similar conclusion. Translation is characterised by ongoing confrontation at two levels: firstly, the linguistic systems of the languages involved in the process of translation collide, and secondly the text taking shape in the process of text creation fights against the constraints imposed on it by the source text, an existing end-product. The output of this sophisticated confrontation is translated text which carries concomitantly the features of both source and target language. If this is a characteristic of translated text in general, the statement has to be valid also with respect to the typical building blocks or patterns of translated text. If the process of translation did not leave traces on translated text, texts comparable in terms of genre and topic for example, would have to use the same building blocks in their text production, irrespective of whether they are the result of authentic text production or of translation. There is more and more evidence however that the 'prefabs' used to formulate translated text are different: translations contain patterns of both source and target language (in respect of lexis and linguistic structure), as well as patterns which are a result of the combination of the two (e.g. differences in frequency).

### 3 Redefining Translationese

*Traduttore – tradittore; traduction – trahison; ein übersetztes Buch – ein verletztes Buch*: these sayings have become commonplace in most cultures affected by translation and reflect a fairly negative attitude towards translated text in that they

express the view that translations may only be misrepresentations of original, that is, non-translated texts. This way of reasoning is responsible for the fact that translations have long been viewed as second-rate texts and as such not suitable for, not even worthy of scientific enquiry. Even two decades ago, the newly emerging discipline corpus linguistics reinforced this idea by deliberately excluding translations from the various corpus projects on the grounds that they did not represent "real" language, but only distorted our image of it. While it must be admitted that the process of translation does indeed leave traces on translated text making it differ from non-translations, the position stated above is extremely biased and therefore unacceptable.

The features which characterise languages in contact have been well documented in the literature on second language acquisition and have been termed interference, interlanguage (Selinker 1972), quasi-correctness (Papp 1984), lack of naturalness (Sinclair 1984), while in the context of translation a similar phenomenon has come to be known as the *third code* (Frawley 1984) or *translationese* (Toury 1980, Gellerstam 1986):

Theoretical considerations highly verify it, and even lead to hypothesising that the language used in translation tends to be interlanguage (sometimes designated "translationese"), or that a translation is, as it were, an "inter-text" *by definition*. (Toury 1980: 75)

In other words: "the translation itself [...] is essentially a third code which arises out of the bilateral consideration of the matrix and target codes: it is, in a sense, a sub-code of each of the codes involved" (Frawley 1984: 168).

While interlanguage often becomes manifest in formal equivalencies which do not exist in the target language as deviation from the target *code*, translationese is primarily characterised by the occurrence of unusual forms as deviation from the *norm of language use*. If, for example, in a target text the frequency of certain linguistic items or forms which have formal counterparts in the source language is higher or lower than is usual in the target language, it may mark a particular text as a translation.

The term *translationese* has often been used to refer to the difference in the distribution of particular lexical items or to the translated version of the target language (see for example Gellerstam 1986). Yet translationese is a complex phenomenon which manifests itself on each and every level of the linguistic system, and even beyond. The language of translations does not bear only features of source language, however. Like interlanguage in second language acquisition, translation also shows universal traits – that is "features which typically occur in translated text rather than original utterances and which are not the result of interference from specific linguistic systems" (Baker 1993: 243). Such features are present even in translations which House (1977) has termed *covert*. In contrast to *overt* translations, which can easily be identified as such, covert translations sound like natural target language texts. Research (e.g. Laviosa 1997) has proven that even these texts carry features which differ from those of genuine texts in the same language.

That translation is qualitatively different from authentic text production – although they both belong to the same linguistic code – is an idea which has received pro-

minence in recent translation studies. More and more studies support the view that there is only partial overlap between the two modes of text production. Correspondence is found where translated text reads like authentic text, and what lies beyond the common area belongs to the *third code*.

Used in a non-evaluative way, then, translationese refers to all those features, overt or covert, at each level of the linguistic structure which distinguish translated text from original, genuine language. The third code is used as a synonym while others like interlanguage or quasi-correctness are ignored because of their implicit evaluative stance. In line with Toury (1980) and Frawley (1984) who claim that translationese is the product of the translational process itself, which results from the confrontation of the source and target language under circumstances specific to the process of translation, translationese is regarded here as a particular code with its own characteristic features. By no means does the term cover faulty translations, but the regular influence of the source language on the target language.

Translationese is made up of two sets of components: one set consists of features of target language which behave differently from what is typical of target language usage. The second is made up of what has been referred to in descriptive translation studies as universals of translation – e.g. explicitation, simplification, normalisation. The third code as defined here is illustrated by a slightly adapted version of Schmiel and Schäffler's model (1996: 44):

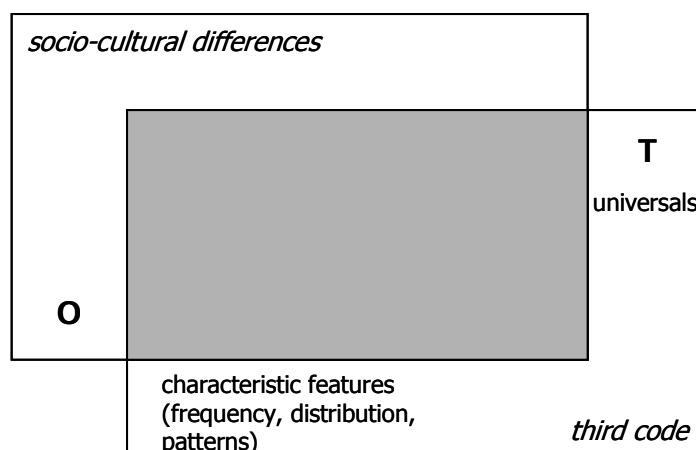


Figure 1: The domain of the third code (O: original language, T: translated language)

The idea that the language of translation differs from "normal" language has been present in translation studies for some time now. Nevertheless, until the emergence of corpus linguistics in the study of translation, only a few isolated studies have attempted to unveil the nature of the third code. According to Chesterman (1998) the most important change in today's translation studies is the changeover from philosophical conceptual analysis to empirical research. Wilss (1996) even talks about the re-empirisation of the discipline, which in his view is an answer to the "pragmatic constraints" that almost all disciplines have to face nowadays. The emergence of

corpus linguistics and the study of corpora have contributed a great deal to the re-empiricism of translation studies. With access to huge amounts of data and with the methods and the linguistic approach characteristic of corpus linguistics, the study of the language of translation can be approached in a novel way.

#### 4 A Corpus-driven Look at *ábra* 'figure'

In what follows, a corpus of comparable academic texts will be put under scrutiny in order to identify instances of translationese. This paper, which is part of a larger study on the behaviour of a series of key words which carry metatextual meaning, is focused on *ábra* 'figure', the first metatextual lexical item in the alphabetical word list. Metatext as understood throughout this paper has the function of indicating the organisation of the text, of providing comments on the referential or propositional content of the text, or of serving as a guide to the reader (Mauranen 2000: 120).

##### 4.1 The Corpus

Translationese will be investigated in a corpus which, since it is still under development, is fairly small. The one-million-token Hungarian Translational Corpus (Magyar Fordítási Szövegtár) consists of three subcorpora: one of original Hungarian texts (EHU), another of original English texts (EAN), and still another of the Hungarian translations of the English originals (FHU). The component texts are from similar domains (academic texts), are similar in length, and cover a similar time span (published after 1980). The analysis is carried out using WordSmith Tools software (Scott 1998).

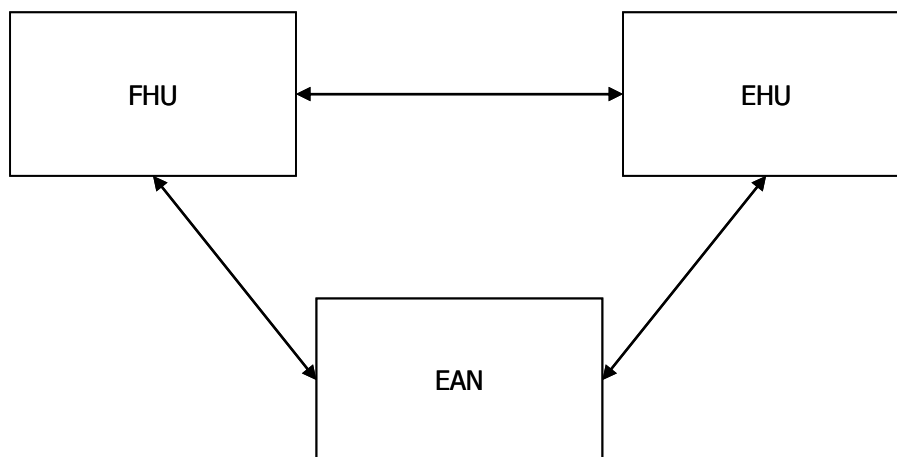


Figure 2: The structure of the Hungarian Translational Corpus (Magyar Fordítási Szövegtár)

##### 4.2 A Note on Corpus-driven

Starting with the 1990s, corpus linguistics has seen the separation of two distinctive trends: the corpus-based and the corpus-driven approaches (Francis 1993, Hunston/Francis 2000, Tognini-Bonelli 2001). This differentiation was brought about by the roles

assigned, respectively, to corpora and corpus data: in the corpus-based approach the corpus is a collection of examples used to test or exemplify given theoretical statements. By contrast, the corpus-driven approach views corpus data as the starting point of a theory-independent investigation. The theoretical statement is to be formulated in the presence of corpus evidence and is fully accountable to it. The linguistic categories are based on recurring *language patterns* and on the frequency of their occurrence.

Empirical and bottom-up *per definitionem*, while searching for universals, descriptive translation studies has undergone a transformation. Still remaining within a descriptive framework, it has become top-down in its attempt to identify, occasionally even to fish for the universals of translation. Corpora have often been used to prove the existence of phenomena like explicitation, simplification, normalisation and the like. Yet this approach is not in accordance with the aim of descriptive translation studies which is to *describe* the language of translation. A corpus-driven approach – as opposed to a corpus-based one – combined with a bottom-up methodology is more suitable in this endeavour.

The study of translation with the help of corpora, termed at times corpus translation studies, was made possible not only by access to large amounts of data. It was also enabled by the fact that this field of study is located at the convergence of descriptive translation studies and corpus linguistics – two disciplines with similar conceptual frameworks in that both regard description rather than prescription as their primary task, both are interested in what is typical in the language and in recurrent patterns, and both hold the same view about the role of theory: “If theory cannot account for every phenomenon that actually occurs in translation, what is defective, and what should be corrected, is the theory and not the ‘problematic’ phenomenon” (Touy 1980: 62).

#### 4.3 A Note on Hungarian

Hungarian is a basically agglutinative language, in other words grammatical relations are expressed through affixes. Hungarian nouns for example consist of a stem followed by three slots each of which can be filled by an inflectional suffix. Nouns inflect for number, person (possessor) and grammatical case, with the relevant suffixes attached in that order (Abondolo 1992). The case system is very complex and involves 16 (in other opinions 24) different forms which express grammatical and/or semantic relations. Word composition and derivation are very common and highly productive strategies in present-day Hungarian. These features prevent Hungarian from being straightforwardly accessible for automatic analyses.

#### 4.4 The Word Form *ábra* ‘figure’

The presence of *ábra* ‘figure’ is statistically significant in both subcorpora: with the derived forms, including compound words, it has 153 occurrences (0.01%) altogether in the translational subcorpus (FHU) and 122 (0.02%) in the original Hungarian subcorpus (EHU).

FHU				EHU			
49	ÁBRA	33	0,01	46	ÁBRA	67	0,02
50	ÁBRAALÁÍRÁS	1		47	ÁBRÁBA	1	
51	ÁBRÁBA	1		48	ÁBRÁBAN	1	
52	ÁBRÁBAN	5		51	ÁBRÁIT	1	
53	ÁBRÁK	2		52	ÁBRÁJA	2	
54	ÁBRÁN	26		53	ÁBRÁJÁT	1	
55	ÁBRÁT	2		54	ÁBRÁJÁVAL	1	
56	ÁBRÁZOL	1		55	ÁBRÁK	2	
57	ÁBRÁZOLÁS	8		56	ÁBRÁKBAN	2	
58	ÁBRÁZOLÁSA	8		62	ÁBRÁT	5	
59	ÁBRÁZOLÁSÁHOZ	1		63	ÁBRÁVAL	2	
60	ÁBRÁZOLÁSÁI	1		64	ÁBRÁZOLÁS	3	
61	ÁBRÁZOLÁSÁIK	1		65	ÁBRÁZOLÁSA	2	
62	ÁBRÁZOLÁSÁIN	1		66	ÁBRÁZOLÁSÁBAN	1	
63	ÁBRÁZOLÁSÁKOR	1		67	ÁBRÁZOLÁSÁI	1	
64	ÁBRÁZOLÁSÁN	1		68	ÁBRÁZOLÁSÁNAK	1	
65	ÁBRÁZOLÁSÁNAK	2		69	ÁBRÁZOLÁSÁT	1	
66	ÁBRÁZOLÁSÁRA	4		70	ÁBRÁZOLÁSÁVAL	1	
67	ÁBRÁZOLÁSÁT	2		71	ÁBRÁZOLÁSÁBÓL	2	
68	ÁBRÁZOLÁSÁVAL	2		72	ÁBRÁZOLÁSI	1	
69	ÁBRÁZOLÁSÁBAN	2		73	ÁBRÁZOLÁSOKBÓL	1	
70	ÁBRÁZOLÁSÁBÓL	1		74	ÁBRÁZOLHATJUK	2	
71	ÁBRÁZOLÁSI	2		75	ÁBRÁZOLHATÓ	1	
72	ÁBRÁZOLÁSMÓD	1		76	ÁBRÁZOLJA	2	
73	ÁBRÁZOLÁSMÓDNA+	1		77	ÁBRÁZOLJÁK	1	
74	ÁBRÁZOLÁSOK	2		78	ÁBRÁZOLJUNK	1	
75	ÁBRÁZOLÁSOKRA	1		79	ÁBRÁZOLNAK	1	
76	ÁBRÁZOLÁSTÓL	1		80	ÁBRÁZOLNI	2	
77	ÁBRÁZOLHATJUK	2		81	ÁBRÁZOLÓ	6	
78	ÁBRÁZOLHATÓ	2		82	ÁBRÁZOLT	3	
79	ÁBRÁZOLHATÓK	1		83	ÁBRÁZOLTA	1	
80	ÁBRÁZOLJA	6		84	ÁBRÁZOLVA	3	
81	ÁBRÁZOLJÁK	2					
82	ÁBRÁZOLJON	1					
83	ÁBRÁZOLJUK	3					
84	ÁBRÁZOLNI	3					
85	ÁBRÁZOLÓ	1					
86	ÁBRÁZOLÓDTHAT	1					
87	ÁBRÁZOLÓDIK	1					
88	ÁBRÁZOLT	4					
89	ÁBRÁZOLTA	6					
90	ÁBRÁZOLTAK	1					
91	ÁBRÁZOLTÁK	1					
92	ÁBRÁZOLTUK	1					
93	ÁBRÁZOLUNK	1					
94	ÁBRÁZOLVA	2					

Table 1: The distribution of the lemma ÁBRA in the FHU and the EHU (number of occurrences; type/token ratio)



The alphabetical list clearly shows that the paradigm of the noun is richer in the original Hungarian subcorpus (EHU) than in the subcorpus of translated Hungarian (FHU). The different oblique case forms have at most two occurrences. The FHU subcorpus, on the other hand, contains the type *ábrán* 'in the figure' with quite a few occurrences, which is even more striking since the type is absent from the EHU subcorpus altogether. A look at the concordance of the more frequent forms will give a clearer picture of their behaviour.

1	PONTjakként fogja használni.	8.	<b>Az</b>	1. ábra	<b>azt mutatja</b> , ahogyan a megmunkáló v	
2	atók az őket jelző vonalak, amint	<b>az</b>	1. ábra	felső részében	<b>látható</b> . A műveletek	
3	csatlakoznak a szövegvilág-modell (11. ábra) csomópontjai a séma (13. ábra) cso					
4	zín-állapotba juttatja a rakétát.	<b>A</b>	13. ábra	a 'repülés'-sémát	<b>mutatja</b> hálózati	
5	ell (11. ábra) csomópontjai a séma (13. ábra) csomópontjaihoz.	26.	A séma hat			
6	ra szolgáló hathatós mechanizmus.	<b>A</b>	14. ábra	képszerűen	<b>szemlélteti</b> , hogyan csat	
7	, de fentről lefelé szedünk szét).	<b>A</b>	2. ábra	<b>mutatja</b> a példánkban szereplő főnév		
8	tlenül a névelőt (a 'egy') követi.	<b>A</b>	3. ábra	<b>mutatja</b> az igei csoport elemzését á		
9	függőségi viszonyokat képviselik.	<b>A</b>	4. ábra	eszerint	<b>mutatja</b> a hálózatot. Egy i	
10	yen hálók lazák és szétszórtak.	[5.1	ábra]	Egyszerű hálóviszonyok	Ez a	
11	és a maradvány területektől is.	[6.1	ábra]	A rajnai legyező	Az egyes fone	
12	a nyelv standard változatával.	[6.2	ábra]	Egymást kereszterő izoglosszák		
13	r-es ejtés használatának vágya.	[7.1	ábra]	Az (r) százaléka; (r] a fourth		
14	emben bármely két átlag eltérése.	<b>A</b>	7.2	ábra	áthidalása lehet számítási hibák er	
15	és a szópároknál is előfordul.	[7.2	ábra]	r-es ejtés New York Cityben tár		
16	álata az adott beszédstílusban.	[7.3	ábra]	A thing-ben, three-ben stb. lév		
17	detroiti munkásosztály között.	[7.4	ábra]	Az egyes szám harmadik személy,		
18	lálható különbség szignifikáns.	[7.5	ábra]	Az [r] hiánya olyan szavakban,		
19	dul elő a (bebor)ban ('Vágd!').	[7.6	ábra]	A hasonult magánhangzók szándék		
20	lyamatnak az időbeli lezajlása.	[8.1	ábra]	A lexikai diffúzió időbeli leza		
21	ás. (Fel kell figyelniük a 8.1 és a 8.2	ábra	hasonlóságára.)	[8.2	ábra]	Le
22	1 és a 8.2 ábra hasonlóságára.)	[8.2	ábra]	Lexikai diffúzió az egyéni besz		
23	bekezdésnek a modelltere nagyjából	<b>a</b>	9. ábra	szerinti. A 'rakéta' 'emelkedő' moz		
24	n látható megoszlást tapasztalta.	<b>Ez az</b>	ábra	<b>azt mutatja</b> , hogy a Saksnál jobban		
25	oszlopban, maguk az állapotok pedig	<b>az</b>	ábra	közepén. Amikor a fejet megtalálta,		
26	erzsében informális beszéd esetén	<b>Az</b>	ábra	<b>mutat átfedést</b> a nemek között, de a		
27	brában foglalta össze.	[Ábra]	<b>Az</b>	ábra	tehát M.A.K. Halliday nyelvészettel	
28	ogalmaz meg vele kapcsolatban.	<b>A</b>	7.2-es	ábra	<b>azt mutatja</b> , hogyan használják a kü	
29	önböző területet jelent, <b>ahogy a</b>	6.2-es	ábra	<b>mutatja</b> : (r) és (u); (r) és (A); 0		
30	evez, akik erre a döntésre jutottak.	[Ábra]	Látható, hogy az alexandriaiak			
31	ó sematikus ábrában foglalta össze.	[Ábra]	Az ábra	tehát M.A.K. Halliday		
32	mok a megfelelő szűtrákra utalnak):	[Ábra]	Csak az utolsó ábrázolás az ön			

Table 2: Concordance to *ábra* from the FHU subcorpus



1 családon belüli nyelvválasztása]	[1. ábra. Magyarországi (kétegyházi) románok
2 -as tanácsai minősítés szerint)]	[10. ábra. Szlovák nemzetiség Magyarországon
3 forrás: Arday-Hlavik 1988: 62.)	[11. ábra. Délszláv nemzetiségek (horvátok, s
4 forrás: Arday-Hlavik 1988: 64.)	[12. ábra. Román nemzetiség Magyarországon (1
5 olyamatát <b>így is ábrázolhatjuk:</b>	[13. ábra. A nyelvváltás folyamata] A nyelv
6 t Skutnabb-Kangas 1983 <b>alapján!</b>	[14. ábra. A kétnyelvűség foka, típusa és kog
7 in nyelvjött létre s van ma is.	[15. ábra. A pidzsin és a kreol nyelvek terül
8 kínálta nyelvüvelői javakból.	[16. ábra. A rádió nyelvüvelői műsorainak hal
9 esetében viszont fordítva van.	[17. ábra. Nyelvek gazdasági és demográfiai s
10 elemzés <b>összefoglaló ábrázolását 1.</b>	[18. ábra.) A nem és az életkor nagyban befo
11 t (a főnt említett okok miatt).	[18. ábra. Egy vállalathoz tartozó, nem egy s
12 55 év közötti emberek beszélnek.	[2. ábra. Életkor- és nemzedékspecifikus nye
13 mmunikációs modellt (1972: 234-244):	[Ábra] <b>Ez a séma így értelmezhető: a</b>
14 <b>zöképpen szemléltethető</b> (1981: 248):	[Ábra] <b>Amint ebből a szemléltetésből</b>
15 ányak is szólnak" (Aczél 1984: 25).	[Ábra] <b>Jól látható, hogy ebben a költ</b>
16 Deme 1975: 122, Wacha 1978: 22, 28):	[Ábra] Az elhangzó szöveg szinte mind
17 az orosz volt a hivatalos nyelv.	[3. ábra. Életkori különbségek a litvánok li
18 b társadalmi helyzetét <b>mutatják.</b>	[4. ábra. Nemek szerinti különbségek középső
19 ználatának a köznyelvesítésében.	[5. ábra. Életkor és nemek szerinti különbsé
20 zt egy sydney-i vizsgálat <b>ábrájával</b>	(5. ábra) <b>szemléltetjük, amely fiatalabb és</b>
21 on't segédige használata alapján. <b>A</b>	6. ábra hat-, nyolc- és tízéves lányok és f
22 ] (A forrás: Coates 1993: 151.)	[6. ábra. Nemek szerinti különbségek gyermek
23 igény(esség) alapján (1979a: 64-65):	[Ábra] A szöveg mint teljes halmaz -
24 pedig szakmunkásképzőbe járnak.	[7. ábra. Magyarországi középiskolások iskol
25 őleg milyen környezetben nő föl.	[8. ábra. Az anyanyelvi és a matematikai tel
26 eti elhelyezkedése és számaránya	[9. ábra. Német nemzetiség Magyarországon (1
27 gyik osztályozási lehetőségét <b>az alábbi</b>	ábra <b>mutatja. Az osztályozás kritériumai</b>
28 <b>gy foglалható össze:</b> [Ábra] <b>Ez az</b>	ábra bonyolult és részletező tipizálási
29 nyelvi kiejtést, a fiúk nem egészen. <b>Az</b>	ábra másik tanulsága az, hogy az iskolai
30 em köznyelviessen (au, illetőleg u:). <b>Az</b>	ábra világosan <b>mutatja</b> a nemek szerinti
31 • A grammatikai jellemzők [Ábra] [Ábra]	Nyelvspecifikus sajáttság, hogy
32 at szimbolizáló szarv <b>ábrázolásával:</b>	[Ábra] 3. A látható nyelv a költészet
33 Sematikusan <b>így lehetne ábrázolni:</b>	[Ábra] • Többfelé ágazó réma [Szöv
34 ét és szinteződését <b>mutató ágrajzon:</b>	[Ábra] Az igeneves szerkezetek felold
35 újdonsága révén a beszéd folyamatot:	[Ábra] A magyar mondatok egy része -
36 ra. <b>Következzék most szemléltetésül</b> egy	ábra! A litván lakosság anyanyelvi és or
37 gét, mint a szöveg többi egységében:	[Ábra] Ha gondolatban összehasonlítju
38 ze a szöveg és kép együttesére épül:	[Ábra] Az élet napos oldala metaforik
39 an <b>ábrázolni</b> a bekezdés felépítését:	[Ábra] Valójában ezt az ábrát csak s
40 efoglalva a <b>következőképpen festene:</b>	[Ábra] Természetesen ezt a szövegfelé
41 g aktív olvasása (Ingarden fogalma).	[Ábra] Ebben a versben a sarló alakú
42 pedig a személynév más 0 formája:	[Ábra] Ha a téma különböző mondatrés
43 ekciója a költői morfondírozás jele:	[Ábra] A kritikai kiadás adatai szeri
44 csolatban áll a szöveg jelentésével.	[Ábra] Ez a versrajz vagy rajzos vers
45 a jelentést vagy jelentésváltozatot:	[Ábra] Ebből a sokváltozós "rejtvényb
46 i pályán. • A grammatikai jellemzők	[Ábra] [Ábra] Nyelvspecifikus saját
47 lesztetők közös metszetbe kerülnek:	[Ábra] Az olvasmány szó jelentésmezéj
48 i fizikai dolgozók voltak. <b>A következő</b>	ábra magyarországi középiskolások szülei
49 s: Haarmann 1988: 1667.) <b>A következő</b>	ábra pedig <b>azt mutatja</b> , hogy a mai magya
50 jelszerveződésű szöveg a <b>következő:</b>	[Ábra] Az újság a nyelvi közlést az o
51 "az elhangzott beszéd átiratában külön	ábra nélkül, a betűsoron alkalmazva, egy
52 tőből a <b>következőképpen vezetik le:</b>	[Ábra] Ezek a számok Panini nyelvtaná
53 a költő <b>példát is mutat a másodikra:</b>	[Ábra] 3.2.2. Konkrét költészet Az ö
54 elemek is jelzik: egyszer másodsor:	[Ábra] Ebben az esetben szövegépítkez
55 ségét akarják ezáltal megvalósítani:	[Ábra] 5. A multimediális szövegek ér
56 ágyazva s külön jellel (négyzettel):	[Ábra] c) A szociolingvisztikában gya
57 jra, viszont a befogadó odagondolja:	[Ábra] Kitűnt az ábrázolásokból, hogy
58 gazási <b>ábrában így foglалható össze:</b>	[Ábra] Schweiger a standard szöveget
59 mba. Rendszere <b>így foglалható össze:</b>	[Ábra] Ez az ábra bonyolult és részle
60 kobson egy-egy alapfunkciót rendelt:	[Ábra] A jakobsoni kommunikációelméle
61 óját a <b>következőképpen szemlélteti:</b>	[Ábra] (Fülei-Szántó, 1981: 249.)
62 k (x) szerint felépülő szövegegység:	[Ábra] Ez a bekezdés nem "nyitott" ke
63 éseket (Deme 1974: 114): [Szöveg]	[Ábra] Ez a gondolat nyomatékos formá
64 ő részeit árnyaló elemek találhatók:	[Ábra] 3.2 A szövegmondatok kapcsolód
65 vagy a változatlan elem törlésével:	[Ábra] • Fejlődés levezetett témával:
66 s kérdésre a <b>kép adja meg a választ:</b>	[Ábra] A következő Kossuth-nóta szöve
67 krét helyzetekben való viselkedésre.	[Ábra] A nyelvi attitűdök - első pill

Table 3: Concordance to *ábra* from the EHU subcorpus

The concordances show that while the type *ábra* is a mere caption in most cases, its co-text and the context in which it appears are different in the two subcorpora. In the greatest part of the concordance lines from the FHU subcorpus, *ábra* refers to concrete, numbered pictures and is determined by the definite article (*az*). The verb *mutat(ja)* or its synonym *szemlélteti* 'show' is repeated several times, building up a recurrent pattern: *az (x) ábra mutat(ja) 'the figure (x) shows'*. In the concordances, the figure referred to in the text unveils itself as an organic part of, or as an active participant shaping the argumentation.

The EHU subcorpus does not reveal a similar recurring pattern. Nevertheless we can observe a loose pattern containing interchangeable elements best represented in concordance line 36: *következzék most szemléltetésül egy ábra* 'here follows as illustration a figure' – variations 5, 14, 28, 33, 40, 52, 57, 58, 61. The use of *ábra* 'figure' in the original Hungarian texts is different from the typical behaviour of *ábra* in the FHU subcorpus of translated Hungarian. *Ábra* 'figure' is expressed as the conclusion of something mentioned earlier: it is the thing being referred to, that is being explained or, alternatively, something which is being exemplified. The word form *ábra* usually follows something that is signalled in writing with a colon (:). If in the translated Hungarian subcorpus the type was found to belong to the argumentation, in the original Hungarian subcorpus it seems to be merely a supplement to, or a complement of the argumentation.

#### 4.5 *Ábrán* 'in the figure'

The alphabetical list reveals the absence from the EHU subcorpus of the word form which has the highest number of occurrences in the FHU subcorpus.

1 ormán kötődnek-e egymáshoz, **mint az** (1) ábrán, szorosan kötődnek, de úgy, hogy a  
2 potok rendszereként, **mint amilyen az** 1. ábrán a főnévi csoporté volt. 10. Ami  
3 tát a nagy magasságban, megkapjuk a 10. ábrán **látható** hálózatot. 37. Az utols  
4 egvilág végső modelljét, **amint az a** 11. ábrán **látható**. Figyeljük meg, hogy ezútt  
5 gfelelője olyasmí lehet, **mint ami a** 12. ábrán **látható**, ha megpróbáljuk a 11. ábr  
6 lí kapcsolat a legfontosabb, **mint a** (2) ábrán; gyengén kötődnek, s e kötődésben  
7 ben A az egyetlen kapcsolat, **mint a** (3) ábrán, vagy a kapcsolat A és E között C-  
8 t C-n keresztül valósul meg, **mint a** (4) ábrán? Többszörös egy háló akkor, ha eg  
9 ogalmi hálózatot a IV. 10. pontban **a** 4. ábrán **látható** nyelvtani hálózattal. Bár  
10 g lefelé, majd fölfelé mutatóval, **az** 5. ábrán **látható** négyféle minta lehetséges  
11 hálóban kapcsolódik össze, **mint az** 5.1 ábrán, egyformán kötődnek-e egymáshoz, m  
12 hez juthatunk, ha összehasonlítjuk **a** 6. ábrán **látható** fogalmi hálózatot a IV. 10  
13 . Mindezeket a fogalmi viszonyokat **a** 6. ábrán **látható** hálózatban ábrázolhatjuk.  
14 egyezőszerűen szétterülnek, **ahogy a** 6.1 ábrán **látható**. A legyezőn belül minden t  
15 asági helyzet kategóriái azonosak **a** 7.2 ábrán **használtakkal**) A grafikonról La  
16 rtok nyelvi viselkedésében, **ahogy a** 7.2 ábrán **látható**. Az erős törést mutató meg  
17 afikonokat mutassanak be, **amilyet a** 7.2 ábrán **látunk**: eredményeiket bemutató gra  
18 középső pedig közbülső változat. **A** 7.3 ábrán **látható**, **hogyan** oszlanak meg a (th  
19 k, hogy az [r] hiányának esetében **a** 7.5 ábrán lejtős rétegzettséggel **van dolgunk**  
20 110. p.). A szükséges információt **a** 7.5 ábrán **találhatjuk meg**. A statisztikai el  
21 tatkoznak kevésbé, mint a nőknél. **A** 7.6 ábrán a 7.7 táblázat adatai vannak máské  
22 adék 20 százaléka is megváltozik. **A** 8.1 ábrán **látható** ennek a lexikai diffúziós  
23 lvasásakor) (Labov, 1966, 240. p.). **Az** ábrán **láthatjuk**, hogy az r használatát n  
24 itani a dologgal kapcsolatban. **A** 7.2-es ábrán **látható** vonalak csoportátlagokat r  
25 megismételt válaszként), Labov **a** 7.1-es ábrán **látható** megoszlást tapasztalta. Ez  
26 l pedig a többszörös tagadást. **A** 7.4-es ábrán például az egyes szám harmadik sze

Table 4: Concordance to *ábrán* from the FHU subcorpus

The concordance lines clearly reveal a multiword pattern:

mint / amilyen / a(z) x ábrán látható / láthatjuk 'as shown in figure x'

This pattern is obviously very typified in the FHU subcorpus. However both the word type and the multiword pattern are completely absent from the authentic Hungarian subcorpus. Since it is assumed that this pattern is the product of the translation process, we have carried out a concordance of the potential initiator of this type from the EAN subcorpus of original English texts:

1 CENTRE for the whole macro-state. 8. Figure 1 **shows** the processor moving along  
2 the rocket at great heights, we obtain Figure 10. 37. The single sentence of  
3 giving us the final text-world model **in** Figure 11. Notice that we have to use ar  
4 or our sample might look **something like** Figure 12, if we try to keep the same ba  
5 he rocket into a new state of location. Figure 13 **shows** the 'flight'-schema as a  
6 ng general strategies to current tasks. Figure 14 **offers a graphic illustration**  
7 order and removed in the reverse order. Figure 2 **illustrates** the order for stack  
8 eterminer ('a') immediately ('desert'). Figure 3 **shows** the parsing out of the ve  
9 tes and the links are the dependencies. Figure 4 **illustrates** the network in this  
10 for the same two part utterance, **as in** Figure 5 (examples from Brazil 1975: 6)  
11 D, and E are linked in a network, **as in** figure 5.1, are they all equally linked  
12 ns into a network such as that **shown in** Figure 6. Link labels announce the type  
13 Rhine. At that point they 'fan', **as in** figure 6.1. Each area within the fan has  
14 the four distinct areas **illustrated in** figure 6.2: [r] and [u]; [r] and [A]; O,  
15 above found the distribution **reported in** figure 7.1. This figure shows that r-pro  
16 tistical sense. The cross-over **shown in** figure 7.2 could conceivably be the resu  
17 necessary. Such displays **as we find in** figure 7.2 are displays of group means.  
18 rticularly strong claims concerning it. Figure 7.2 **shows** the use of r by various  
19 ent us with displays **such as the one in** figure 7.2, displays which show their re  
20 d middle-class groups, **as we can see in** figure 7.2. He calls a distribution with  
21 three, etc. in New York City. SEC **as in** figure 7.2.] Finally, what is impress  
22 on what he calls the (th) index, **as in** figure 7.3. The higher the index score,  
23 boy's, and goes; and multiple negation. Figure 7.4, for example, **shows** group mea  
24 th [r] absence (Wolfram, 1969, p. 110). Figure 7.5 **gives** us the necessary inform  
25 rgue that in the case of [r] absence **in** figure 7.5 we have an example of what th  
26 le less than do females. The display **in** figure 7.6 reorganizes the data of table  
27 rackets, we obtain the pattern given **in** Figure 8. Again, the links have labels e  
28 cent of relevant words show the change. Figure 8.1 **shows** this process of lexical  
29 . (Note the similarity of figure 8.2 to figure 8.1.) [Figure 8.1. Lexical dif  
30 half the words. (Note the similarity of figure 8.2 to figure 8.1.) [Figure 8.  
31 nd another set will show both variants. Figure 8.2 **shows** how such a process will  
32 his paragraph might look **something like** Figure 9. The 'rising' motion of the 'ro  
33 larity of figure 8.2 to figure 8.1.) [Figure 8.1. Lexical diffusion over time  
34 ast to either a focal or relic area. [Figure 6.1. The Rhenish Fan] The iso  
35 link from A to E achieved through C? [Figure 5.1. Simple network relationship  
36 e class and the upper working class. [Figure 7.5. Percentage of [r] absence i  
37 ocial class, the upper middle class. [Figure 7.2. R-pronunciation in New York  
38 litically, socially, and culturally. [Figure 6.2. Intersecting isoglosses]  
39 and working-class people in Detroit. [Figure 7.4. Percentage of [z] absence i  
40 e 8.1. Lexical diffusion over time ] [Figure 8.2. Lexical diffusion by indivi  
41 in casual speech in Tehran Persian] [Figure 7.6. Percentage of vowels assimi  
42 desire to adopt such pronunciation. [Figure 7.1. Percentage of (r); [r] in f  
43 social and stylistic stratification. [Figure 7.3. Stylistic and social strati  
44 emarking [z]. A close inspection of the figure **shows** that, whereas it is quite p  
45 tribution reported in figure 7.1. This figure **shows** that r-pronunciation was fa

Table 5: Concordance to figure from the EAN subcorpus

The concordance of *figure* reveals two more or less stable multiword lexical units in the English texts: *figure x shows / illustrates* and *as (we can find / see) in figure x*. These patterns are almost literally equivalent with the patterns of the types *ábra* and *ábrán* in

the concordances from the FHU subcorpus of translated Hungarian: *figure x shows / illustrates* vs. *az (x) ábra mutat(ja)*, and *as (we can find / see) in figure x* vs. *(mint) (amilyen) / a(z) x ábrán (látható / láthatjuk)*. What we have here is a clear example of translated text taking over patterns from the English texts.

The concordances from the FHU subcorpus reveal the manifestation of translationese at two levels: at the lexico-grammatical level, the target text makes use of the patterns of the source language. Since they do not break the rules of the target language code, these patterns are acceptable in the target language; however, they are not natural in genuine language production. "Not natural" here means that, while grammatically correct, the frequency of usage is different from what is usual or typical. It is possible that – even if subliminally registered – the differences would have passed unnoticed had the concordance lines not revealed the patterns typical of the two subcorpora.

At a textual level, translationese manifests itself in the building up of scientific text. As mentioned earlier, reference to *figure* in the translated subcorpus is part of the argumentation, of text dynamics, whereas in the original subcorpus it completes the text and closes down the argumentation with an emphasis. The metalinguistic reference has an inclusive/exclusive relation to the translated and original text respectively. This example seems to reveal a difference between English and Hungarian text production, between the ways the two languages build up discourse in academic texts – a problem which has not been tackled so far in Hungarian text linguistics. If awareness of the particular patterns is not raised, the translator oversees them even if in principle s/he is aware of the differences. This kind of argumentation raises the question of the methods of scientific inquiry as the examples suggest that the differences in making reference are the result of culturally different approaches to knowledge building: an inductive, bottom-up approach in the English-speaking cultural environment – which, in our examples, leaves obvious traces in the translated Hungarian texts –, and a deductive, top-down approach in the Hungarian cultural environment. Due to the small amount of data, the phenomenon described above may be accidental, therefore we cannot draw general conclusions at this stage. Nevertheless the hypothesis is worth further, in-depth investigation.

## 5 *Ábra* and Its Synonyms

A look at the different forms of the lemma *ÁBRA* reveals that the number of nominalised verb forms is not evenly distributed between the original and the translated corpus components: while in the translational subcorpus there are 39 occurrences with a rich paradigm, the original Hungarian subcorpus contains only 11 occurrences with an unproductive paradigm (see table 1). If we compare the distribution of the noun forms of the lemma *ÁBRA* in the two subcorpora by taking the form *ábrázolás* 'illustration' as the synonym of the form *ábra* (Kiss 1998), the picture of the distribution of the noun forms in the two subcorpora improves. Moreover, if we take into consideration additional synonymic forms listed in *Magyar szókincstár* (Hungarian thesaurus): *illusztráció, kép, ábrázolás* the ratio becomes even: in the translated Hungarian subcorpus the forms

*ábrázolás* and *ábra* cover the whole range of meaning whereas in the original Hungarian subcorpus the meaning is carried by several synonymic variants. It is to be noticed that the foreign loan word *illusztráció* 'illustration' is absent from the translated Hungarian texts, while in the EHU it is used in a natural way (see table 6).

FHU			EHU		
17770	KÉP	6	21498	KÉP	38 0,01
17771	KÉPALÁÍRÁSÁBAN	1	21499	KÉPALKOTÁSI	1
17772	KÉPBEN	1	21500	KÉPBE	1
17773	KÉPE	1	21501	KÉPBEN	8
17774	KÉPÉBE	1	21502	KÉPBŐL	1
17775	KÉPECSKÉJÉBŐL	1	21503	KÉPE	25
17776	KÉPECSKÉK	1	21504	KÉPÉBŐL	1
17777	KÉPEK	4	21505	KÉPEGÉSSZEL	1
17778	KÉPEKBŐL	1	21506	KÉPEGÉSZBŐL	1
17779	KÉPEKET	3	21507	KÉPEI	1
17780	KÉPEN	1	21508	KÉPEINK	2
			21509	KÉPEIT	1
15556	ILLUSZTRÁLJA	3	21510	KÉPEIVEL	1
15557	ILLUSZTRÁLT	1	21511	KÉPEK	8
			21512	KÉPEKBEN	1
			21513	KÉPEKET	2
			21514	KÉPEKKEL	1
			21515	KÉPÉNEK	3
			18507	ILLUSZTRÁCIÓI	1
			18508	ILLUSZTRÁCIÓJA	1
			18509	ILLUSZTRÁCIÓK	1
			18510	ILLUSZTRÁCIÓKÉ+	1
			18511	ILLUSZTRÁLÁSÁR+	2
			18512	ILLUSZTRÁLJA	3
			18513	ILLUSZTRÁLJAM	1
			18514	ILLUSZTRÁLNI	1
			18515	ILLUSZTRÁLT	1
			18516	ILLUSZTRÁLVA	3

Table 6: The distribution of the synonyms of *ábra* (number of occurrences; type/token ratio)

## 6 Conclusions

The corpus-driven intralingual analysis carried out above has shown that in the different components of a comparable corpus the lemma *ÁBRA* observes different patterns of behaviour on both the paradigmatic and the syntagmatic level. It was the qualitative analysis of the corpus data and of the concordance lines that brought to light the double life of the lemma. Since the patterns identified are not unusual or uncommon, they would have remained hidden for our attention which is caught mainly by peculiarities. Nevertheless the existence of these covert phenomena makes translated text differ from authentic text in the same language.

A look at the alphabetic lists of the subcorpora of the comparable corpus revealed that translated and authentic texts use different forms of the lemma *ÁBRA* with a different number of occurrences. Moreover the translational subcorpus contains forms which are completely absent from the authentic subcorpus. A closer analysis of the concordance of *ábra* exposed different patterns of usage in the translational and the authentic corpus component. The stability of these patterns was contrary to our expectations: the patterns in the authentic texts were looser than those in the translated texts.

The interlingual analysis of KWIC-concordances of the word form revealed a pattern of language use which has proven to be a source language pattern. The keyword is part of a multiword lexical unit which does not turn up in original texts. In general, it can be said that the patterns of language use in the translated subcorpus and the original subcorpus clearly differ.

The by-product of this small-scale investigation using a corpus-driven and bottom-up approach is the proof of two universals of translation: simplification and normalisation. The data threw up evidence that the same meaning was conveyed by fewer synonyms with more occurrences in the translated subcorpus, whereas the subcorpus of original texts used a wider range of synonyms with fewer occurrences. Similarly, the subcorpus of translated texts is characterised by the absence of foreign loan words, which are used naturally in the subcorpus of original texts.

## References

- Abondolo, Daniel (1992): "Hungarian." William Bright (ed.): *International Encyclopedia of Linguistics*. Vol. 2. New York/Oxford: Oxford University Press, 182-187
- Altenberg, Bengt; Mats Eeg-Olofsson (1990): "Phraseology in Spoken and Written English: Presentation of a Project." Jan Aarts, Willem Meijs (eds): *Theory and Practice in Corpus Linguistics*. Amsterdam/Atlanta: Rodopi, 1-26
- Baker, Mona (1993): "Corpus Linguistics and Translation Studies. Implications and Applications." Mona Baker, Gill Francis, Elena Tognini-Bonelli (eds): *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins, 233-250
- Bolinger, Dwight (1976): "Meaning and Memory." *Forum Linguisticum* 1 [2]: 1-14
- Chesterman, Andrew (1998): "Causes, Translations, Effects." *Target* 10 [2]: 201-230
- Francis, Gill (1993): "A Corpus-driven Approach to Grammar: Principles, Methods and Examples." Mona Baker, Gill Francis, Elena Tognini-Bonelli (eds): *Text and Technology. In Honour of John Sinclair*. Philadelphia-Amsterdam: John Benjamins, 137-156
- Frawley, William (1984): "Prolegomenon to a Theory of Translation." William Frawley (ed.): *Translation: Literary, Linguistic, and Philosophical Perspectives*. London/Toronto: Associated University Press, 159-175
- Gellerstam, Martin (1986): "Translationese in Swedish Novels Translated from English." Lars Wollin, Hans Lindquist (eds): *Translation Studies in Scandinavia. Proceedings from The Scandinavian Symposium on Translation Theory (SSOT)*. (Lund Studies in English 15.) Lund: CWK Gleerup, 88-95
- Hanks, Patrick (1996): "Contextual Dependency and Lexical Sets." *International Journal of Corpus Linguistics* 1 [1]: 75-98
- House, Juliane (1977): *A Model for Translation Quality Assessment*. Tübingen: Narr



- Hunston, Susan; Gill Francis, (2000): *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam/Philadelphia: Benjamins
- Kiss, Gábor (ed.) (1998): *Magyar szókincstár*. Budapest: Tinta
- Kjellmer, Göran (1994): *A Dictionary of English Collocations*. Oxford: Clarendon Press
- Kuiper, Koenraad; Douglas Haggio (1984): "Livestock Auctions, Oral Poetry, and Ordinary Language." *Language in Society* 13 [2]: 205-234
- Laviosa, Sara (1997): "How Comparable Can 'Comparable Corpora' Be?" *Target* 9 [2]: 289-319
- Mauranen, Anna (2000): "Strange Strings in Translated Language. A Study on Corpora." Maeve Olohan (ed): *Intercultural Faultlines*. Manchester/Northampton: St Jerome Publishing, 119-142
- Moon, Rosamund (1998): *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press
- Papp, Ferenc (1984): "The Russian of Hungarian People Whose Russian Is First-Rate." Ferenc Papp (ed.): *Contrastive Studies Hungarian-Russian*. Budapest: Akadémiai Kiadó, 139-154
- Pawley, Andrew; Frances Hodgetts Syder (1983): "Two Puzzles for Linguistic Theory: Nativelike Selection and Nativelike Fluency." Jack C. Richards, Richard W. Schmidt (eds): *Language and Communication*. London: Longman, 191-226
- Schmied, Josef; Hildegard Schäffler (1996): "Approaching Translationese Through Parallel and Translation Corpora." Carol Percy, Charles Meyer, Ian Lancashire (eds): *Synchronic Corpus Linguistics. Papers from the Sixteenth International Conference on English Language Research on Computerized Corpora (ICAME 16)*. Amsterdam/Atlanta: Rodopi, 41-56
- Scott, Mike (1998): *WordSmith Tools*. Oxford: Oxford University Press
- Selinker, Larry (1972): "Interlanguage." *IRAL* 10 [3]: 209-231
- Sinclair, John M. (1984): "Naturalness in Language." Jan Aarts, Willem Meijs (eds): *Corpus Linguistics: Recent Developments in the Use of Computer Corpora in English Language Research*. Amsterdam: Rodopi, 203-210
- Sinclair, John M. (1991): *Corpus Concordance Collocation*. Oxford: Oxford University Press

#### **trans-kom**

**trans-kom** ist eine wissenschaftliche Zeitschrift für Translation und Fachkommunikation.

**trans-kom** veröffentlicht Forschungsergebnisse und wissenschaftliche Diskussionsbeiträge zu Themen des Übersetzens und Dolmetschens, der Fachkommunikation, der Technikkommunikation, der Fachsprachen, der Terminologie und verwandter Gebiete.

Beiträge können in deutscher, englischer, französischer oder spanischer Sprache eingereicht werden. Sie müssen nach den Publikationsrichtlinien der Zeitschrift gestaltet sein. Diese Richtlinien können von der **trans-kom**-Website heruntergeladen werden. Alle Beiträge werden vor der Veröffentlichung anonym begutachtet.

**trans-kom** wird ausschließlich im Internet publiziert: <http://www.trans-kom.eu>

#### Redaktion

Leona Van Vaerenbergh  
Hogeschool Antwerpen  
Hoger Instituut voor Vertalers en Tolken  
Schilderstraat 41  
B-2000 Antwerpen  
Belgien  
[leona.vanvaerenbergh@scarlet.be](mailto:leona.vanvaerenbergh@scarlet.be)

Klaus Schubert  
Fachhochschule Flensburg  
Studiengang Internationale Fachkommunikation  
Kanzleistraße 91-93  
D-24943 Flensburg  
Deutschland  
[schubert@fh-flensburg.de](mailto:schubert@fh-flensburg.de)



- Sinclair, John M.; Antoinette Renouf (1988): "A Lexical Syllabus for Language Learning." Ronald Carter, Michael McCarthy (eds): *Vocabulary and Language Teaching*. Harlow: Longman, 140-158
- Tognini-Bonelli, Elena (2001): *Corpus Linguistics at Work*. Amsterdam/Philadelphia: Benjamins
- Toury, Gideon (1980): *In Search of a Theory of Translation*. Tel Aviv: The Porter Institute for Poetics and Semiotics
- Weinert, Regina (1995): "The Role of Formulaic Language in Second Language Acquisition: A Review." *Applied Linguistics* 16 [2]: 180-205
- Wilss, Wolfram (1996): *Übersetzungsunterricht. Eine Einführung. Begriffliche Grundlagen und methodische Orientierungen*. Tübingen: Narr

#### *Author*

Maria Balaskó is Reader in Applied Linguistics in the Institute of Intercultural Studies at the University of West Hungary. Main research interests: Translations studies, corpus linguistics (especially the interface between lexis and grammar), corpus-based translation studies, intercultural communication.

E-mail: [balaskomaria@btk.nyme.hu](mailto:balaskomaria@btk.nyme.hu)

Website: <http://alknyelv.bdf.hu/>