



DOI: 10.2478/topling-2015-0004

# Translating the representation of the tourist landscape: A corpus-based study

Stefania Gandin  
University of Sassari, Italy

## Abstract

This paper will present a corpus-based study on the translated language of tourism, focusing in particular on the stylistics of tourist landscapes. Through a comparative analysis of a specifically designed corpus of travel articles originally written in English (namely the *TourEC - Tourism English Corpus*) and a corpus of tourist texts translated from a variety of languages into English (namely the *T-TourEC - Translational Tourism English Corpus*), the study will investigate a selection of collocates, concordances and keywords related to the description and representation of tourist settings in both corpora. The aim will be that of identifying differences, aspects or practices to be potentially improved that characterize the translated language of tourism with respect to tourist texts originally written in English. Results will show that the discursive patterns of translated texts differ from the stylistic strategies typically employed in native English for the linguistic representation of landscape and settings due to phenomena of translation universals, and that these differences may affect the relating communicative functions, properties and persuasive effects of tourist promotional discourse.

## Keywords

corpus linguistics, language of tourism, stylistics, translation studies.

## 1. Introduction: corpus linguistics, translation and tourism discourse

The application of corpus linguistics to descriptive and applied translation studies represents nowadays a consolidated, but still evolving, research trend. Before the nineties and even afterwards, the vast majority of stylistics and corpus linguistics studies<sup>1</sup> excluded translations from the material to be analysed, since translated texts were considered merely as alternative versions of original texts written in other specific languages. It was only with the development of Zohar's polysystem theory (1979), Toury's Descriptive Translation Studies (1995) and Baker's foregrounding papers (1993, 1995, 1996) that the application of corpus analysis to translation studies finally found its solid theoretical framework.

The first corpus-based translation analyses<sup>2</sup> were mainly dedicated to the description of the potential research paths offered by the application of corpus linguistics methodologies to the study of translation as a distinct linguistic phenomenon, in order to define the specific linguistic features, patterns and communicative functions of the so-called 'language of translation' or 'translationese' – a term used to indicate (often in a pejorative sense) how 'the language of translated texts may differ from that of other texts produced in the same language' (Zanettin, 2012: 12). The subsequent and increasing number of

<sup>1</sup> Cfr. Jeffries and McIntyre (2010); Sinclair (2004); Simpson (2004); Tognini-Bonelli (2001).

<sup>2</sup> Cfr. Baker (1993, 1995, 1996, 1999 and 2000); Bowker (2002); Bowker and Pearson (2002); Johansson (2003); Kenny (2001); Laviosa (1998 and 2002); Olohan and Baker (2000); Tymoczko (1998); Zanettin (2000).

studies<sup>3</sup> employing corpus linguistics tools in translation led to larger-scale projects focusing on the discovery and description of 'universals of translation', that is the macro-linguistic features characterizing 'the distinctive nature of translation as a communicative event (...) shaped by its own goals, pressures and context of production' (Baker 1996: 175), thus recognizing translation as a variety of language with its own 'regular patterns of behaviour' (Zanettin 2012: 11), 'not in order to criticise or evaluate individual translations, but in order to understand what actually happens in the process of translation' (Baker 1996: 175). Translation universals include:

- the phenomenon of explicitation, represented by all those devices used to spell things out rather than leave them implicit in translation, by means of explanatory, grammatical and/or lexical items in the specification of terms conjunctions or any supplementary information added to the text;
- the concept of interference which refers to features of the SL that get transferred in target texts during the process of translation;
- the phenomenon of normalization represented by the tendency to exaggerate the features of the target language and to conform to its grammatical and collocational patterns, through the normalization of grammatical structures, punctuation collocational patterns and lexical creativity in terms of suffixes and ST unique words;
- the phenomenon of levelling out, represented by the tendency of translated texts to move away from any extreme of oral or literate markedness involved both in the source and target language;
- the phenomenon of simplification, expressed through the use of simplified language resulting in a lower degree of lexical density and a narrower range of type-token ratios, or by means of shorter sentences, alteration of the punctuation from weaker to stronger marks, omission of redundant or repeated information and shortening of complex collocations;

- the notion of unique items, which entails the investigation of unusual TL specific lexical items which are not common in the standard TL and may turn out to be even less frequent in translated texts
- and, finally, untypical collocations, which are those word combinations that although possible in the TL, are rare or absent in standard TL texts.

Beside the study of translation universals, the application of corpus linguistics methodologies has also addressed language combinations involving translations of a variety of textual genres to and from English, through research projects aimed at identifying practices and norms characterizing the translation of standard language, the stylistic features of individual translators<sup>4</sup> and the translation of languages for special purposes. Amongst the largest and most important translational corpus studies, we can recall parallel corpora dedicated mainly or exclusively to fictional texts, such as CEXI and COMPARA, or other kinds of corpora (parallel, monolingual and/or comparable) comprising a larger variety of textual typologies, including fictional and specialist texts, such as ENPC, MLCC, TEC and TRANSEARCH (Gandin, 2009). However, when considering corpus-based research projects dedicated to the translation of languages for special purposes, we can notice that there are not many corpora dedicated to study of the language of tourism and its translational features in a comparative and/or parallel perspective.

Hence, the aim of this paper will be that of presenting the results of a comparable corpus analysis aimed at investigating whether and to what extent universals of translation may affect the communicative functions, stylistic features and persuasive effects typically featured in a specific language for special purposes, namely the language of tourism, focusing in particular on the description and representation of tourist settings. The language of tourism represents a unique type of specialized language made up of a wider range of stylistic, pragmatic and lexical features intertwined with and influenced by different registers and different specialized languages; its linguistic and sociolinguistic

<sup>3</sup> Cfr. Mauranen and Kuyamaki (2007); Olohan (2004); Zanettin (2012).

<sup>4</sup> Cfr. Baker (2000); Bosseaux (2004); Hermans (1996), Johansson (2004), Kenny (2001).

features have been studied by numerous scholars through extensive and interdisciplinary approaches.<sup>5</sup> The main theoretical model followed in this research is the one developed by Graham Dann (1996), who classified the features of tourism

discourse into convergent properties (which are the ones that the language of tourism shares with other specialized languages), divergent properties (those that differentiate the language of tourism from other types of discourses) and verbal and visual techniques (that can also be used in combined forms), as shown in Table 1 below.

<sup>5</sup> Cfr.: Dann (1996); Dann & Parrinello (2009); Francesconi, (2007); Gandin (2013); Gotti (2007); Kang and Qiaofeng (2011), Nigro (2006); Palusci and De Stasio (2007).

CONVERGENT PROPERTIES		DIVERGENT PROPERTIES	
Functions (Jacobsonian classification)		Lack of sender identification	
Structure: combination of text/image or textual functions in order for the product to be purchased binary language of opposites		Monologue	
Tense: binary opposition btw present and future		Euphoria: use of hyperbolic language	
Magic: misrepresenting time and nature through language and images.		Tautology: stereotypes, pre-packaged expectations about the destination	
VERBAL TECHNIQUES		VISUAL TECHNIQUES	
comparison	humour	colour	visual cliché
key words & keying	linguaging	format	connotation procedures
testimony	ego-targeting		
VERBAL AND VISUAL TECHNIQUES COMBINED			
puzzles		ousting the competition	
temporal contrast		infraction of taboo	
collage		significant omission	

**Table 1. Dann’s model (adapted): properties and techniques of the language of tourism**

Such a multilevelled characterization of tourism discourse implies that its translational practices must take into account the complex lexical and pragmatic features of this language, particularly when dealing with tourist texts in which the representation of landscape and tourist places plays a key role in the construction and development of discourse and whose description in terms of either natural settings or urban spaces performs a two-fold referential and phatic function.

**2. Methodology**

The research carried out a comparative analysis of specific keywords and relating collocates and concordances taken from two monolingual corpora of English translated and non-translated texts dedicated to the language of tourism, focusing on the linguistic representation of landscape and tourist settings.

The first monolingual corpus employed in the analysis was the TourEC (Tourism English Corpus). It was compiled between

2011 and 2012 and comprises 468 254 tokens and 36 498 types (type/token ratio: 7.79), with over 500 travel articles downloaded from the web (BBC travel web site), written in English by a variety of authors and dealing with a vast array of typical tourist topics and locations worldwide as shown in Table 2.

TourEC TOPICS			
Adventure	Food & Drink	Nature & Outdoors	
Arts & Architecture	Health	Road Trips	
Beaches	Hiking	Romance	
Budget	History	Shopping	
Business	Holidays	Snow and skiing	
Cruises	Hotels	Sports	
Cultural Activities	Living in...	Tours & Classes	
Eco-tourism	Luxury	Travel Tips	
Family	Music	Weekends	
TourEC LOCATIONS			
Abu Dhabi	Costa Rica	Istanbul	Rio de Janeiro
Alaska	Croatia	Italy	San Francisco

Amsterdam	Dubai	Japan	Seattle
Argentina	Ecuador	Kenia	Seoul
Australia	Egypt	Las Vegas	Singapore
Barcelona	Finland	Los Angeles	South Africa
Beijing	France	Malaysia	Spain
Berlin	Germany	Mexico	Tahiti
Brazil	Great Britain	Miami	Thailand
Buenos Aires	Greece	Morocco	Tokyo
Cambodia	Hawaii	Moscow	Turkey
Canada	Hong Kong	New York City	Ukraine
Caribbean	India	New Zealand	USA
Chicago	Indonesia	Poland	Vietnam
China	Ireland	Prague	Washington DC

**Table 2. TourEC topics and locations**

The second corpus - the T-TourEC (Translational Tourism English Corpus) - was created in 2013 and comprises 361 198 tokens and 23 144 types (type/token ratio: 6.41), with over 800 texts divided into a set of three source-language related sub-corpora exclusively dedicated to travel texts translated into English from Italian, Norwegian and Japanese<sup>6</sup>. Texts were downloaded from the institutional web sites of the national Japanese, Italian and Norwegian tourist boards and describe the typical tourist attractions of the countries taken into consideration, with specific sub-topics corresponding to those included in the TourEC, so as to avoid potential topical bias in the comparison of the two corpora (see Table 3 below).

T-TourEC TOPICS		
action and adventure	lakes & mountains	sightseeing areas
active holidays	leisure and cultural activities	sports and adventure
art and history	nature and wildlife	tourist facilities
attractions and culture	recreation and sports	tours and safaris
culture and entertainment	religion and spirituality	UNESCO World Heritage Sites
eating and drinking	sea and beaches	well-being and health
family and fun	shopping	where to go

**Table 3. T-TourEC topics**

<sup>6</sup> These languages were chosen because of their distant and different linguistic origins, in order to avoid biased data deriving from the potential repetition of linguistic and translational patterns implied in source languages coming from the same or similar language family.

By means of Wordsmith (Concord and Wordlist tools) and PoS-tagging procedures (CLAWS part-of-speech tagger - tagset: C5) we identified a series of specific collocates and relating concordances referring to landscape and settings representations in both corpora, and the comparison of their data allowed the identification of a series of stylistic differences characterizing the translated language of tourism with respect to tourist texts originally written in English, determined by specific phenomena of translation universals.

By analysing the Wordlist data of both corpora, we isolated the first ten most frequent nouns (minimum frequency: 30) relating to landscape and tourist settings description, excluding toponyms or highly specialized terms referring to geographical, natural or urban attractions. We thus identified and selected what we deemed to represent the most semantically broad lemmas relating to the persuasive description of tourist landscapes and settings in both native and translated English tourism discourse, that is: *landscape(s)*, *space(s)* and *scenery(ies)*.

	TourEC	#		T-TourEC	#
1.	CITY(ies)	1126		CITY(ies)	749
2.	PLACE(s)	531		REGION(s)	504
3.	REGION(s)	143		PLACE(s)	359
4.	SPACE(s)	116		LANDSCAPE(s)	170
5.	LANDSCAPE(s)	92		LAND	131
6.	LOCATION	90		SCENERY(ies)	128
7.	DESTINATION(s)	78		SITE(s)	111
8.	SCENERY(ies)	60		DISTRICT(s)	89
9.	PARADISE	58		WILDERNESS	50
10.	WILDERNESS	39		SPACE(2)	37

**Table 4. Most frequent lemmas relating to tourist landscape**

We identified their most relevant lexical collocates (5L and 5R with a minimum frequency of 2) in terms of adjectives, common nouns and verbs (excluding therefore specific toponyms which could have led to topic-biased data, functional words such as articles, prepositions, conjunctions, or auxiliaries with no relevant lexical value) in order to better understand their stylistic peculiarities in both corpora. Each collocate was carefully examined through a further and detailed concordance check, in order to filter out meanings from other conceptual domains, verify once again and categorize its relating PoS valence and

link the theoretical categories introduced in Dann’s model to the analytical data of both corpora. Non-relevant conceptual meanings filtered out from the analysis included:

- *landscape*, as noun denoting pictures representing an area of countryside, or the relating painting genre;
- *space*, as noun denoting time, or the physical universe beyond the Earth’s atmosphere.

Finally, data were normalized in order to compare the results on the base of a desired corpus size of 500 000 tokens<sup>7</sup>.

### 3. Analysis

As previously outlined, the analysis focused on a selection of concordances and keywords taken from two monolingual corpora of English translated and non-translated texts, with the aim of verifying the presence of translation universals and their consequences on the communicative functions, stylistic features and persuasive effects characterizing the language of tourism and, more specifically, the linguistic representation of landscape and tourist settings. The choice of comparing a native English corpus with a translational monolingual corpus has been inspired by the research methodology of the TEC-project. TEC is an annotated corpus of contemporary translational English: it comprises a selection of written texts translated into English from European and non-European source languages, including a detailed set of extra-linguistic information relating to the gender and nationality of the translator, the time and place of production of both ST and TT, and so on. The aim of TEC is to provide linguistic data to study the way in which the patterning of translated text might be

<sup>7</sup> The normalization ratio for the Tour-EC resulted in 1.07, while for the T-TourEC was 1.38.

different from that of non-translated text, due to phenomena of translation universals.<sup>8</sup> Another field of investigation is the study of stylistic variation across individual translators, analysed by integrating the TEC linguistic data with all the extra-linguistic information included in the corpus.

The monolingual translational corpus employed in our analysis (T-TourEC) was designed precisely to follow – on a smaller scale – one of the main research goals of the TEC project: the discovery of the most frequent universals of translation in translated tourism discourse, particularly in the representation of tourist landscapes. Therefore, the study entailed a detailed examination of collocates and concordances characterizing the translational context of three specific lemmas [i.e. *landscape(s)*, *space(s)*, *scenery(ies)*] and a further comparative analysis with another monolingual reference corpora in English dedicated to tourism discourse (TourEC). The analysis could not include parallel investigations with the source languages from which the T-TourEC texts originated, nor the study of individual translational behaviour and processes since, when consulted, all the three national tourist boards involved in the production of the texts did not provide access to these extra-linguistic data due to privacy policies.

#### 3.1 LANDSCAPE(S)

The results of the collocational patterns for the lemma *landscape(s)* are reported in tables 5 and 6 below:<sup>9</sup>

<sup>8</sup> Cfr. Baker (1993, 1995, 1996, 1999 and 2000); Laviosa (1998); Olohan (2004); Olohan and Baker (2000); Tymoczko (1998); TEC web site.  
<sup>9</sup> Key to acronyms: ADJ = adjectives / NN = common nouns / VB = verbs / PoS (#) = Parts of speech (frequency) / NRM = normalized datum.

N.	TourEC- <i>landscape(s)</i> [TOTAL node word frequency: 92 (NRM → 98.44)]	PoS	PoS #	N.	TourEC- <i>landscape(s)</i> [TOTAL node word frequency: 92 (NRM → 98.44)]	PoS	PoS #
1)	BEAUTIFUL	ADJ	2	25)	HILLS	NN	3
2)	BIG	ADJ	2	26)	HISTORY	NN	2
3)	CHANGING	ADJ	3	27)	LAKES	NN	2
4)	CIRCULAR	ADJ	2	28)	MIX	NN	2
5)	COLOURED	ADJ	2	29)	MOUNTAIN	NN	2
6)	CULTURAL	ADJ	3	30)	MUSEUM	NN	3
7)	DIFFERENT	ADJ	2	31)	PEOPLE	NN	2
8)	DRAMATIC	ADJ	4	32)	PAINTINGS	NN	2

9)	EPIC	ADJ	3	33)	PARK	NN	2
10)	GREEN	ADJ	2	34)	RENAISSANCE	NN	2
11)	HARSH	ADJ	2	35)	RIVER	NN	2
12)	LUNAR	ADJ	2	36)	SPECTACLE	NN	2
13)	NATIONAL	ADJ	3	37)	SUNSET	NN	2
14)	NEOLITHIC	ADJ	2	38)	THOUSAND	NN	2
15)	PUBLIC	ADJ	2	39)	VINEYARDS	NN	2
16)	REMARKABLE	ADJ	2	40)	WAY	NN	3
17)	SURREAL	ADJ	2	41)	WEATHER	NN	2
18)	THEIR	ADJ	3	42)	WILDLIFE	NN	2
19)	UNIQUE	ADJ	2	43)	MADE	VB	2
20)	ART	NN	2	44)	SEE	VB	3
21)	BEAUTY	NN	2	45)	SEEMS	VB	2
22)	CITY	NN	2	46)	SET	VB	2
23)	CYPRESS	NN	2	47)	TAKE	VB	2
24)	GARDENS	NN	2				

**Table 5. TourEC - collocates of *Landscape(s)***

N.	T-TourEC - <i>landscape(s)</i> [TOTAL node word frequency: 177 (NRM → 244.26)]	PoS	PoS #	N.	T-TourEC - <i>landscape(s)</i> [TOTAL node word frequency: 177 (NRM → 244.26)]	PoS	PoS #
1)	ALPINE	ADJ	4	57)	GARDEN	NN	4
2)	ARCTIC	ADJ	2	58)	GLACIER	NN	2
3)	BEAUTIFUL	ADJ	9	59)	GORGE	NN	2
4)	BREATHTAKING	ADJ	2	60)	GUESTS	NN	2
5)	CHANGING	ADJ	2	61)	HILLS	NN	2
6)	CHARACTERISTIC	ADJ	2	62)	HISTORY	NN	3
7)	CHARMING	ADJ	3	63)	HORSE	NN	2
8)	COASTAL	ADJ	6	64)	HOTEL	NN	9
9)	CULTURAL	ADJ	15	65)	ISLANDS	NN	2
10)	DIFFERENT	ADJ	3	66)	LEAVES	NN	2
11)	DRAMATIC	ADJ	3	67)	LIGHT	NN	2
12)	FANTASTIC	ADJ	3	68)	MOUNTAIN	NN	3
13)	FASCINATING	ADJ	2	69)	MOUNTAINS	NN	4
14)	FRESH	ADJ	2	70)	NATURE	NN	3
15)	HISTORICAL	ADJ	2	71)	NORWEGIANS	NN	2
16)	IMPORTANT	ADJ	2	72)	OPENNESS	NN	2
17)	JAGGED	ADJ	2	73)	PARK	NN	2
18)	JAPANESE	ADJ	3	74)	PARKS	NN	2
19)	LUCANIAN	ADJ	2	75)	PART	NN	2
20)	LUNAR	ADJ	3	76)	PATHS	NN	2
21)	MAIN	ADJ	2	77)	ROAD	NN	2
22)	MOUNTAINOUS	ADJ	2	78)	ROUTE	NN	2
23)	NATIONAL	ADJ	2	79)	SCENERY	NN	2
24)	NATURAL	ADJ	5	80)	SEA	NN	2
25)	NEWEST	ADJ	2	81)	SNOW	NN	2
26)	NORTHERN	ADJ	2	82)	STYLES	NN	2
27)	NORWEGIAN	ADJ	6	83)	TERRITORY	NN	2
28)	OPEN	ADJ	2	84)	TRAILS	NN	2
29)	PICTURESQUE	ADJ	5	85)	TREE	NN	2
30)	RICH	ADJ	4	86)	TRIP	NN	2

31)	ROLLING	ADJ	2	87)	VALLEY	NN	2
32)	SMALL	ADJ	2	88)	VIEW	NN	2
33)	SPECTACULAR	ADJ	3	89)	WAY	NN	2
34)	SPLENDID	ADJ	2	90)	WILDLIFE	NN	2
35)	STEEP	ADJ	2	91)	WINNER	NN	3
36)	STRONG	ADJ	2	92)	WINTER	NN	4
37)	STUNNING	ADJ	2	93)	CHARMED	VB	2
38)	UNIQUE	ADJ	4	94)	DOMINATED	VB	2
39)	VARIED	ADJ	6	95)	ENJOY	VB	2
40)	VOLCANIC	ADJ	2	96)	EXPERIENCE	VB	2
41)	WONDERFUL	ADJ	2	97)	EXPLORE	VB	2
42)	ACTIVITIES	NN	2	98)	FEATURES	VB	2
43)	AREA	NN	3	99)	FEEL	VB	2
44)	AREAS	NN	2	100)	FIND	VB	3
45)	ART	NN	4	101)	HIKING	VB	2
46)	AUTUMN	NN	3	102)	INCLUDES	VB	2
47)	BUILDINGS	NN	2	103)	KEPT	VB	2
48)	CENTRE	NN	3	104)	KNOWN	VB	2
49)	CLIFFS	NN	2	105)	MADE	VB	2
50)	COAST	NN	3	106)	MAKE	VB	2
51)	COLOURS	NN	3	107)	OFFERS	VB	2
52)	CULTURE	NN	3	108)	OPENS	VB	3
53)	DOLOMITES	NN	2	109)	SHOWS	VB	2
54)	EXPERIENCE	NN	2	110)	SURROUNDED	VB	3
55)	FAUNA	NN	2	111)	SHAPED	VB	2
56)	FJORD	NN	11	112)	VARIES	VB	2

Table 6: T-TourEC - collocates of *Landscape(s)*

Just by considering the frequency of the lemma *landscape* in both corpora, we can notice a much higher rate in T-TourEC, almost doubling (and tripling when normalized) the preference for this lemma in the translational choices of tourist texts. Consequently, in terms of lexical variation there are far more variants in T-TourEC with respect to TourEC: 112 collocates (NRM 154.56) against 47 (NRM 50.29), almost tripling the amount of lexical words accompanying the collocational pattern of the term *landscape* in translational tourism discourse. These data appear even more remarkable if we analyse and compare the overall distribution of normalized PoS in each corpus (as shown in Figure 1) and might be interpreted in the light of different translation universals.

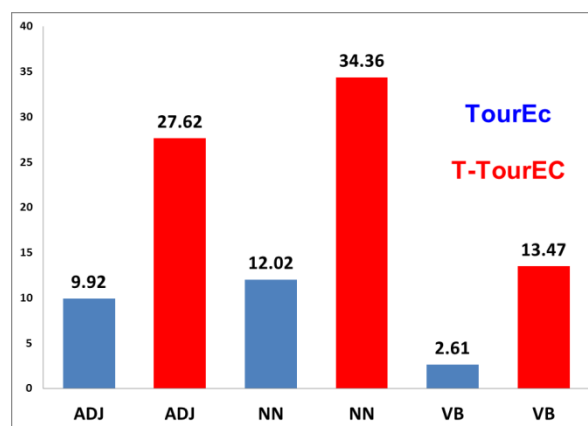


Figure 1. *Landscape(s)*: normalized PoS %

Most likely indeed, the larger range and normalized number of different collocates characterizing the lexical patterning of *landscape(s)* in T-TourEC may reflect an expression of explicitation phenomena, represented by all those devices used to ‘spell things out rather than leave them implicit in translation’ (Baker 1996:180), as the high proportion of normalized

adjectives, nouns and, mostly, the five-times higher percentage of verbs seems to suggest. Moreover, the relevant percentage of nouns in T-TourEC may indicate universals of simplification through nominalization processes that contribute to the shortening of complex collocations and the creation of less complex grammatical and syntactical structures (see following example).

(1) *A flight to remember. A fantastic landscape and experience. Fly over one of Norway's finest outdoor area by helicopter and let the magic of the Arctic scenery pass by beneath you.* (T-TourEC)

The wide variety and number of normalized adjectives in T-TourEC may be considered as the realization of another translation universal, that is the phenomenon of normalization, defined as 'the tendency to exaggerate the features of the target language and to conform to its grammatical and collocational patterns' (Ibid.: 183). In this sense, the high proportion of evaluative and non-evaluative adjectives in translated tourist texts may be considered as the tendency to exploit those collocates able to convey some of the most typical persuasive effects of tourism discourse even more frequently with respect to 'native' tourist texts, thus leading to an over-employment of:

- euphoria, that is a property the language of tourism employed 'to speak only in positive and glowing terms of the service and attractions it seeks to promote' (Dann, 1996: 65) by means of superlatives, hyperbole and other linguistic devices able to emphasize the uniqueness of the attraction/place to be promoted, as shown in the following example:

(2) *Sicily, too, is covered in natural reserves and brehtaking landscapes.* (T-TourEC)

- keywords and keying, defined by Dann (1996: 174-175) as 'a series of attributes of the destination [...] which correspond to the requirements of the potential tourist [...] through 'the use of appropriate language' able to give 'an aura of genuineness' and 'authenticity' to the destination promoted (i.e. through the use of words such as *away, adventure, dream, unique, exotic, authentic, genuine* etc., as reported in the example below);

(3) *Italy offers endless and amazing opportunities [...] for those who want to*

*be in close contact with nature both during the summer and the winter, surrounded by breathtaking scenery and unique landscapes.* (T-TourEC)

- the concurrent attainment of the conative function, in order to attract the potential tourist's attention by means of persuasive language, such as superlative adjectives, positive evaluative adjectives, combined with peculiar verbal forms such as imperatives, the use of past participle in thematic position etc. (see example 4);

(4) *Take part in a trip and experience spectacular landscape, nature and birdlife, a fantastic bluish light, and the possibilities to watch the orcas in the Vestfjord.* (T-TourEC)

- the achievement of the referential function, through the employment of non-evaluative adjectives and also common nouns (see example 5);

(5) *Sea kayaking trips provide a different view of the Northern Norwegian coastal landscape* (T-TourEC).

### 3.2 SPACE(S)

Regarding the lemma *space(s)*, the representation of this notion across the TourEC and T-TourEC collocates displays a rather opposite situation. First of all, the term *space(s)* is far less frequent in T-TourEC, probably because of phenomena of interference<sup>10</sup> linked to the abstractness of the term<sup>11</sup>, which could lead to ambiguities in the translational choices taken for the representation of tourist scenarios. As an almost inevitable consequence, the lexical range of collocates accompanying the lemma

<sup>10</sup> The presence of universals of interference can be evidently inferred also from the specific lexical collocates relating to unit of measurement in T-TourEC, since imperial units (which should be the preferable target-reader oriented translational option when translating a text into English) are missing and replaced by metric system referents typically used in the corpus source-languages measurement systems.

<sup>11</sup> The notion of *space* is also linked to time and duration in English - see the most relevant definitions provided by the Oxford English Dictionary [*Space (n)*]: I. Denoting time or duration. \* Time which is free or available for doing something; leisure; [...]/ II. Denoting area or extension. \* General or unlimited extent. [...] Physical extent or area; extent in two or three dimensions. The physical expanse which surrounds something; extent in all directions from a given point or object.] and the WordNet tool [**Noun** - **space**: the unlimited expanse in which everything is located / an empty area (usually bounded in some way between things) / an area reserved for some particular purpose].



*space(s)* is more limited in T-TourEC, with just 14 (NRM: 19.32) variants against the 51

(NRM: 54.57) different collocates retrieved in TourEC, as reported in tables 7 and 8.

N.	TourEC- <i>space(s)</i> [TOTAL node word frequency: 119 (NRM → 127.33)]	PoS	PoS #	N.	TourEC- <i>space(s)</i> [TOTAL node word frequency: 119 (NRM → 127.33)]	PoS	PoS #
1)	CHINESE	ADJ	2	27)	LOFT	NN	2
2)	CONTEMPORARY	ADJ	2	28)	MUSEUM	NN	2
3)	ENCHANTING	ADJ	2	29)	NIGHTS	NN	2
4)	GREAT	ADJ	2	30)	PLENTY	NN	2
5)	GREEN	ADJ	2	31)	OFFICE	NN	2
6)	INDUSTRIAL	ADJ	2	32)	RETAIL	NN	2
7)	LITTLE	ADJ	2	33)	PEOPLE	NN	2
8)	OLD	ADJ	2	34)	PLAYGROUND	NN	2
9)	OUTDOOR	ADJ	2	35)	PUBS	NN	2
10)	PUBLIC	ADJ	2	36)	ROOMS	NN	2
11)	SMALL	ADJ	2	37)	STAGE	NN	2
12)	SQ	ADJ	2	38)	STUDIO	NN	2
13)	SQUARE	ADJ	3	39)	STYLE	NN	4
14)	WHITE	ADJ	2	40)	SUN	NN	2
15)	ART	NN	5	41)	TREE	NN	2
16)	AUDIENCE	NN	2	42)	VEHICLE	NN	2
17)	BAR	NN	2	43)	CARRIED	VB	2
18)	CENTER	NN	4	44)	JOSTLE	VB	2
19)	CITY	NN	2	45)	LEAVE	VB	2
20)	COAST	NN	3	46)	SAVING	VB	2
21)	END	NN	2	47)	SHARE	VB	4
22)	EXHIBITION	NN	4	48)	SPEND	VB	2
23)	FOOD	NN	2	49)	TRANSFORM	VB	2
24)	GLASS	NN	2	50)	TURNED	VB	2
25)	HIGHLIGHT	NN	2	51)	USE	VB	2
26)	ICE	NN	2				

Table 7. TourEC – collocates of *Space(s)*

N.	T-TourEC- <i>space(s)</i> [TOTAL node word frequency: 37 (NRM → 51.06)]	PoS	PoS #	N.	T-TourEC- <i>space(s)</i> [TOTAL node word frequency: 37 (NRM → 51.06)]	PoS	PoS #
)	ALL	ADJ	2	)	EXHIBITION	NN	2
)	FIRST	ADJ	2	)	METER	NN	3
)	OPEN	ADJ	3	0)	MUSEUM	NN	4
)	SHOPPING	ADJ	2	1)	TIME	NN	3
)	SQUARE	ADJ	2	2)	VISITORS	NN	2
)	AREA	NN	3	3)	WORLD	NN	4
)	ART	NN	2	4)	CREATED	VB	2

Table 8. T-TourEC – collocates of *Space(s)*

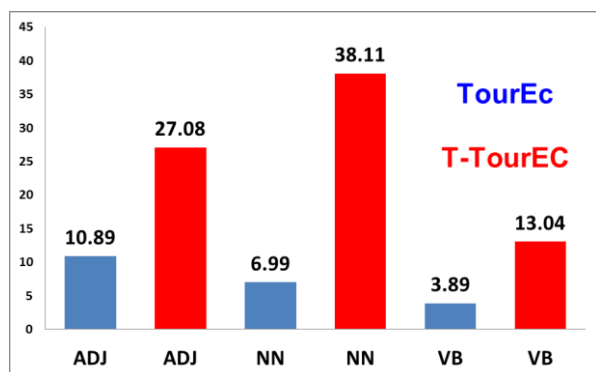


Figure 2. *Space(s)*: normalized PoS %

By comparing the specific collocates, TourEC presents a wider degree of lexical variation, both in nouns (predominant PoS) and adjectives. Therefore, notwithstanding the lower degree of specificity of the term in general English, the use of *space(s)* in native tourism discourse seems to be characterized by collocational patterns referring to the various physical features of the lemma which represent a source of tourist attraction (specific social settings, geographic origins, historical details etc.), and thus allow the application of the typical persuasive properties of tourism discourse, including euphoria, keywords and keying and the

ordinary presence of the referential function, as shown in the following examples.

(6) *The premier space for contemporary art is the Hamburger Bahnhof, while Picasso fans gravitate to the Museum Berggruen.* (TourEC)

e.g. 7) *What started in 1900 as the dream of a Barcelonian magnate, Eusebi Güell, for an English-style 'garden city' for the hoity-toity ended up as an enchanting public space.* (TourEC)

On the contrary, the collocational patterns of the lemma *spaces(s)* in translational tourism discourse seem to be limited to PoS with a mere referential function (see e.g. 8 and 9): T-TourEC collocates do not include any evaluative adjective (which on the contrary can be retraced in TourEC), and even verbal collocates are far less frequent compared to TourEC data (cfr. tables 7, 8 and Figure 2 above), thus presenting a strong limitation of the most persuasive stylistic properties of tourism discourse.

(8) *The most recent member of this group is Rygge Storsenter, with 25,000 square metres of shopping space.* (T-TourEC)

(9) *The building is 83 metres long and has been reconstructed as a living museum with spaces dedicated to the exhibition of findings, reconstructions, and domestic animals.* (T-TourEC)

### 3.3 SCENERY(IES)

Concerning *scenery(ies)*, T-TourEC reports higher frequency rates with respect to TourEC, more than doubling the preference for this notion in the stylistic representation of tourist attractions (in particular natural settings, as can be observed from the typology of collocates most frequently accompanying the term in both corpora - see tables 9 and 10, almost to counterbalance the scarce presence of the previously analysed lemma *space(s)*).

N.	TourEC- <i>scenery(ies)</i> [TOTAL node word frequency: 60 (NRM → 64.2)]	PoS	PoS #	N.	TourEC- <i>scenery(ies)</i> [TOTAL node word frequency: 60 (NRM → 64.2)]	PoS	PoS #
1)	ALL	ADJ	2	15)	ALPS	NN	2
2)	BEAUTIFUL	ADJ	2	16)	CLIFFS	NN	3
3)	BEST	ADJ	2	17)	FOREST	NN	2
4)	BREATH TAKING	ADJ	4	18)	LAKE	NN	2
5)	CHANGING	ADJ	3	19)	MOUNTAIN	NN	4
6)	DRAMATIC	ADJ	3	20)	REGION	NN	2
7)	EPIC	ADJ	2	21)	SUN	NN	2
8)	GORGEOUS	ADJ	2	22)	TREES	NN	2
9)	LASHED	ADJ	3	23)	WALKS	NN	2
10)	ROCKY	ADJ	2	24)	INCLUDING	VB	2
11)	ROOFED	ADJ	2	25)	INSPIRED	VB	2
12)	SPECTACULAR	ADJ	2	26)	KNOWN	VB	2
13)	STUNNING	ADJ	3	27)	MAKE	VB	2
14)	SUBLIME	ADJ	2	28)	PROVIDES	VB	3

Table 9. TourEC – collocates of *Scenery(ies)*

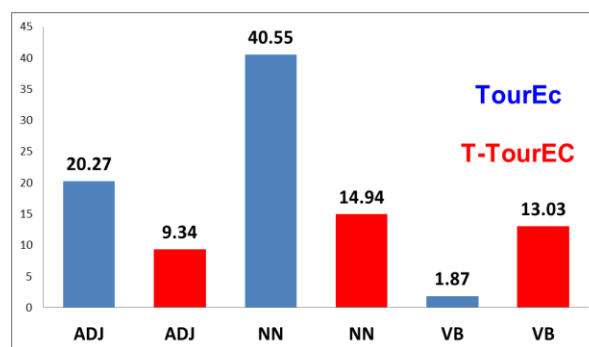
N.	T-TourEC- <i>scenery(ies)</i> [TOTAL node word frequency: 129 (NRM → 178.02)]	PoS	PoS #	N.	T-TourEC- <i>scenery(ies)</i> [TOTAL node word frequency: 129 (NRM → 178.02)]	PoS	PoS #
1)	AMAZING	ADJ	3	40)	FJORDS	NN	2
2)	ARCTIC	ADJ	5	41)	FORESTS	NN	2
3)	ASTOUNDING	ADJ	3	42)	HERITAGE	NN	2
4)	BEAUTIFUL	ADJ	23	43)	HISTORY	NN	2
5)	BREATH TAKING	ADJ	2	44)	ISLAND	NN	2
6)	COASTAL	ADJ	5	45)	ISLANDS	NN	2
7)	DRAMATIC	ADJ	3	46)	KITING	NN	2
8)	EASY	ADJ	3	47)	LANDSCAPES	NN	2
9)	EXTREME	ADJ	24	48)	MIDST	NN	2
10)	FABULOUS	ADJ	3	49)	MOUNTAIN	NN	7

11)	FANTASTIC	ADJ	2	50)	MOUNTAINS	NN	3
12)	IMPRESSIVE	ADJ	2	51)	NATURE	NN	4
13)	INCREDIBLE	ADJ	2	52)	PEAKS	NN	4
14)	MAGNIFICENT	ADJ	9	53)	PEOPLE	NN	2
15)	NATURAL	ADJ	5	54)	PICTURE	NN	2
16)	NORWEGIAN	ADJ	2	55)	ROUTE	NN	2
17)	NOSTALGIC	ADJ	3	56)	SEA	NN	2
18)	PERFECT	ADJ	5	57)	STREETS	NN	3
19)	RICH	ADJ	2	58)	TIME	NN	2
20)	SPECTACULAR	ADJ	2	59)	TOWNS	NN	2
21)	STRIKING	ADJ	4	60)	TRIP	NN	2
22)	STUNNING	ADJ	10	61)	VARIETY	NN	2
23)	SURROUNDING	ADJ	2	62)	WALK	NN	2
24)	UNSPOILT	ADJ	2	63)	WAY	NN	2
25)	WESTERN	ADJ	3	64)	WINTER	NN	12
26)	WILD	ADJ	3	65)	WORLD	NN	2
27)	WONDERFUL	ADJ	2	66)	BOASTING	VB	2
28)	AREA	NN	2	67)	BOOK	VB	2
29)	BACKDROP	NN	3	68)	BORED	VB	3
30)	BACKGROUND	NN	2	69)	CREATING	VB	2
31)	BRIDGE	NN	3	70)	DISCOVER	VB	2
32)	CIRCLE	NN	2	71)	ENJOY	VB	10
33)	CITY	NN	2	72)	EXPERIENCE	VB	4
34)	COAST	NN	2	73)	EXPLORE	VB	2
35)	CONTRAST	NN	2	74)	INCLUDING	VB	2
36)	CULTURE	NN	10	75)	OFFER	VB	3
37)	DAY	NN	2	76)	PROVIDE	VB	2
38)	FILMS	NN	2	77)	SURROUNDED	VB	4
39)	FJORD	NN	10	78)	WAITING	VB	3

**Table 10. T-TourEC – collocates of *Scenery(ies)***

In terms of lexical variations, T-TourEC presents far more variants with respect to TourEC, similarly to the previous results of the term *landscape(s)*. However, the specific distribution of PoS in each corpus differs (Figure 3), since TourEC has a larger proportion of adjectives (mostly evaluative ones) with respect to its other PoS, probably due to the semantic frame of the word *scenery* that is perceived as a very suggestive word<sup>12</sup> in Standard English. Therefore, in the promotional context of native tourism texts, *scenery(ies)* is accompanied by a larger number of adjectives (i.e. *dramatic, epic, gorgeous* etc.) relating to the properties of euphoria and keywords and keying, which boost more effectively the persuasive effects entailed in the lemma. On the contrary, adjectives are less frequent in the inner distribution of the T-TourEC PoS but compensated by a larger proportion of nouns, thus confirming a

predominant tendency towards nominalization processes in translated tourism discourse, linked to the universal of simplification.



**Figure 3. *Scenery(ies)*: normalized PoS %**

Finally if we consider the overall PoS ratios, the higher percentage of adjectives and verbs in T-TourEC with respect to TourEc seems to suggest phenomena of explicitation (by adding lexical devices in order to express things more clearly in translation) and normalization, which, as in the previous case of *landscape(s)*, manage to over-exploit the properties of:

<sup>12</sup> See also the definition provided by the Oxford English Dictionary: ‘Scenery: [MASS NOUN] 1The natural features of a landscape considered in terms of their appearance, especially when picturesque: *spectacular views of mountain scenery.*’

▪ euphoria:  
 (10) *The train journey provides some of Norway's wildest and most magnificent scenery.* (T-TourEC)

▪ keywords and keying:  
 (11) *From Rondane National Tourist Route, you have easy access to the Rondane Mountains with an excellent network of waymarked trails. Join activities in wild and unspoiled scenery – summer as well as winter.* (T-TourEC)

▪ the conative function:  
 (12) *Enjoy the beautiful scenery between Oslo and Bergen as part of a round trip.* (T-TourEC)

▪ the referential function, by means of non-evaluative adjectives and – mostly – common nouns:

e.g. 13) *Katsura-hama Beach is famous for the pleasant contrast between the beach's pine tree forests and rocky coastal scenery.* (T-TourEC)

e.g. 14) *Shin-Noboritbetsu-onsen hot springs, and has a rich variety of scenery including forests, lakes and marshes.* (T-TourEC).

## Conclusions

The aim of this paper was that of presenting a corpus-based comparative analysis of two corpora of native and translated tourist texts in English, concerning the representation of landscapes and tourist settings, in order to investigate the specific discursive patterns of translated texts and identify the role and effects of their relating translation universals. By applying on a smaller scale the research methods of the TEC-project, the analysis was carried out through a quantitative and qualitative comparison between translated and non-translated texts, focusing on collocates and concordances of three specific lemmas best representing the description of tourist settings: *landscape(s)*, *space (s)* and *scenery(ies)*. What emerged from this research is that phenomena of universals of translation can influence and even alter to a certain extent the communicative functions, stylistic properties and persuasive effects typically employed in Standard-English tourist texts. Results demonstrated that *landscape(s)* was the most recurring term in translated tourism discourse, representing a sort of 'passepartout' or 'universal' word for the identification of urban and natural tourist places. Its translational and stylistic features showed significant differences with respect to the characteristics of native-English tourism discourse, and included:

▪ a larger number of different collocates used in the lexical patterning of the term (triple amount of normalized collocates with respect to TourEC), linked to the universal of explicitation;

▪ a relevant percentage of nouns (34.36%), indicating patterns of nominalization and related to the universal of simplification;

▪ a wider variety and proportion of evaluative and non-evaluative adjectives, linked to the properties of euphoria, keywords and keying, and the conative and referential functions, which can be considered as the realization of normalization phenomena.

The frequency of the lemma *space(s)* was much lower in the translated tourist texts because of phenomena of interference from the original source languages, also relating to the abstractness of the term. Consequently, the lexical range of collocates accompanying the lemma *space(s)* was more limited in translated tourism discourse, resulting in the absence of evaluative adjectives and in a low number of verbal collocates. The main nominalization processes characterizing the collocational patterning of *space* in translated tourism discourse seemed to indicate the expression of universals of interference and simplification, leading to collocational patterns with a mere referential function and strongly limiting the persuasive effects of the other typical properties of the language of tourism. Finally, the collocational patterning of the word *scenery(ies)* confirmed the preference of this term for the description of natural tourist attractions in both corpora. Nonetheless, the analysis outlined several differences between TourEC and T-TourEC. *Scenery(ies)* was more frequent in translated tourism discourse in comparison with native tourist texts, resulting as the second most recurrent term in the description of tourist settings, thus balancing somehow the lower incidence of *space(s)*. The overall higher ratio of adjectives and verbs in T-TourEC with respect to TourEC seemed to suggest phenomena of explicitation and normalization, thus conforming once again to the properties of euphoria, keywords and keying, and the conative and referential functions. However, the inner distribution of PoS in the two corpora was different with respect to the data concerning *landscape(s)*

and *space(s)*, whose hierarchy always reflected the order NN, ADJ and VB for both lemmas and in both corpora. In the case of *scenery(ies)* TourEC reported a larger proportion of adjectives (mostly evaluative ones) with respect to its other PoS, probably due to the semantic frame of the word *scenery*, while adjectives were less frequent in the T-TourEC but compensated by a larger proportion of nouns, thus confirming and further enhancing the predominant tendency towards nominalization processes in translated tourism discourse, linked to the universal of simplification.

In conclusion, universals of simplification, normalization, explicitation and interference seem to represent the main reasons for the collocational divergences between translated and non-translated tourism discourse in the representation of tourist settings able to affect collocational variability and the standard employment of the persuasive properties of the language of tourism. Even

if the analysis was limited to three generic lemmas, and could not be integrated with extra-linguistic data entailing the commercial and narrative features involved in the process of translation (in-house or external translations, gender, nationality and occupation of the translators, direction of translation or other procedural aspects), it would be worth further developing the study in the future by analysing other lemmas (including for instance more specific categories of natural attractions, urban spaces and even catering and accommodation spaces), or other PoS that can be relevant in the description of tourist settings, in order to identify the potential role of other universals of translation and thus fully understand and improve the strategies and features of translated tourism discourse.

## References

- BAKER, M., 1993. Corpus Linguistics and Translation Studies — Implications and Applications. In M. Baker, G. Francis and E. Tognini-Bonelli, eds. *Text and Technology: In honour of John Sinclair*. Amsterdam: John Benjamins, pp. 233 -250.
- BAKER, M., 1995. Corpora in Translation Studies An Overview and Some Suggestions for Future Research. *Target* 7:2, pp. 223-243..
- BAKER, M., 1996. Corpus-based Translation Studies. The Challenges that Lie Ahead. In H. Somers, ed. *Terminology, LSP and Translation*. Amsterdam: John Benjamins, pp. 175-86.
- BAKER, M., 1999. The Role of Corpora in Investigating the Linguistic Behaviour of Professional Translators. *International Journal of Corpus Linguistics* 4(2), pp. 281-298.
- BAKER, M., 2000. Towards a Methodology for Investigating the Style of a Literary Translator. *Target* 12(2), pp: 241-266.
- BOSSEAU C., 2004. Translating point of view: A corpus-based study. *Language Matters* 35 (1), pp. 259-274.
- BOWKER, L., 2002. *Computer-Aided Translation Technology-A Practical Introduction*. Ottawa: Ottawa University Press
- BOWKER, L. and PEARSON J., 2002. *Working with Specialized Languages: A practical guide to using corpora*. London and New York: Routledge.
- DANN, G., 1996. *The Language of Tourism. A Sociolinguistic Perspective*. Wallingford: CAB International.
- DANN, G. and PARRINELLO LIEBMAN G. eds., 2009. *The Sociology of Tourism - European Origins and Developments*. Bingley: Emerald Group Publishing.
- FRANCESCONI, S., 2007. *English for Tourism Promotion*. Milano: Hoepli.
- GANDIN, S., 2009. Linguistica dei corpora e traduzione: definizioni, criteri di compilazione e implicazioni di ricerca dei corpora paralleli. *Annali della Facoltà di Lingue e Letterature Straniere* , 5, pp. 133 - 152.
- GANDIN S., 2013. Translating the language of tourism. A corpus-based study on the Translational Tourism English Corpus (T-TourEC). *Procedia - Social and Behavioral Sciences* 95, pp: 325-335.
- GOTTI, M., 2006. The Language of Tourism as Specialized Discourse. In O. Palusci and S. Francesconi, eds. *Translating Tourism: Linguistic / Cultural Representations*. Università di Trento: Dipartimento di Studi Letterari, Linguistici e Filologici, 15-34.
- HERMANS, T., 1996. The Translator's Voice in Translated Narrative. *Target* 8(1), 23-48.
- JEFFRIES, L. and McINTYRE D., 2010, *Stylistics*, Cambridge: Cambridge University Press.

- JOHANSSON, S., 2003. Reflection on Corpora and their Uses in Cross-linguistic Research. In F. Zanettin, S. Bernardini and D. Stewart, eds. *Corpora in Translator Education*. Manchester: St. Jerome, pp. 135-144.
- JOHANSSON, S., 2004. Why change the subject? On changes in subject selection in translation from English to Norwegian. *Target* 16[1], pp. 29-52.
- KANG, N. and QIAOFENG Y., 2011. Corpus-based Stylistic Analysis of Tourism English. *Journal of Language Teaching and Research*, 2[1], pp. 129-136.
- KENNY, D., 2001. *Lexis and Creativity in Translation: A Corpus-based Study*, Manchester: St. Jerome.
- LAVIOSA, S., 1998. The English Comparable Corpus: A Resource and a Methodology. In L. Bowker, M. Cronin, D. Kenny and J. Pearson, eds. *Unity in Diversity? Current Trends in Translation Studies*. Manchester: St. Jerome, pp. 101-112.
- LAVIOSA, S., 2002. *Corpus-based Translation Studies. Theories, Findings, Applications*. Amsterdam/New York: Rodopi.
- MAURANEN, A. and KUYAMAKI P. eds., 2004. *Translation Universals: Do they exist?*. Amsterdam: John Benjamins.
- NIGRO, M.G., 2006. *Il Linguaggio Specialistico del Turismo. Aspetti storici, teorici e traduttivi*. Roma: Aracne.
- OLOHAN, M., 2004. *Introducing Corpora in Translation Studies*. London and New York: Routledge.
- OLOHAN M. and BAKER M., 2000. Reporting that in Translated English: Evidence for Subconscious Processes of Explication?. *Across Languages and Cultures*, 1, pp. 141-172.
- OXFORD ENGLISH DICTIONARY online: <http://www.oed.com/> [Accessed September 2014].
- PALUSCI, O. and DE STASIO C. eds., 2007. *The Languages of Tourism: turismo e mediazione*. Milano: Unicopli.
- PRINCETON UNIVERSITY. "About WordNet." *WordNet*. Princeton University. 2010. <http://wordnet.princeton.edu> [Accessed January 2015].
- SINCLAIR, J., 2004. *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- SIMPSON, P., 2004. *Stylistics – A Resource Book for Students*. London: Routledge.
- TEC web site – University of Manchester – Centre for Translation and Intercultural Studies: <http://www.llc.manchester.ac.uk/Research/Centres/CentreforTranslationandInterculturalStudies/ResearchProgrammesPhDMPhil/TranslationEnglishCorpus/> [Accessed January 2015].
- TOGNINI-BONELLI, E., 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- TYMOCZKO, M., 1998. Computerized corpora and the future of translation studies. *Meta*, 43, pp. 652-659.
- ZANETTIN, F., 2000. Parallel Corpora in Translation Studies: Issues in Corpus Design and Analysis. In Olohan M. ed. *Intercultural Faultlines. Research Models in Translation Studies I: Textual and Cognitive Aspects*. Manchester: St. Jerome, pp. 105-118.
- ZANETTIN, F., 2012. *Translation-Driven Corpora Corpus Resources for Descriptive and Applied Translation Studies*. Manchester: Saint Jerome Publishing.

#### **Author's address and contact details**

Stefania Gandin  
(MA), Lecturer in English Language and Translation  
Dipartimento di Scienze Umanistiche e Sociali  
Università degli Studi di Sassari  
Via Roma 151, Sassari  
07100, Italy  
Phone: +39079229691  
E-mail: [sgandin@uniss.it](mailto:sgandin@uniss.it)