

## Investigating the Complementary Polysemy and the Arabic Translations of the Noun *Destruction* in EAPCOUNT

Hammouda Salhi

Volume 58, Number 1, April 2013

URI: <https://id.erudit.org/iderudit/1023818ar>

DOI: <https://doi.org/10.7202/1023818ar>

[See table of contents](#)

Publisher(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (print)

1492-1421 (digital)

[Explore this journal](#)

Cite this article

Salhi, H. (2013). Investigating the Complementary Polysemy and the Arabic Translations of the Noun *Destruction* in EAPCOUNT. *Meta*, 58(1), 227–246.  
<https://doi.org/10.7202/1023818ar>

Article abstract

This article investigates a topic at the intersection of translation studies, lexical semantics and corpus linguistics. Its general aim is to show how translation studies can benefit from both lexical semantics and corpus linguistics. The specific objective is to capture the semantic and pragmatic behavior of the noun *destruction* and its different translations into Arabic. The data are obtained from an English-Arabic parallel corpus made from UN texts and their translations (EAPCOUNT). The analysis of the data shows the polysemy of the word *destruction* as a number of semantic and pragmatic alternations can be captured. These findings are discussed in the frame of the Generative Lexicon (GL) theory developed by James Pustejovsky. The paper concludes with some concrete suggestions on how to enhance the relationship between linguists and translators and their mutual cooperation.

# Investigating the Complementary Polysemy and the Arabic Translations of the Noun *Destruction* in EAPCOUNT

HAMMOUDA SALHI

University of Tunis El Manar, Tunisia

hammouda\_s@hotmail.com

## RÉSUMÉ

Le présent article aborde un sujet à la croisée de la traductologie, de la sémantique lexicale et de la linguistique de corpus. Son objectif général est de montrer comment la traductologie peut tirer profit de la sémantique lexicale et de la linguistique de corpus. Plus spécifiquement, il vise à capturer le comportement sémantique et pragmatique du substantif *destruction* et de ses différentes traductions en arabe. Les données sont extraites d'un corpus parallèle anglais-arabe, constitué de textes de l'ONU et de leurs traductions (EAPCOUNT). L'analyse des données recueillies met en relief la polysémie du mot *destruction* dans la mesure où un certain nombre d'alternances sémantiques et pragmatiques peuvent être saisies. Ces résultats sont discutés dans le cadre de la théorie du Lexique Génératif (LG) de James Pustejovsky. L'article se conclut sur certaines propositions concrètes pour améliorer la relation et la coopération mutuelle entre linguistes et traducteurs.

## ABSTRACT

This article investigates a topic at the intersection of translation studies, lexical semantics and corpus linguistics. Its general aim is to show how translation studies can benefit from both lexical semantics and corpus linguistics. The specific objective is to capture the semantic and pragmatic behavior of the noun *destruction* and its different translations into Arabic. The data are obtained from an English-Arabic parallel corpus made from UN texts and their translations (EAPCOUNT). The analysis of the data shows the polysemy of the word *destruction* as a number of semantic and pragmatic alternations can be captured. These findings are discussed in the frame of the Generative Lexicon (GL) theory developed by James Pustejovsky. The paper concludes with some concrete suggestions on how to enhance the relationship between linguists and translators and their mutual cooperation.

## MOTS-CLÉS/KEYWORDS

traductologie fondée sur corpus, linguistique de corpus, sémantique lexicale, ambiguïté lexicale, polysémie  
corpus-based translation studies, corpus linguistics, lexical semantics, lexical ambiguity, polysemy

It is my belief that the time is now ripe for a major redefinition of the scope and aims of translation studies, and that we are about to witness a turning point in the history of the discipline. I would like to argue that this turning point will come as a direct consequence of access to large corpora of both original and translated texts, and of development of specific methods and tools for interrogating such corpora in ways which are appropriate to the needs of translation scholars.

(Baker 1993: 235)

Over the history of both linguistics and translation studies, the relationship between linguists and translators had been plagued by reciprocal disparagement in spite of the fact that there exists many areas of common interest. Malmkjær (1998: 535) is right when she asserts that “linguists and translators ought to be the best of friends; their areas of interest, however one wants to look at them, and however they may differ, have language and linguistic activity at the centre.” Mutual cooperation can undoubtedly bring many advantages to both linguistic research and translation studies. It can even be argued that the area at the intersection of the two disciplines is wide enough to make such cooperation a necessity.

One of the major achievements of linguistic research has come from the development of corpus linguistics, and corpus semantics in particular (Stubbs 2001), as a method allowing large-scale, systematic and empirical studies to be performed. These studies are now possible thanks to the copious amounts of data that large corpora can make available. Related computational tools, such as concordancers and alignment tools are also an offshoot of this “new paradigm” (Laviosa 1998), which marked a new massive collaboration mood generated by globalization. Now it is largely accepted that computers have paved the way to more reliable methods in the study of language and, owing to computational tools, linguistic research has reached an interesting turning point (Pustejovsky 1995:5). Corpus linguists, for instance, extol the use of computers in linguistic research and equate them with telescopes and microscopes (Kenny 2001), where in the past data were “limited to what a single individual could experience and remember” (Sinclair 1991:1). The ever increasing number of lexical databases and the increasing availability of both original and translated texts in digital format have made such data all the more accessible and have enabled the construction of all kinds of corpora, including *parallel* corpora — see, for example, Kenny (2001; 2009) and Laviosa (1997) who present a clear typology of corpora.

A parallel corpus consists of original texts alongside their translation into another language. The advantages of using parallel (or translation) corpora in both translation studies and linguistics have become apparent and the corpus-based approach has become one of the most fashionable trends in both disciplines (Zanettin 2002; Salhi 2010). Since the beginning of the 1990s, corpus-based work has given rise to a “paradigm shift” in the ways through which language can be investigated and translation process can be conceptualized (Laviosa 1998) thanks to the work carried out by such pioneers in corpus linguistics as Sinclair (1991; 1998; 1999), and Church and Hanks (1990), scholars who introduced the corpus approach to translation studies like Baker (1993; 1996; 1998; 2003), and other researchers who followed her lead such as Kenny (2001), Laviosa (2002), Zanettin (2000; 2009) and Zanettin, Bernardini et al. (2003), to name but a few. Baker (1993) explains that corpus work in translation is motivated by the need to explore large amounts of texts after the decline of the semantic view in favor of the pragmatic view. In fact parallel corpora are often used to reveal translation solutions in the target language.

Parallel corpora, however, include, alongside translated texts, original ones. Yet, it is surprising that research in this new trend has focused almost exclusively on the investigation of translated text. There is a scarcity of studies that use parallel corpora to provide fresh insights into original texts (Salhi 2005). Not only do parallel corpora enable translation scholars to unveil the nature of translated texts (Laviosa 1998), but

they may provide informations about the original language, which may be used to complement the work of linguists. But though mindful of the fact that any attempt to apply the parallel corpus approach to investigate the properties of lexical items in original text rather than in translated text is bound to raise questions about the feasibility of the exercise itself, we argue here that it might show how translation can inform linguistic research. We therefore investigate the noun *destruction* as a case study in order to demonstrate practically how the thesis is tenable.

We note that the description of nouns like *destruction* may vary from one dictionary to another and that it may sometimes appear as if it were monosemic. The lack of contexts of usage does not help the translators. We will see how knowledge of the semantic and pragmatic behavior of *destruction* is better informed when the item is looked at in real contexts and through its different translations into another language like Arabic. The data are taken from an *English-Arabic Parallel Corpus of United Nations Texts* (henceforth EAPCOUNT). This corpus, which will be described in detail below, contains texts and translations that can be regarded as reliable—an important consideration in an era when the reliability of texts is not always self-evident, especially when it comes to online texts and translations.

Through EAPCOUNT data, we hope to address the main question of the present article, namely how parallel corpora can come to the rescue of linguists and translation scholars alike in their investigation of the semantics, and *pragmatics*, of source language (SL) words. We also hope to achieve the broad aim of this article, which is to attempt to “endear” linguists to the world of translators and translation studies and vice versa. In order to try to achieve these goals the article is divided into six sections tackling the main aspects of the question. Section one shows how the gap between translation theory and translation practice is being bridged with the use of parallel corpora. Section two is devoted to the approach to lexical ambiguity and complementary polysemy taken from a translational perspective. Section three outlines how the item *destruction* has been handled in bilingual and monolingual dictionaries. Section four describes the methodology adopted when using EAPCOUNT data to account for the item *destruction*. Section five attempts to reveal the semantic and pragmatic behavior of *destruction* in EAPCOUNT. In section six a brief illustration is provided of how the findings of this paper are in line with the Generative Lexicon (GL) theory developed by the semanticist and generativist James Pustejovsky.

Before moving to the first section, a brief word is said here about the presentation of data in the current paper: Arabic-English and English-Arabic data are presented in the examples below as they appear in the fragments excerpted from EAPCOUNT. The data generally support discussion of contextually determined aspects of interpretation. However, morphemic glossing (left to right) is supplied where appropriate to facilitate analysis of selected features of the Arabic text. IPA phonological transcription for Arabic is provided when required by the context and is accompanied by a literal back translation of Arabic examples. Diacritics indicating emphatics and long vowels are therefore sometimes displayed.

### 1. Parallel corpora, translation theory and translation practice

The introduction of the corpus approach into translation studies can be seen as a response to, and a recovery from, the shortcomings of the previous linguistically-

oriented approaches to translation (see Snell-Hornby 1988). The principal contribution of this approach lies in its concern with the systematic study of authentic examples of language in use as opposed to older approaches, which were concerned with language as a 'static' mental construct. Carter (1998: 80) argues that if one wants to analyse lexical items in discourse, one should "move beyond constructed examples to a consideration of real texts." Carter is obviously in favor of an approach to lexical items that deals with contextually defined units of meaning, a method which John Sinclair (1991) considers central to the understanding language. But Baker (1993: 237) argues that studies that take the context into account and "attempt to investigate usage, are, by definition, only feasible if access is available to real data, and, in the case of usage, to substantial amounts of it." She goes even further to argue that corpus work has emerged as a response to that need. On the other hand, it should also be pointed out that the emergence of the corpus approach has also been motivated by the purposes of applied linguistics.

Translation is now regarded as an act of communication rather than a mere linguistic exercise. Therefore we strongly believe that *translation theory* is coming to be used to mean the body of knowledge that we have about the process of translating, including the decisions and strategies that translators can adopt to solve the many translation problems they encounter. Indeed developments which are taking place in the world of theoretical studies owe much of their strengths to practice. In many instances corpus-based studies have helped bridge the gap between theory and practice as well as the academic and professional gap (see Salhi 2010), which have plagued translation for a long time as outlined earlier. Parallel corpora, for instance, have already been used as repertoires of strategies deployed by on the past by translators, as well as repertoires of translation translations (Zanettin 2002). The high demand for, and the large-scale use of, parallel corpora can be explained by the need to address questions related to equivalence at lexical level (see Baker 1992) on the basis of pragmatic meaning or word usages, rather than on the basis of prototypical meanings of words. The satisfaction of such demand would be practically impossible without computers and copious amounts of data such as those which may be elicited from parallel corpora. When updated, such corpora can help researchers keep pace with current real usages of words better than any regularly updated dictionary. Parallel corpora can provide information that neither bilingual nor monolingual dictionaries usually contain. The use of corpora can give trainee translators a higher degree of confidence, for they are enlightened with the examples offered by real-life translations and the wide range of strategies adopted in a specific genre, register or style (Salhi 2010). Finding appropriate translations for words having more than one meaning, without necessarily being contrastively ambiguous, can be presented as a vexing problem for trainee translators as, most of the time, they translate core meanings rather than the actual pragmatic meaning in the given context.

## 2. Investigating lexical ambiguity from a translational perspective

Investigating lexical ambiguity in a bilingual context is a complicated but informative and useful exercise. Such complexity is due to the fact that the semantic space is divided up by languages quite differently and that complete and stable equivalences are rare.

### 2.1. Lexical ambiguity, homonymy and polysemy

It is almost certain that the majority of words in language carry more than one meaning, a phenomenon traditionally known as *lexical ambiguity*. Lexical ambiguity is an inherent problem of language because humans are impelled to assign to a finite resource of meaningful items an unlimited set of applications (Pustejovsky 1995; Sinclair 1998). According to Pustejovsky (1995: 27), Weinreich (1964) distinguishes two types of lexical ambiguity. The first type is that of contrastive ambiguity where, synchronically speaking, there is no relation between the different meanings of a word, i.e. cases of homonymy: in fact, this means that homonyms are actually independent lexical units. Well-known examples provided by Pustejovsky (1995: 27) are the cases of *bank* as in *Mary walked along the bank of the river*, vs. *HarborBank is the richest bank in the city*, or *taxi, line, bar*.<sup>1</sup> In dictionaries, a word exhibiting two distinct meanings as such would have two distinct entries *bank*<sub>1</sub> and *bank*<sub>2</sub> or *line*<sub>1</sub> and *line*<sub>2</sub> and so on. Pustejovsky (1991, 1995) calls a lexicon of this kind a *sense enumerative lexicon*, which he criticizes for undermining the richness of the knowledge that lexical items contain.

The second type of lexical ambiguity is what Pustejovsky calls *complementary polysemy* (CP), which applies where a word carries two (or more) meanings that relate to a common basic (or core) meaning. As for *bank*, compare *the bank raised its interest rates yesterday*, *the store is next to the newly constructed bank* and *the bank appeared first in Italy in the renaissance* (Pustejovsky 1995: 28). It is interesting to note that contrastive polysemy is essentially a pragmatically-constrained ambiguity that requires knowledge of the situational context to be resolved. For instance, in a sentence like *we finally reached the bank*, only the context of situation can tell us which bank is pointed to here, *the financial institution* or *the edge of a river* (see Stubbs 2001: 14 for more details on the ambiguity of the item *bank*). The complexity arising from a case of CP, on the other hand, is that it often cannot be disambiguated unless both semantic knowledge and pragmatic information are combined and made to complement each other. Pustejovsky calls these two concepts lexical knowledge and commonsense knowledge respectively. Thus the polysemous meanings in this case are in effect the different usages of lexical items in context.

### 2.2. Prototypical meaning vs. word usage and translation

According to Pustejovsky (1995), the problem of lexical ambiguity and polysemy is very central to language. This means that if you take any item in the way *bank*, *window* and *university* are handled above, you will find out that it exhibits at least two different senses, one of which is a core or prototypical meaning and the other is contextual. The various dimensions of lexical meaning are too complex to allow an easy answer (Stubbs 2001), but for the sake of simplicity, we will consider a word sense as either prototypical or contextual, though the contextual meaning overlaps sometimes to a great extent with the prototypical one.

A prototypical meaning is the meaning being immediately recalled in the memory when a word is uttered in isolation from any context, although some variations may be observed among speakers (Rosch 1973). Prototypical meanings of words constitute most of the time only a preliminary stage in the constitution of the final

contextual meaning which can never be expressed or perceived out of context. In some cases the prototypical meaning is hardly found through the final usage of a given word. There is no doubt that people rarely communicate only with core meanings of words. Communication depends heavily on contextual, pragmatic, meanings that words may convey through their lexical environment (Firth 1957) and the situational context. Therefore, meaning, as such, does not consist of something stable and pre-existing that can be elicited from texts in one language and rendered in texts of another language in the same way as one may copy a text from one file to another, as if translation were a copy and paste exercise and the translator a processing tool (Baker 1993: 236). Equivalence in word meaning is closely linked with correspondence in word usage, which goes beyond its prototypical meaning(s). This is a prerequisite for understanding texts and utterances and, therefore, for translating. Inferring pragmatic meanings of items is not, however, always easy and straightforward, especially when it comes to written texts whose authors are rarely accessible when more information about their intentions and implicatures is sought.

Inferences are especially problematic for novice translators who, when frustrated by the challenging and demanding translation task, fall back on monolingual and bilingual dictionaries which provide nothing more than brief definitions or prototypical equivalents (Salhi 2012). Therefore they tend to attach more importance to core meanings than the latter really merit. The inherent shortcomings of bilingual dictionaries come from the fact that they suggest translations on the basis of more or less prototypical meanings of SL items and potential usages of target language [TL] items, while the strength of parallel corpora is that they offer translations on the basis of actual usages of words following an approach which handles language as a dynamic mental construct (Carter 1998). While bilingual dictionaries tend to concentrate on the most obvious and prototypical and translations, parallel corpora offer researchers and translators a wider and richer set of real translations, reflecting context-bound decisions and strategies of professional translators. They can thus be used in the processes of teaching and learning translation, in that they encourage students to think of a wider range of possible translations that go way beyond what strictly theoretical linguistics or lexicography could offer them. Full reliance on sources of lexical data such as monolingual and bilingual dictionaries may therefore impact their translational competence and may result in non-fluent TL output. Another advantage of parallel corpora, and translations in general, is that they can be used to explore SL items through TL translations, an exercise which is less probable using dictionaries.

### *2.3. Using parallel corpora to rediscover the source text*

Although parallel corpora are now widely used to investigate translated texts through techniques based on comparisons with the original texts, as already mentioned, they are less often used to study the *original*. It is true that any translation is bound to its original, but once a translation is final, it becomes a fully-fledged language unit and a text in its own right (Baker 1993) as long as the rules of the TL and principles of coherence and cohesion are respected, though there is much controversy about the degree of equivalence which may be reached. It is not argued here that a translation should reach the status of the original—a translation remains always a translation



with its own specific features—but instead the aim is to draw attention to the service a translation may give to a researcher in investigating an original text or discourse.

### 3. *Destruction* in monolingual and bilingual dictionaries

Before moving on the use of EAPCOUNT as to explore the behavior of the word *destruction*, we now this item has been handled in bilingual and monolingual dictionaries.

#### 3.1. *Destruction* in English-Arabic dictionaries

A wide variety of English-Arabic dictionaries, either printed or electronic (online), suggest some translation solutions without any further information about the possible context of usage, which can be quite confusing sometimes. This is the case for the noun *destruction*. *LingvoSoft online Dictionary* and *Sakhr online Dictionary* (Table 1) are some representative samples of what electronic English-Arabic bilingual dictionaries generally tend to offer. They present nothing more than a list of de-contextualized equivalents, without any information on their possible usages. They also reflect a general failure to utilize frequency information as a criteria for determining the order of the different meanings of polysemous words. Having اجتياح (literally invasion) set as the fourth equivalent of *destruction* in the *Sakhr Dictionary*, while دمار (literally, [*devastation/state of destruction*]) is listed as the 17<sup>th</sup> equivalent is a good example of the somehow anarchical state of this kind of dictionary. In fact, the equivalents are just listed randomly, without any serious attempt to categorize them in accordance with some specific criterion. What is even worse is that some of the proposed equivalents are items which are less likely to be any translation of *destruction*, such as إبادة (literally, [*genocide*]) محب الهدم (literally, [*demolition lover*]).

TABLE 1  
*Destruction* in some online English-Arabic dictionaries

1	<i>LingvoSoft dictionary</i> <sup>2</sup>	إبادة، هدم، إغناء، دمار، فتك، مستحق الإباده، تدمير، إهلاك، خراب، هلاك، سبب الهلاك [annihilation, demolition, extermination, destruction (result), lethality, worthy of extermination, destruction (process), to destroy, ruin, perdition, the reason for the loss]
2	<i>Sakhr online Dictionary</i> <sup>3</sup>	إتلاف، إفساد، إهلاك، إجتياح، إنهدام، بوار، تباب، تبار، تحطم، تحطيم، تخريب، تدمير، تهديم، تهشيم، خراب، دك، دمار، سحق، فتك، محق، ملاشاة، موبقة، نقض، هد، هدم، هلاك، هلكة، هورة [damage, spoil, depreciation, invasion, collapse, doom, perdition, crash, smash, sabotage, destruction, process of destruction, demolition, crushing, ruins, bomb, destruction (result), crush, lethality, right, neglect, mortal, sin, veto, pluck, demolition, decimation, fatalism]
3	<i>Ectaco UK</i> <sup>4</sup>	إبادة، هدم، إغناء، دمار، فتك، مستحق، أباده، تدمير، إهلك، خراب، هك، سبب الهلاك [annihilation, demolition, extermination, destruction (result), lethality, exterminable, destruction (process), to destroy, ruins, bomb, the reason for the loss]



Another limitation of English-Arabic bilingual dictionaries is the lack of illustrative examples. Most of them are likely to suffer such weakness, though larger and more thorough dictionaries may be less prone to these charges. But even printed bilingual dictionaries lack the necessary contextual information (CI). Table 2 shows the entry for *destruction* in the *Al-Mawrid English-Arabic Dictionary* (2005), one of the most comprehensive and reliable dictionaries for this language pair.

TABLE 2

**Destruction in the Al-Mawrid English-Arabic Dictionary<sup>5</sup>**

1.	هدم تدمير تخريب إهلاك إبادة إلخ	demolition, tearing down destroying devastation, ravage ruining, annihilation genocide etc.	ACT or PROCESS of destruction
2.	دمار خراب هلاك إلخ	devastation, ruin ravage, wreck wreckage, doom, perdition etc.	STATE of being destroyed
3.	قوة مدمرة مهلكة أو سبب الهلاك إلخ	Alcohol is likely to be his (cause of) <i>destruction</i>	CAUSE of destruction

What is interesting about this account of *destruction* is that equivalents have been classified into three sense-based categories. The first set, which refer to the act of destroying, is kept separate from the second set where *destruction* refers to a state. In the third set, the equivalents denote the cause of destruction, more precisely a destroying agency. The authors of the dictionary provided an example probably because they felt that this sense might be confusing. Although better than the randomly set equivalents in the English-Arabic e-dictionaries cited above, such a treatment is still very reductive, and translations are generally not exemplified. If they were to rely fully on such dictionaries, trainee translators would face clear challenges in translating less typical usages of *destruction* like in Examples (1a) and (1b).

- (1) a. the destruction of wheat crops  
محاصيل الحبوب  
[cereal crops??]  
b. the destruction of rights and freedoms  
الحقوق والحريات  
[freedom and rights??]

None of the equivalents found in the bilingual dictionaries would faithfully express the meaning of *destruction* when collocating with wheat crops or human rights. Thus, while bilingual dictionaries are useful for delivering quick references for trainee translators, this same asset may be their main weakness, as they can deliver non-fluent and hasty translations instead.

### 3.2. *Destruction in English Monolingual Dictionaries*

While the majority of bilingual dictionaries are criticized for providing no more than proposed equivalents, monolingual dictionaries may also be criticized for undermining the real richness of the semantic and pragmatic knowledge that can be elicited from polysemous lexical items. Table 3 shows a sample of definitions for *destruction* reproduced from often-cited English dictionaries, the *Oxford Advanced Learner's Dictionary* (online), the *Cambridge Advanced Learner's Dictionary* (online), and the *Oxford English Dictionary* (1989; printed).

TABLE 3

#### *Destruction in some English monolingual dictionaries*<sup>6</sup>

1	<i>Oxford Advanced Learner's Dictionary</i> <sup>7</sup>	[UNCOUNTABLE] the act of destroying something; the process of being destroyed <ul style="list-style-type: none"> <li>• the destruction of the rainforests</li> <li>• weapons of mass destruction</li> <li>• a tidal wave bringing <b>death and destruction</b> in its wake</li> <li>• The central argument is that capitalism <b>sows the seeds of its own destruction</b>(= creates the forces that destroy it).</li> </ul>
2	<i>Cambridge Advanced Learner's Dictionary &amp; Thesaurus</i> <sup>8</sup>	noun <b>the act of destroying something, or the fact of being destroyed:</b> <i>Many people are very concerned about the destruction of the rainforests.</i> <i>Unusually high winds left a trail of destruction over the area.</i> <i>weapons of mass destruction (= those which kill or hurt large numbers of people)</i>
3	<i>Oxford English Dictionary</i> <sup>9</sup>	1) a) <i>The action of demolishing a building or structure of any kind, of pulling to pieces, reducing to fragments, undoing, wasting, rendering useless, putting an end to, or doing away with anything material or immaterial; demolition.</i> b) <i>The action of ravaging or laying waste; havoc, ruin.</i> c) <i>The action of putting to death, slaughter, now chiefly said of multitudes of men or animals, and of noxious creatures.</i> 2) <i>The fact, condition, or state of being destroyed, ruin.</i> 3) <i>A cause or means of destruction</i>

The first two monolingual dictionaries in Table 3 provide some illustrative examples where *destruction* is used in potential contexts. However, they tend to limit their range of focus to the most prototypical senses, and in a similar way. This might reflect the fact that the latest editions of these dictionaries are taking advantage of the availability of huge corpus resources, so as to best reflect current English. These dictionaries, however, still cannot be exhaustive by any means. Most of them fall short of offering examples which can go far beyond the scope of prototypical meanings.

Other monolingual dictionaries, such as the third one in Table 3 (*Oxford English Dictionary* 1989), are more comprehensive in their description of *destruction* but still are doomed to go amiss when attempting to reflect the real richness of the lexical item as well as when trying to keep pace with its semantic development and dynamism. In addition, because the item belongs to a general register with some technical and legal uses in certain contexts, these dictionaries, due to size constraints, provide nothing more than a general definition fitting a wide range of contexts.

In the above definitions, there is too much emphasis on prototypical senses of the item. Buildings are among the objects most prototypically associated with the act of destruction and they are presented either explicitly or implicitly throughout the whole description. This same concern is raised by Pustejovsky who deplores current monolingual dictionaries for their reductionist approach to meaning. Current dictionaries, according to him, reflect a particularly *static* approach in their handling of the concept of context by enforcing certain readings of a word instead of others. He argues that “the numbers of and distinctions between senses within an entry are ‘frozen’ into a fixed grammar’s lexicon. Furthermore, definitions hardly make *any* provisions for the notion that boundaries between word senses may shift with context” (Pustejovsky 1993a: 73). This state of affairs strongly suggests that a corpus-based approach is highly appropriate in a translation training environment, especially when complexities are generated by this systematic ambiguity.

#### 4. Methodology

This section describes the main features of EAPCOUNT corpus, gives a brief overview of the methodology adopted for its development and describes the procedures used to analyse and interpret its data.

##### 4.1. Description of EAPCOUNT

EAPCOUNT, as used in the research presented here,<sup>10</sup> comprised 261 English-Arabic bitexts aligned on a paragraph basis. The English subcorpus contained 2,749,876 word tokens, with 75,606 word types. The Arabic translation subcorpus had a slightly fewer number of word tokens (2,642,615), yet differed greatly in terms of the word types, which is 122,154. The whole corpus contained 5,392,491 tokens.

EAPCOUNT consists mainly, but not exclusively, of resolutions and annual reports issued by different UN organizations and institutions. Some texts are taken from the authoritative publications of another UN-like institution, namely the Inter-Parliamentary Union (IPU), representing 2.18% of the total number of tokens in the English subcorpus. But the great majority of texts are issued by the General Assembly and Security Council (66.44% SL tokens). The assumption here is that TL texts produced by these selected international bodies can be considered as translations of a high degree of reliability. All texts have been downloaded from first-hand sources (official websites of these agencies) in order to make sure that the publications are all kept in their original form. Table 4 shows the sources of the texts as well as the number of texts and tokens in the English subcorpus.<sup>11</sup>

TABLE 4

**Number of texts and word tokens of the English Subcorpus per organization**

Source		Number of texts	Percentage	Number of tokens	Percentage
United Nations Agencies	General Assembly and Security Council	185	70.88%	1,660,875	66.44%
	UNICEF	8	3.06%	47,419	1.72%
	UNIDO	4	1.5%	150,570	5.47%
	UN Office for Outer Space Affairs	3	1.14%	38,381	1.39%
	UNESCO	1	0.38%	19,564	0.71%
	Economic and Social Council	11	4.21%	22,077	0.80%
	WHO	1	0.38%	83,349	3.03%
	IMF	6	2.29%	483,638	17.58%
	Others	12	4.59%	183,797	6.68%
Non-UN Agencies	IPU	30	11.49%	60,206	2.18%
Total		261	100%	2,749,876	100%

EAPCOUNT covers a time-frame of about 14 years (Table 5) and may be considered to be synchronic, even though Meyer (2002:46) maintains that “a time-frame of 5 to 10 years seems reasonable” for a corpus to be synchronic. We argue that almost all original texts as well as the translations are issued by the same bodies and are governed by strict norms and standards of writing and translation, which may likely mean that language change happens at a slower pace. In addition, 22.6% of the texts were produced in 2009, 16% in 2007, and 13.4% in 2005, while 93.87% of the texts were produced over a period of 9 years, namely from 2001 to 2009, i.e. within the reasonable time-frame set by Meyer.

TABLE 5

**The time-frame of EAPCOUNT texts**

Years	96	97	98	99	00	01	02	03	04	05	06	07	08	09
Number of texts	1	1	2	5	7	13	5	14	24	35	31	42	22	59

**4.2. Compiling EAPCOUNT**

The compilation and alignment of EAPCOUNT involved several steps. First, the texts were cleaned up and extra material like pictures, tables, footnotes, and other untranslated parts were removed. Segmentation was then carried out manually at the level of paragraphs, with respect to the original texts, and Arabic translations were divided accordingly. We made sure that all the paragraphs could stand by themselves as meaningful complete units. Because we did not have at our disposal any alignment software that would work with Arabic scripts, the texts were manually aligned in a single file containing the English paragraphs and their Arabic counterparts in alternance. After completing this process, EAPCOUNT was ready for use.

**4.3. Software and method used to analyze EAPCOUNT data**

We used a concordance software, AntConc3.2.1w,<sup>12</sup> which provides a general purpose tool for conducting a wide range of investigations of copious amount of linguistic

data (Anthony 2006). It was used in the process of retrieval of the occurrences of *destruction* and its Arabic translations.

The method consisted in four main steps. We first retrieved the occurrences of the word *destruction* in the English subcorpus (133 occurrences). We then eliminated all the concordance lines that were similar: for example, the phrase *weapons of mass destruction* is always translated as أسلحة الدمار الشامل.

Then we isolated the Arabic translations for *destruction*. We then searched all the Arabic contexts containing an occurrence of these translations, as to find out if they were used to translate some other English lexical units.

The third step consisted in investigating the collocations of *destruction*, as to provide a significant insight into the behavior of this word. In other words, we categorized the occurrences the word *destruction* in their respective contexts to examine the word's English collocational behavior based on the different Arabic translations.

The final step consisted in the capture of the regularities of *destruction* use and perception. In the following section, we move on to look at how some bilingual and monolingual dictionaries deal with the entry *destruction* to be able to draw some conclusions based on a comparisons between such dictionaries and EAPCOUNT.

**5. Investigating the semantic and pragmatic behavior of destruction in EAPCOUNT**

**5.1. Occurrences of destruction in EAPCOUNT**

The concordancer AntConc 3.2.1w found 133 instances of the noun *destruction* in EAPCOUNT (See Table 6 for the ten first contexts). Four different translations are found for *destruction* in the corpus, namely *damār* (دمار), *tadmīr* (تدمير), *taqwid* (تقويض) and *itlāf* (إتلاف).

TABLE 6  
The ten first instances of *destruction* in EAPCOUNT

1 e the commission of the grave breach of extensive destruction of property, not justified by military necessity 1.txt
2 acks on the foundations of civilian life in Gaza: destruction of industrial infrastructure, food production, wa1.txt
3 sion investigated several incidents involving the destruction of industrial infrastructure, food production, wa1.txt
4 sued on previous days. The Mission finds that its destruction had no military justification. The nature of the 1.txt
5 the Fourth Geneva Convention. Unlawful and wanton destruction which is not justified by military necessity amou1.txt
6 s to a war crime. The Mission also finds that the destruction of the mill was carried out to deny sustenance to1.txt
7 oncludes that this was a deliberate act of wanton destruction not justified by any military necessity and draws1.txt
8 the same legal conclusions as in the case of the destruction of the flour mill. ٥١ - 1.txt
9 za Strip, the Mission witnessed the extent of the destruction of residential housing caused by air strikes, mor1.txt
10 gathered by the Mission strongly suggest that the destruction of housing was carried out in the absence of any 1.txt

## 5.2. Occurrences of target words

### 5.2.1. Case of تدمير [tadmīr]

This target word has 75 occurrences. The overwhelming majority of these occurrences (70 instances) are the translations of *destruction* while *devastation* (2), *demolition* (2; Example [2]) and *damage* (1) account for the rest.

- (2) وقد جرى بنجاح تدمير 5,529 جهازا متفجرا في موقع تجمع الجيش الماوي في شيتوان.  
[wa qad zarā bi-naḡāhin **tadmīru** 5,529 ḡihāzan mutafaḡiran fi mawqiʔi tazamuʔ  
al-ḡajf al-māwi fi šitwān]  
The *demolition* of 5,529 explosive items was successfully carried out at the  
Maoist cantonment site in Chitwan.

### 5.2.2. Case of دمار [damār]

This target word has 45 occurrences, 37 of which are translations of *destruction*. Seventeen (17) of these occurrences are the translations of *destruction* in the expression *weapons of mass destruction*. The other eight occurrences are the translations of *devastation* (5; Example [3]), *damage* (2) and *havoc* (1).

- (3) وفي الإنتعاش من الدمار الذي سببه عقدان من الحرب.  
[wa fi al-intiṣāfi mina **ad-damār** alaḡi sababahou ṣaḡd-āni mina al-harb]  
[and in the recovery from the *destruction* that caused-it decade-dual from the war]  
recovering from the *devastation* caused by over two decades of war

### 5.2.3. Case of تقويض [taqwidʕ]

The third word occurs 31 times, with only 1 occurrence being a translation of *destruction*. The other occurrences comprise 25 translations of *to undermine* (Example [5]), 2 translations of *erosion*, 1 translation of *to compromise* and 1 translation of *to unravel* (Example [6]).

- (4) قد نجحت في تقويض التحالفات المثيرة للقلق.  
[qad naḡahat fī **taqwidʕi** at-tahalouf-āti almoḡīrati lel-qalaqi]  
[emphasis- succeeded in *undermining* the allying-pl the arising to worry]  
It had successfully *undermined* the recent worrying alliances.
- (5) سيعرض عملية وقف إطلاق النار الهشة لخطر التقويض.  
[sajuṣaridʕu ṣamalijata waqfi itʕlāq an-nār alhaṣā liḡatʕari **at-taqwidʕi**]  
[will expose process cease shooting the fire the fragile to danger *the undermining*]  
The fragile ceasefire would start *to unravel*

### 5.2.4. Case of إتلاف [itlāf]

This fourth word occurs 5 times. Only 1 occurrence is a translation of *destruction*, 2 occurrences are the translations of *damage* (Example [6]), 1 is a translation of *annihilation* and 1 is a translation of *to vandalize*.

- (6) إتلاف مركبة تابعة لمكتب الامم المتحدة  
[**Itlāf** markaba tābiṣa li-maktabi al-omam al-motaḡid-a]  
[ravaging vehicle belonging to office the nations the united-fem/pl]  
A BINUB vehicle was *damaged*

These results provide an insight concerning the most prototypical usage of *destruction* in this, more or less, legal context. The polysemous sense<sup>13</sup> expressed by the Arabic lexical item *tadmīr* (تدمير) is more likely to correspond to the most prototypical usage of *destruction* in this context. This means that most usages of *destruction* have a semantic link with the sense which expresses the *act* of destroying. We must recall here that the repetitive instances of the expression *weapons of mass destruction* were weeded out from the analysis. For unlike other instances of collocations, which are all the result of much freer combinations – though they are still habitual cooccurrences, this expression and its Arabic equivalent أسلحة الدمار الشامل are rather *frozen expressions*, i.e. *compound nouns*. With such fixed expressions no variation is possible at all. The choice not to include them follows from the assumption that compounds should not be taken into account if we are to give an accurate proportional representation of the distribution of the different translations of *destruction* as found in the corpus.

The next step was to analyse the collocations where *destruction* appears and try to classify them according to the different translations proposed, in an attempt to understand why *destruction* is translated differently each time and with the aim of uncovering the underlying ambiguity of this lexical item in the SL.

### 5.3. Collocations of destruction

Firth, in a famous statement, argues that “you shall know a word by the company it keeps” (Firth 1957: 11). But despite the fact that there has been a great deal of theoretical and applied work related to collocations, there seems to be no agreed-upon definition for this concept (e.g. Allerton 1984; Cruse 1986; Sinclair 1991). Depending on their interests and points of view, researchers have focused on different aspects of collocations and given different definitions. One of the often-quoted definitions has been proposed by Sinclair: a collocation is “the occurrence of two or more words within a short space of each other in a text” (Sinclair 1991: 170). One obvious conclusion that can be drawn from this definition is the fact that the meaning of a word is, to a greater or lesser extent, inter-dependent with other words’ meanings. Studying the collocations of an item like *destruction* corresponding to each Arabic translation may help us to grasp the extent of the polysemy of *destruction*, and to explore the specific meaning of the Arabic equivalents.

The translation of *destruction* by تدمير [tadmīr] occurs when it denotes an act by an animate agent who destroys something for either evil or good purposes (Example [7])

- (7) تدمير بعض الهياكل الأساسية  
[tadmīr baʕḍ al-hajakil al-asāsi-ja]  
[Destroying some the – structures the-base-fem/pl]  
Destruction of basic infrastructure

The word دمار [damār] is used to refer to the result or the final state of a destructive event or act either by an animate agent or by nature (Example [8]).

- (8) الدمار نتيجة المجاعة  
[ad-damār natizata al-mazāʕa]  
[the destruction due to the-famine]  
the destruction caused by famine



In the case of تقويض [taqwīḍ], what is meant is the destruction of non-concrete objects (abstract and moral objects) such as *rights*, *freedom* (Example [9]) or *identity*.

- (9) تقويض الحقوق و الحريات  
[taqwīḍ al- huquq w al- horij-āt]  
[undermining the-right and the-freedom-pl]  
the *destruction* of rights and freedoms

Example [10] shows that when destruction is translated by إتلاف [itlāf], it is a case of partial and not complete demolition or devastation. This term is derived from the root أَتْلَف [atlaḥ], which means to deteriorate, to corrode, or to puncture, etc.

- (10) إتلاف محاصيل الحبوب  
[itlāf mahāsīl al- huboub]  
[Ravaging crops the wheat]  
The *destruction* of wheat crops

These findings from EAPCOUNT clearly support Zanettin's (2002: 11) claim that "while dictionaries favour a synthetic approach to lexical meaning (via definitions) corpora offer an analytic approach (via multiple contexts)." While analyzing the behavior of *destruction* through (1) its collocates in English source texts, (2) its various translations in the Arabic target texts, and (3) the other English lexical items than *destruction* translated by these latter, we noticed that a number of systematic alternations of the meaning of *destruction* could be captured. Across EAPCOUNT, *destruction* is used in a systematic way that makes it possible to analyse the extent of its meaning according to a number of common *behavioural traits*.

#### 5.4. Contextual alternations found in EAPCOUNT

##### 5.4.1. Negative or Positive Connotations

Through its lexicographical descriptions, *destruction* is generally perceived to refer to a negative act, with pejorative or at best neutral accounts of its meaning (see upper sections). In EAPCOUNT, however, it may also present some positive connotation, for example, in contexts in which it corresponds to the destruction of harmful and/or destructive objects. Example (11) comes from the report of the Secretary-General on the situation concerning Western Sahara in which he welcomes the efforts of the parties to clear the Territory of mines and unexploded ordnance, and goes on to express his pleasure with the notable humanitarian mine action progress achieved in this region.

- (11) I commend the Royal Moroccan Army for its efforts in the *destruction* of hazardous explosive ordnance and highlight its request for more resources, such as mechanical and detection equipment, in order to increase the pace and efficiency of its clearance work.

Still we should admit that rating some acts as negative and even destructive and others as positive may be intuitive, subjective or even ideological, exactly like the variable rating of fruits or birds as more or less prototypical. However, we feel that the CI provided for each example helps the reader to grasp the peculiarities of each usage.

#### 5.4.2. Extensive or Partial Act of Destruction Alternations

Contexts found in EAPCOUNT reveal a number of situations where the act of destruction is either complete or partial. A context-dependent case of complete destruction becomes all the more obvious when considering the English original of the translated expression 'إتلاف بالكامل' [total devastation] as *annihilation* and not *destruction* (taken from a letter dated 29 December 2005 from the Permanent Representative of Italy to the United Nations addressed to the Secretary-General).

- (12) With an adequate fund, too, the United Nations would be able to act quickly before an emergency – such as growing numbers of locusts in Africa – became a full-blown catastrophe, such as the annihilation of crops.

إتلاف المحاصيل الزراعية بكاملها

[Itlāf al-mahas'il azirašija bi-kāmiliha]

[ravaging the-crops agricultural completely]

In contrast, in Example (12) (taken from the 18<sup>th</sup> Progress report of the Secretary-General on the United Nations Mission in Liberia) إتلاف [devastation], is chosen as the translation of *destruction*. Therefore *destruction* when translated as 'إتلاف' was meant to be partial rather than complete.

- (13) Nuclear weapons of mass **destruction** can hit earth and destroy everything at a blink.

In Example (13), extensive destruction is clear without further contextual information.

#### 5.4.3. Concrete / Abstract Alternation

*Destruction* is often associated with concrete objects, as is reflected in almost all monolingual and bilingual dictionaries, except the *Oxford English Dictionary* (1989) which states that *destruction* may be used when referring to immaterial objects (see Table 5). *Material property, bridges, schools, farms, weapons, factories*, and so on are all common prototypical objects of destruction. Destruction of abstract things, however, may not come to the mind immediately. Yet, findings from the empirical work using EAPCOUNT show that destruction of abstract objects such as *identity, rights, freedoms, spirit, will, life* and so on, is quite common and repetitive enough to catch our attention (see Example [10]).

#### 5.4.4. Deliberate (even Systematic)/ Natural (Non-Deliberate) Alternation

*Destruction* is often taken as a deliberate action, but alternations found in EAPCOUNT indicate that it can be natural or accidental as well. Example (14) is taken from the report of the UN Secretary-General on the situation concerning Western Sahara, in which the Royal Moroccan Army has deliberately destroyed anti-personnel mines and other unexploded ordnance under the control of the United Nations Mission for the Referendum in Western Sahara (MINURSO).

- (14) MINURSO monitored the *destruction* by the Army of 70 anti-personnel mines, 62 anti-tank mines and 1,644 pieces of large-calibre unexploded ordnance, as well as of 1,500 mines and close to 10,000 items of other explosive ordnance from stockpiles.

Destruction can also be accidental, as in the case of natural disasters. Example (15) is taken from IPU Document No 17 (113th IPU Assembly).

- (15) Also distressed at the loss of life and *destruction* caused by famine and other natural disasters in parts of Africa.

## 6. Discussion

### 6.1. Methodology

The identification of regularities in the semantic properties of *destruction* is partially based on EAPCOUNT data, which is contextual in nature, supplemented by introspective interpretation. Another limitation is related to the alternations found in EAPCOUNT. Some of the classifications involve a sort of value judgements such as positive/negative or deliberate/natural. What may be seen as positive by some parties can be seen as negative by others, while what is deliberate or natural is not left to any interpretation. However, when a certain amount of subjectivity was at stake, the primary reference was the meaning and the connotation as they were expressed in the contexts under analyse – in another words, the intention of the author.

### 6.2. Interpretation of the data on the frame the Generative Lexicon (GL) theory

The results obtained in this article are in line with the GL theory developed by the formal semanticist James Pustejovsky (1995), especially when it comes to the extent to which the real richness of the lexical knowledge is mirrored in both representations. First, Pustejovsky proposed meaning alternations as to illustrate the concept of complementary polysemy, such as the Producer/Product alternation for the lexical item *newspaper*: *The newspaper fired its editor*, vs. *John spilled coffee on the newspaper* (Pustejovsky 1995: 31). This is reminiscent of the meaning alternations observed in the case of *destruction*.

The ambiguous behavior of *destruction* may also be described through the concept of *coercion*, which is defined by Pustejovsky (1995) and Pustejovsky and Jezek (2008) as a semantic operation which compels an argument to denote a required meaning through a head lexical item. Pustejovsky claims that some sense variation can be captured through a mechanism which makes a shift in the semantic meaning of an entity. *Coercion*, for instance, explains how a fast car is one that can be driven quickly, while a fast typist is one who can type quickly. By the same token, a good pilot will be someone who is “good at navigating air vessels,” while a good knife is a knife that “cuts well” (Fillmore 1969:123). Likewise, this same mechanism can also explain how ‘the destruction of villages’ will be an act of ravaging buildings, while ‘the destruction of rights’ is the violation of rights, ‘the destruction of water catchment areas’ is the natural end of these areas, and ‘the destruction of wheat crops’ is only a partial kind of devastation or deterioration and so forth. Therefore, in order to investigate *coercion* phenomena, corpora can be very useful as a source of data. Corpora can help us to conduct research on how words are actually used as opposed to the way they should be used. These results can never be achieved if we relied solely on dictionaries, or they would have to be completely exhaustive. Being so, we can briefly list the following advantages of parallel corpora over dictionaries. First, while dictionaries are intended to provide the user with the general prototypical meanings

of headwords, parallel corpora provide more reliable and up to date prototypical meanings in specific knowledge domains. Second, parallel corpora are of valuable help to translators and linguists alike and can help bridge the gap between linguists and translation theorists. Finally, this type of analysis may be carried out with any kind of language pair, subject area or corpora.

## 7. Concluding remarks and suggestions

In this article we have attempted to demonstrate how translation (or parallel) corpora come to the rescue of linguists in their semantic and pragmatic investigations of lexical items. In particular, we have tried to show how EAPCOUNT can reveal the ambiguous behavior of the noun *destruction* through its different translations into Arabic. Our data show that such behavior is too complex and fuzzy to accept an account from one perspective, within one language, by one researcher and restricted to one field of study. For instance it is found that there should be no clear-cut distinction between questions of polysemy, synonymy and collocations and analyses of the semantics of words and pragmatics of words, which are so much overlapping that they seem to be just different sides of the same coin. No researcher can conduct a serious empirical investigation of one phenomenon without seeking information from the other. This state of affairs makes it necessary to seek better cooperation between linguists, on the one hand, and translators and translation scholars on the other. It is high time that translators, translation theorists, translation teachers and linguists became the best of friends, no longer endorsing traditional conflicting views, and eased the reciprocal tension that has plagued linguistic and translation research in the history of both fields of enquiry.

Interestingly, the findings of this study prompt me to make some suggestions on how to enhance such cooperation among linguists, professional translators, translation teachers and trainee translators. (1) Linguists, for instance, are called upon to review their views about translation and translation studies. Translation can now be appraised as a fruitful exercise and thanks to corpora it can be re-associated with the excitement of new discoveries. (2) Professional translators are called upon to make their translations available to novice translators, linguists, and translation theorists. (3) Translation teachers are called upon to introduce recent linguistic approaches and technology in the translation classroom. (4) As for translation students, they are called upon to invest some time and effort in collecting their own corpora, to protect their translation products from inappropriate prototypical or dictionary meanings, and to dispel the misconceptions that still exist about one-to-one equivalence across languages. Therefore, the current study has shown a case where cooperation from all concerned parties is needed, whether they are students, teachers, researchers, domain experts, terminologists, linguists, translators or writers of various text types, in order not to reinvent the wheel (Salhi 2010).

## NOTES

1. See also Pustejovsky (1995: 27):  
*First we leave the gate, then we taxi down, vs. John saw the taxi down the street.*  
*Drop me a line when you are in Boston, vs. We built a fence along the property line.*  
*The judge asked the defendant to approach the bar, vs. the defendant was in the pub at the bar.*

2. *LingvoSoft Online*. Long Island City: Ectaco Corporate Center. Visited on 7 January 2012, <<http://www.lingvozone.com/LingvoSoft-Online-English-Arabic-Dictionary>>
3. <<http://dictionary.sakhr.com/SearchResults.aspx>>.
4. *English <-> Arabic Online Dictionary – Ectaco UK*. Long Island City: Ectaco Corporate Center. Visited on 7 January 2012, <[http://online.ectaco.co.uk/main.jsp?do=e-services-dictionaries-word\\_translate&status=translate&lang1=23&lang2=ar&source\\_id=2118988](http://online.ectaco.co.uk/main.jsp?do=e-services-dictionaries-word_translate&status=translate&lang1=23&lang2=ar&source_id=2118988)>.
5. BA'LABAKKĪ, Munir (2005): *al-Mawrid: qāmūs Inkilizī-'Arabī*. Bayrūt: Dār al-'Ilm lil-Malāyīn.
6. The typography for these entries is kept in its original form and the content is similar to that of their printed forms.
7. *Oxford Advanced Learner's Dictionary*. Oxford: Oxford University Press. Consulted on 10 december 2012, <<http://oald8.oxfordlearnersdictionaries.com/dictionary/destruction>>
8. *Cambridge Advanced Learner's Dictionary & Thesaurus*. Consulted on 10 december 2012, <<http://dictionary.cambridge.org/dictionary/british/destruction?q=destruction#>>
9. MURRAY, James, SIMPSON, J. A., and WEINER, E. S. C. (1989). *Oxford English Dictionary*. 2nd ed. Oxford: Clarendon Press. Note that it does exist an online version as well. See <<http://www.oed.com/view/Entry/51117?redirectedFrom=destruction#eid>>.
10. At the time of publication of this article, the corpus is in the vicinity of 7.5 million tokens.
11. Here, the fact that a great deal of UN texts, especially those issued by the General Assembly and the Security Council, are already divided into small numbered paragraphs makes their alignment easier and more accurate. Some other texts were easily divided into clear-cut paragraphs though their paragraphs are not already numbered. This is the case with texts produced by organizations such as UNESCO (United Nations Educational, Scientific and Cultural Organization) and UNICEF (United Nations of International Children's Emergency Fund). The only difficulty in the segmentation and alignment process came with long paragraphs in some IMF (International Monetary Fund) and UNIDO (United Nations Industrial Development Organization) publications, which for practical reasons, have been further divided into smaller paragraphs. But the real challenge was to decide at which point and on which basis such paragraphs are to be divided.
12. Anthony Laurence's Website, The AntConc Homepage. Visited on 12 december 2012, <[http://www.antlab.sci.waseda.ac.jp/antconc\\_index.html](http://www.antlab.sci.waseda.ac.jp/antconc_index.html)>.
13. AntConc is not tied to any particular language and can be used with English-Arabic texts. The *Unicode utf8* encoding must be selected instead of *Western Europe "Latin1"* (iso-8859-1).
13. The adjective 'polysemous' is used in Pustejovsky (1995) to qualify both a word and a sense. A polysemous sense is one which is exhibited along with others by the same word.

## REFERENCES

- ALLERTON, David J. (1984): Three (or four) levels of word co-occurrence restriction. *Lingua*. 63:17-40.
- ANTHONY, Laurence (2006): Concordancing with AntConc: An Introduction to Tools and Techniques in Corpus Linguistics. *JACET Newsletter*. (55):2085.
- BAKER, Mona (1992): *In other words: a coursebook on translation*. London/New York: Routledge.
- BAKER, Mona (1993): Corpus Linguistics and Translation Studies: Implications and Applications. In: Mona BAKER, Gill FRANCIS and Elena TOGNINI-BONELLI, eds. *Text and Technology: In Honour of John Sinclair*. Amsterdam/Philadelphia: John Benjamins, 233-250.
- BAKER, Mona (1996): Corpus-based Translation Studies: The Challenges that Lie Ahead. In: Harold SOMERS, eds. *Terminology, LSP and Translation: Studies in Language Engineering in Honour of Juan C. Sager*. Amsterdam/Philadelphia: John Benjamins, 175-186.
- BAKER, Mona (1998): Norms. In: Mona BAKER, ed. *Routledge Encyclopedia of Translation Studies*, London/New York: Routledge, 163-165.
- BAKER, Mona (2003): Corpus-based Translation Studies in the Academy. In: Heidrun GERZYMISCH-ARBOGAST, Eva HAJICOVÁ, Petr SGALL, Zuzana JETMAROVÁ, Annely ROTHKEGEL and Dorothee ROTHFUSS-BASTIAN, eds. *Textologie und Translation: Jahrbuch Übersetzen und Dolmetschen 4/II*, Tübingen: Gunter Narr, 7-15.
- CARTER, Ronald Allan (1998): *Vocabulary: Applied linguistic perspectives*. 2<sup>nd</sup> ed. London: Routledge.
- CHURCH, Kenneth Ward and HANKS, Patrick (1990): Word association norms, mutual information and lexicography. *Computational Linguistics*. 16(1):22-29.

- CRUSE, D. Alan. (1986): *Lexical Semantics*. Cambridge: Cambridge University Press.
- DI MARCO, Chrysanne, HIRST, Graeme and STEDE, Manfred (1993): The semantic and stylistic differentiation of synonyms and near-synonyms. *AAAI Spring Symposium on Building Lexicons for Machine Translation*. AAAI Technical Report SS-93-02, 120-127.
- FIRTH, John Rupert (1957): A synopsis of linguistic theory, 1930-1955. In: PHILOLOGICAL SOCIETY. *Studies in Linguistic Analysis*. Oxford: Basil Blackwell, 1-32.
- KENNY, Dorothy (2001): *Lexis and Creativity in Translation. A Corpus-based Study*. Manchester: St. Jerome Publishing.
- KENNY, Dorothy (2009): Corpora. In: Mona BAKER and Gabriela SALDANHA, eds. *Routledge Encyclopedia of Translation Studies*. 2<sup>nd</sup> ed. Oxford: Taylor & Francis, 59-62.
- LAVIOSA, Sara (1997): How Comparable Can 'Comparable Corpora' Be? *Target*. 9(2):289-319.
- LAVIOSA, Sara (1998): The Corpus-based Approach: A New Paradigm in Translation Studies. *Meta*. 43(4):474-479.
- LAVIOSA, Sara (2002): *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam/New York: Rodopi.
- LYONS, John (1995). *Linguistic Semantics. An Introduction*. Cambridge: Cambridge University Press.
- MALMKJÆR, Kristen (1998): Love thy Neighbour: Will Parallel Corpora Endear Linguists to Translators? *Meta* 43(4):534-541.
- MEYER, Charles F. (2002) *English Corpus Linguistics*. Cambridge: Cambridge University Press.
- PUSTEJOVSKY, James (1991): The Generative Lexicon. *Computational Linguistics*. 17(4):409-441.
- PUSTEJOVSKY, James (1993a): Type Coercion and Lexical Selection. In: James PUSTEJOVSKY, ed. *Semantics and the Lexicon*. Dordrecht: Kluwer Academic Publishers, 73-94.
- PUSTEJOVSKY, James (1993): *Semantics and the lexicon*. Dordrecht: Kluwer Academic Press.
- PUSTEJOVSKY, James (1995): *The Generative Lexicon*, Cambridge: MIT Press.
- PUSTEJOVSKY, James and Jezek, Elisabetta (2008): Semantic Coercion in Language: Beyond Distributional Analysis. *Rivista di Linguistica*. 20(1):181-214.
- ROSCH, Eleanor (1973): Natural categories, *Cognitive Psychology*. 4:328-350.
- SALHI, Hammouda (2005): *Logical Polysemy and Human Translation*. Master thesis, unpublished. Tunis: University of Carthage.
- SALHI, Hammouda (2010): Small Parallel Corpora in an English-Arabic Translation Classroom: No Need to Reinvent the Wheel in the Era of Globalization. In: Said M SHIYAB, Marilyn Gaddis ROSE, Juliane HOUSE, and John DUVAL, ed. *Globalisation and Aspects of Translation*. Newcastle: Cambridge Scholars Publishing, 53-67.
- SALHI, Hammouda (2012): Investigating Lexical Ambiguity in and through Translation: a Corpus-based Study. PhD thesis, unpublished. Tunis: University of Carthage.
- SINCLAIR, John (1991): *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- SINCLAIR, John (1998): Large corpus research and foreign language teaching. In: Robert DE BEAUGRANDE, Meta GROSMAN and Barbara SEIDLHOFER, eds. *Language Policy and Language Education in Emerging Nations*. Advances in Discourse Processes, vol. LXIII. Stamford: Ablex, 79-86.
- SINCLAIR, John (1999): The lexical item. In: Edda WEIGAND, ed. *Contrastive Lexical Semantics. Current Issues in Linguistic Theory*, vol. 17. Amsterdam/Philadelphia: John Benjamins, 1-24.
- SNELL-HORNBY, Mary (1988): *Translation Studies: An Integrated Approach*. Amsterdam/Philadelphia: John Benjamins.
- STUBBS, Michael (2001): *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell Publishers.
- WEINREICH, Uriel (1964): Webster's Third: a critique of its semantics. *International Journal of American Linguistics*. 30:405-409.
- ZANETTIN, Federico (2002): Corpora for translation practice. In: Elia YUSTE-RODRIGO, ed. *Language Resources for Translation Work and Research*. (LREC 2002 Workshop Proceedings, Las Palmas de Gran Canaria, 29-31 May 2002), 10-14.
- ZANETTIN, Federico, BERNARDINI, Silvia and STEWART, Dominic, eds. (2003): *Corpora in Translator Education*. Manchester/Northampton: St Jerome.