redit

Revista Electrónica de Didáctica
de la Traducción e Interpretación

The challenge of teaching the translation of collocations: Extraction and comparison of collocations across languages applied to the training of translators

El reto de enseñar a traducir las colocaciones. Extracción y comparación de colocaciones entre lenguas aplicadas a la formación de traductores

GUADALUPE RUIZ YEPES 0000-0002-8781-1759
Universidad de Ciencias Aplicadas de Heilbronn (Alemania)

ABSTRACT

For anyone interested in the translation of collocations and/or the comparison of collocations across languages there is an essential issue that has to be dealt with beforehand: the successful extraction of collocations from corpora. The aim of this article is twofold. Firstly, to compare collocation extraction methods from corpora, and secondly, to compare the use of collocations in the training of translation using the LSP of marketing for English, German and Spanish. To achieve these two objectives first of all the collocations are defined and clearly distinguished from other phrases. It will be shown how to extract collocations from large corpora focusing on hybrid methods that combine linguistic and statistical information. After describing two of these methods, they will be applied to the corpora. The findings will be analysed and compared in order to show the challenges of translating collocations to future translators.

KEY WORDS: training of translators, contrastive phraselogy, collocations, phraseology of marketing.

RESUMEN

Dado que la traducción de unidades fraseológicas siempre ha sido un reto, la fraseología contrastiva ha pasado a ocupar un papel central en la formación de traductores. Para cualquier persona interesada en la traducción de colocaciones y/o en la comparación de colocaciones entre lenguas, hay una cuestión esencial que debe tratarse de antemano: la extracción satisfactoria de colocaciones a partir de corpus. Por lo tanto, el objetivo de este artículo es doble. En primer lugar, comparar los métodos de extracción de colocaciones a partir de corpus y, en segundo lugar, comparar el uso de colocaciones en el entrenamiento de la traducción utilizando el lenguaje especializado de marketing para el inglés, el alemán y el español. Para lograr estos dos objetivos, en primer lugar se definen las colocaciones y se distinguen claramente de otras frases. Se mostrará cómo extraer colocaciones de grandes corpus centrándose en métodos híbridos que combinan información lingüística y estadística. Tras describir dos de estos métodos, se aplicarán a los corpus. Los resultados se analizarán y compararán con el fin de mostrar los retos de la traducción de colocaciones a los futuros traductores.

PALABRAS CLAVE: formación de traductores, fraseología contrastiva, colocaciones, fraseología de marketing.

1. INTRODUCTION

While native speakers are familiar with fixed expressions such as *to pay attention* or *to fill in a form*, these types of collocations usually are a frequent source of errors for those who dare to write a text in a foreign language. For learners of foreign languages collocations are usually easy to decode owing to their rather transparent meaning, but difficult to encode because they are unpredictable and very often do not preserve the meaning of their components across languages (Seretan 2011: 1). Collocations pose in particular a challenge for translation trainees. The fact that we *cover the table* in German (*den Tisch decken*), but put the table in Spanish (*poner la mesa*) and *lay/set the table* in English is very difficult to explain from a semantic point of view. The semantic compatibility of two words does not guarantee their frequent combination in a language, which does not only apply to general language, but in particular to languages for specific purposes (LSP from now on). The well-known conclusion that LSPs do not only consist of terminology but also of syntactic features and specific text structures is of particular relevance for this study given that the formation of collocations is very often domain specific. Words which do not participate in a collocation in everyday language often do form part of a collocation in an LSP, i.e. the noun *file* collocates with the verbs *create*, *delete*, *save* in texts about computers, but not necessarily in other contexts (McKeown and Radev, 2000: 510). This conclusion opens the doors for research in the field of LSP phraseology. Picht was the first author to show interest in LSP phraseology in the field of terminology research and pointed towards the importance of general verbs when used in specific phrases (1987, 1988).

The main aim of this paper[1] is to compare the collocations extracted from an *ad hoc* multilingual comparable corpus —German, English, Spanish[2]— of marketing texts in order to explain differences in their structure and use that can help translation students overcome the challenge of translating them. However, before being able to compare multilingual collocations these must be extracted from the corpus. Therefore, two different hybrid methods for the extraction of collocations from corpora will be presented and compared in the first place. In order to achieve this aim, the paper is structured as follows. In Section 2 the language of marketing is described. Sections 3 and 4 deal with the definitions of collocations and how to distinguish them clearly from other related phrases. In Section 5 the methods which have so far been used to extract collocations from corpora will be described focusing on hybrid methods which combine linguistic and statistical information. In Section 6 two of these methods will be applied to the corpora and the findings will be analysed. Finally, in Section 7 the methods and the findings will be compared and their usefulness in the training of translators will be highlighted.

---

1   Some results of this study were presented in the Europhras Conference 2017 in London when it was in its pilot stage.

2   We will consider Peninsular Spanish, the German used in Germany and British English. We are not going to focus on linguistic variety differences.

2. THE LANGUAGE OF MARKETING AS A LSP

When looking at LSPs, many research approaches can be highlighted. However, none of them is autonomous, since they always represent areas of research related to each other. In some approaches, LSPs are defined with reference to the subject and the specific goal; some other compare LSPs with general language, and a third group deals with LSPs by looking at the features of the linguistic tools applied (compare Bausch 1976; Beier 1980; von Hahn 1983; Fluck 1996; Hoffmann 1993; Birkenmaier 1991; Cabré 1999 and others).

According to Fluck (1996:11), there has been much discussion in Linguistics on the exact nature of LSP and a widely accepted definition does not exist. However, most authors are of the opinion that LSP is a variety of general language, has developed from general language and uses the grammatical means of general language (Arntz et al. 2014). There is also a wide consensus on the fact that LSP serves the purpose of ensuring understanding among different communication partners on specialized areas as effective, precise and economic as possible (Schmitt 1985: 18).

For the sake of this piece of research we are going to assume that the formation of collocations in LSP follows the same rules as in general language and that any software designed for the extraction of collocations from general language is able to extract collocations from LSP with the same precision.

There have been several attempts to classify business language. Hundt's (1995) typology, which is one of the most frequently quoted, distinguishes between different communicative areas in which people talk about business: daily life, institutions and theory/science (1995:8). Within the area of theory/science, the LSP of macroeconomics and business administration are the core of the language of economics, with marketing language as a subcategory of the LSP of business administration. According to Bongard (2000), Marketing is a subdiscipline of Business Administration.

Meffert's definition of marketing explains its central philosophy: "In its classical interpretation, marketing means planning, coordinating and monitoring all company activities on current and potential markets. Company goals are to be met by the constant satisfaction of customer needs.[3]" (2015: 12). Marketing as a discipline is transnational and, therefore, strongly influenced by the English language. A considerable number of terms and their related expressions and phrases were borrowed from the English language because both English and American theorists were pioneers in this discipline. Not only did they influence the German marketing language, but also the Spanish one very strongly (Feix 1980: 84). The language of marketing has not been examined much so far, therefore the number of

---

3   Own translation.

scientific papers about it is relatively small. Significant studies are limited to the field of terminology. The classification of the terminology of marketing depending on to which LSP the terms belong to is of particular interest. As a subdiscipline of economics, marketing shows a so-called economic basic terminology which can be seen in all other subdisciplines of economics. Based on the scientific character most text types in marketing (experimental studies, case reports, etc.) have, in addition, terms from the specific languages of statistics and mathematics. However, the most important group contains the terms from the field of business psychology which are indispensable for the marketing language. These words, often emotive and sentimental, serve to describe the customer´s behaviour and needs (Konovalova and Ruiz Yepes 2016: 102).

3. WHAT IS A *COLLOCATION?*

There are numerous definitions for the phenomenon *collocation*. Originally the term *collocation* was used in a very broad sense to describe the "general event of recurrent word co-occurrence" (Seretan 2011: 13). But this frequency-based view has been replaced later by a linguistically motivated one, in which the items in a collocation are syntactically and semantically related. Over the years it has been suggested to use the term *co-occurrence* for the recurrent co-appearance of two words, while the term *collocation* is reserved for the phraseological (linguistic) approach. This distinction between co-occurrences and collocations seems to be accepted (Bartsch 2004), and will be adopted in this study.

Since the term collocation was introduced by Firth (1957), researchers from all kinds of research areas have dealt with it. Consequently, there are definitions of *collocation* from phraseology, computational linguistics, corpus linguistics, etc. Each discipline tries to define *collocation* to meet its needs. As Smadja states: "depending on their interests and points of view, researchers have focused on different aspects of collocation" (1993: 145). As a consequence, the term is very vague and not clearly defined. Despite this lack of clarity, two traditions can be distinguished: one follows Firth´s empirical postulate within the British contextualism, the other has its origin in the German-French lexicography (Hausmann 1984, Mel'čuk 2001, 2006). Contextualists consider that the only way the study of language has to be done is considering the context in which the words appear. They argue that the meaning of words is defined by their co-occurrence with other words, as Firth states: "you shall know a word by the company it keeps!" (1957: 179). Hausmann´s focus is on the semantic interrelationship of these words. For him "the collocation consists of a base which is semantically independent and therefore co-creative and a collocator which is affine or collocative to the base" (Hausmann 1984: 401). It can be said that the base "bears most of the meaning of the collocation and triggers the use of the collocator"

(McKeown and Radev, 2000: 512). Hausmann´s collocation typology (1989: 1010) distinguishes six ty-pes of collocations:

i. verb + noun (object) or noun (object) + verb

ii. adjective + noun

iii. noun (subject) + verb

iv. noun + (prepositional phrase) + noun

v. adverb + adjective

vi. verb + adjective or adverb + verb

Hausmann´s typology can be considered a milestone in the research about collocations, since many other authors have made subsequently use of it in their work. For instance, Gloria Corpas Pastor applies this typology to describe collocations in the Spanish Language (1996: 66-77).

In this paper, both Firth´s empiric tradition and Hausmann´s phraseological approach are considered.

4. DISTINCTION BETWEEN THE TERM *COLLOCATION* AND OTHER WORD COMBINATIONS

In Phraseology, collocations are at the interface between free word combinations and idioms. Accor-ding to McKeown and Radev (2000: 508) "an idiom, [...], is a given rigid word combination to which no generalities apply; neither can its meaning be determined from the meaning of its parts nor can it participate in the usual word-order variations". On the other hand, according to Cowie (1981: 223-235) a free word combination can be described using the general rules of grammar, for example, conside-ring the semantic constraints on the words which appear in a certain syntactic relation with a given headword. "Collocations fall between these extremes and it can be difficult to draw the line between categories" (McKeown and Radev, 2000: 508).

Whether collocations are understood as an independent category or as a sub-category of phraseo-logical units depends on whether we are dealing with a "narrow" or a "wider" conception of phraseo-logy (cf. Reder 2006: 44). While Corpas Pastor considers them phraseological units (1996: 52), other authors like Garcia-Page (2008) and Koike (2003) follow the "narrow" perception of praseology leaving collocations unexplored.

Table 1 shows examples of free word combinations, collocations and idioms.

Structure	Free word combinations	Collocations	Idioms
Verb + noun	To see the table	To set/lay the table	To kick the bucket
Noun + noun	The end of the road	Bar of chocolate	Couch potato

Table 1. Examples of free word combinations, collocations and idioms

To determine if and to which extent a word co-occurrence is really a collocation and not a free word combination or an idiom, idiomaticity and stability, together with other aspects must be verified. With this purpose in mind, we present two criteria to identify collocations after the implementation of quantitative methods.

First criterion: verify that the word co-occurrence is not an idiom. As stated above, an idiom is a word combination whose meaning cannot be determined from the meaning of its parts. Therefore the first criterion for identifying a collocation is: word combination whose overall meaning *can* be derived from the meaning of each word. But idioms are not only on the semantic level fixed associations of words. In contrast to collocations they also can display a fixed syntactic behaviour not allowing modifiers, the passive voice, etc.

Second criterion: verify the word co-occurrence does not admit the substitution of one of its components by a synonym without altering the meaning. Research by Pearce (2002) into collocations extraction produced a method based on the substitutions for synonyms within candidate phrases, for example "emotional baggage (a collocation) occurs more frequently than the phrase emotional luggage formed when baggage is substituted for its synonym luggage" (2002: 1530). His method is based on the assumption that in a free word combination, it is possible to substitute one of its components by a synonym without altering too much the meaning. "If a phrase does not permit such substitutions then it is a collocation" (2002: 1533). Therefore, the second criterion is: A collocation can only be identified as such if the speaker has got several collocators available which can be combined with a base from the semantic point of view, but only one of these collocators is preferred in use. This is why the lexical combination *warm greetings* is accepted, but not *hot greetings*, *a running commentary*, but not *a running discussion*, etc.

For non-native speakers collocations are unpredictable because they do not preserve their meaning across languages (compare Seretan 2011: 2). Literal translations would lead to unnatural and awkward sounding formulations called by some linguists anti-collocations (compare Pearce 2001). Table 2 is based on Smadja's (1993) cross linguistic comparison of collocations.

			Literal translation	Correct translation
Free word combinations	German	Den Tisch sehen	To see the table	To see the table
	Spanish	Ver la mesa	To see the table	To see the table
Collocations	German	Den Tisch decken	To cover the table	To lay/set the table
	Spanish	Poner la mesa	To put the table	To lay/set the table

Table 2. Cross linguistic comparison of collocations based on Smadja (1993)

5. EXTRACTION OF COLLOCATIONS FROM CORPORA

Just as there are different definitions of collocation, there are also different methods of extracting collocations from corpora. The following methods have been used so far: methods based on co-occurrence considerations, methods based on collocation patterns and hybrid methods. The rest of this section will present them.

5.1. Methods based on co-occurrence considerations

Depending on research interests and research purposes, different association measures have been developed. So the corpus linguist has to decide in favour of certain association measures. According to Evert (2009: 1236), there is no perfect association measure. Therefore several ones should be used for a study in order to get different results which can be compared. There are two groups of association measures that pursue opposing goals: *effect-size measures* (Mutual Information, Dice, log odds ratio) and *significance measures* (z-score, t-score, chi-squared, log-likelihood). The linguist using effect-size measures is looking to find how much the observed co-occurrence frequency exceeds the expected frequency, wheras significance measures try to determine how unlikely is the null hypothesis that the words are independent (Evert, 2009: 1228).

5.2. Methods based on Collocation Patterns

As mentioned above, Hausmann's collocation typology (1989: 1010) distinguishes six types of collocations. Weller and Heid (2010: 3195) call these types *collocation patterns*, once extracted from corpora they call them *collocation candidates*. In order to extract collocations from corpora using collocation patterns the corpora have to be annotated at least with POS-tags. However, a higher performance can be reached if the corpora are also syntactically analysed (parsed).

5.3. Hybrid methods

Pamies and Pazos (2005: 317-329) compare different association measures and come to the conclusion that mathematical methods alone are not enough and prefer instead the use of *hybrid* methods (2005: 327). There are methods which use a combination of co-occurrence calculation and linguistic criteria in the form of collocation patterns. That is, a hybrid system combines statistical methods and multilingual parsing for detecting accurate collocational information. But, in which order are the filtering methods to be used? There are different approaches:

- First statistical methods are used and —for refining the filtering— additional collocation patterns are applied to the achieved results (Smadja 1993).

- First collocation patterns are used and in the second step the statistical methods are applied to the achieved results (Krenn (2000), Seretan and Wehrli (2006).

Smadja (1993) is the most representative researcher for the first approach. He developed a system called Xtract that retrieved word pairs using a frequency-based metric in the first place. The metric computed the z-score of a pair of words. In addition to the metric, Xtract used three additional filters based on linguistic properties. As a final step, an evaluation of the retrieved collocations was carried out by a lexicographer in order to estimate the number of the true lexical collocations retrieved.

Another example for the first approach is ConcGram 1.0. This software was developed as an inclusive search engine for phraseological units and works on the basis of co-occurrence considerations. It is left to the linguist which of these co-occurrences are significant word combinations and which are chance word co-occurrences. After entering a command, the software compiles a list of *unique words* which are the basis for showing *ConcGrams*. These can be both adjacent and non-adjacent word co-occurrences which can appear in any order in the corpus. As soon as the ConcGram lists are compiled, statistical methods can be applied. They allow a reduction of the lists and provide clear information on non-relevant word combinations which can be ignored. The applied statistical methods are t-score and MI tests whose formulas are explained in detail by Barnbrook (1996: 88-106). When doing computer-assisted corpus analysis though, automatically compiled frequency tables are —despite using associations measures— not always directly usable, but need a human selection input. For this reason, when working with this kind of software the researcher, or in the case of the training of translators the professor, has to apply manually collocation patterns in order to be able to extract collocations candidates. In this paper we also improve the process by adding work stages based on the criteria presented in section 4.

On the other hand, Seretan and Wehrli are the most representative researchers for the second approach. They consider that "syntactic analysis of source corpora is an inescapable precondition for

collocations extraction" (2006: 1). The hybrid method Seretan and Wehrli (2006) developed relies on a deep parser called Fips (Wehrli, 2004) and can be seen as a two-stage process. Firstly the collocation candidates are identified by the parser while POS-tagging and parsing the text corpora. Secondly the candidates are scored and ranked using specific association measures (Seretan and Wehrli, 2006: 2). In this approach the parser is used in the first stage of the extraction in order to identify the collocation candidates and the criterion they employ firstly for the selection of the collocation candidates is the syntactic proximity. As Seretan and Wehrli explain, "as the parsing goes on, the syntactic word pairs are extracted from the parse structures created, from each head-specifier or head-complement relation. The pairs obtained are then partitioned according to their syntactic configuration" (2006: 2). An advantage of this method is that the pairs obtained are partitioned according to the collocations patterns presented by Hausmann (1984). A major disadvantage, however, is the dependence on a specific linguistic theory. Finally, the log-likelihood test is applied.

Since hybrid methods give better accuracy over any single method, they will be adopted in this article.

6. EXTRACTION AND COMPARISON OF COLLOCATIONS FROM A CORPUS USING HYBRID METHODS

John Sinclair, one of the pioneers of Corpus Linguistics, defined a corpus as "… a collection of naturally-occurring language text, chosen to characterize a state or a variety of a language" (1991: 171). It is now well known that there is no such a thing as a perfect corpus. Every researcher compiles his/her corpora of texts according to which the purpose of each research project is.

In this paper a comparable corpus will be analysed. The corpus is comprised of three corpora in three languages (German, Spanish and English) which are comparable because they belong to the same text type. All three corpora consist of articles on marketing topics, published in scientific journals.

In the case of the English corpus, all articles were published in *The Journal of Marketing Management* in the period from 2000 to 2013. The English corpus consists of 1,131,744 words. In the German corpus, all articles were published in *Der Markt – International Journal of Marketing* in the period from 2000 to 2012. The German corpus contains 903,430 words. The Spanish corpus comprises articles which were published in *Revista española de Investigación de Marketing* in the period from 2004 to 2014. It consists of 1,279,954 words.

6.1. Statistical methods first: ConcGram

In the case of the hybrid methods that focus first on statistical information, very often collocation candidates are "selected directly from plain text, with a combinatorial procedure applied to a limited context" (Seretan 2011: 58). This is the case of ConcGram. While this basic procedure works fine for English it has been proven "inefficient for languages such as German or Korean which exhibit richer morphology and a freer word order" (Seretan 2011: 58).

By using the software ConcGram 1.0, all articles of the same language in one corpus have been merged before undergoing analysis. As explained in section 5, concgram lists, which are based on lists of unique words, are compiled by using the software. Once the t-score and MI-tests have been conducted, the manual extraction of collocations by applying Hausmann´s collocation patterns can be started. However, this method is only used with the English corpus because we were not able to extract collocation candidates with ConcGram from the Spanish and German Corpora. The German and the Spanish corpora are analysed semasiologically. That means that the extraction of collocations is only done in the English corpus. Based on the extraction of collocations —and their terminology— in the English language, the German and Spanish counterparts are identified by means of concordances with the format keyword in context so that they can be compared.

Figure 1 shows how to create a ConcGram List. This software does not offer the possibility to select or filter collocations by collocations patterns.
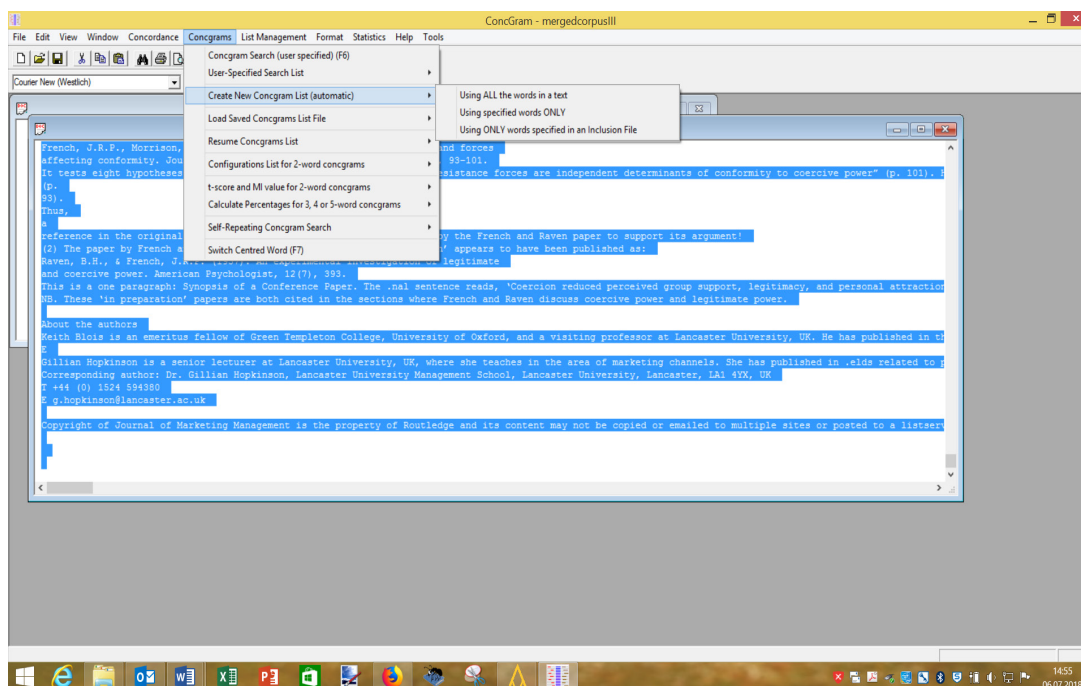

Figure 1. ConcGram screenshot on how to create a ConcGram List

The terms *satisfaction* and *loyalty* occur extraordinarily frequently in the English corpus. The occurrence of the Spanish and German equivalents is similarly frequent. Therefore, certain co-occurrences of the words *satisfaction* and *loyalty* were extracted and compared with their German and Spanish analogue co-occurrences using the software ConcGram. The words *satisfaction* and *loyalty* —in German *Zufriedenheit* and *Loyalität* and in Spanish *satisfacción* and *lealtad*— belong to the category of business psychology in the LSP of marketing. They occur particularly frequently with the words *consumer/customer*, *consumidor/cliente* and *Konsument/Kunde*. To verify if these word combinations were collocations in the light of the two criteria explained in section 4 were applied. Once checked that they met the criteria, a cross linguistic comparison was carried out. Tables 3, 4 and 5 show these collocations in English, German and Spanish.

Base	Collocation	Frequency	Structure
Loyalty	Customer loyalty	83	Noun + noun
	Consumer loyalty	7	Noun + noun
Satisfaction	Customer satisfaction	137	Noun + noun
	Consumer satisfaction	0	Noun + noun

Table 3. Collocations of *satisfaction* and *loyalty* in the English corpus

Base	Collocation	Frequency	Structure
Zufriedenheit	Kundenzufriedenheit	197	Copulative compound
	Konsumentenzufriedenheit	1	Copulative compound
Bindung	Kundenbindung	286	Copulative compound
	Konsumentenbindung	0	Copulative compound
Loyalität	Kundenloyalität	19	Copulative compound
	Konsumentenloyalität	0	Copulative compound

Table 4. Collocations of *Zufriedenheit/Loyalität/Bindung* in the German corpus

Base	Collocation	Frequency	Structure
Lealtad	Lealtad del cliente	17	Noun + prep. + art.+ noun
	Lealtad del consumidor	31	Noun + prep. + art. + noun
Satisfacción	Satisfacción del cliente	79	Noun + prep. + art. + noun
	Satisfacción del consumidor	61	Noun + prep. + art. + noun

Table 5. Collocations of *satisfacción* and *lealtad* in the Spanish corpus

The comparison of the three languages English, German and Spanish has led to the conclusion that the collocations in one language are often expressed by other types of word combinations in the other languages. The English collocations with the structure "noun + noun" are often expressed by compounds or possessive markers in the German language, and by prepositional phrases —either with or without article— in the Spanish language.

These findings also provide evidence of the fact that collocations do not preserve their meaning across languages. While the compound *Kundenloyalität* occurs in the corpus only 19 times, the compound *Kundenbindung* occurs 286 times in the same contexts as *customer loyalty* occurs in the English corpus. Therefore we can assume that *Kundenbindung* is the German equivalent for *customer loyalty* rather than *Kundenloyalität* which is the literal translation.

Last but not least, these findings have also shown that while in the Spanish language the terms *consumidor* and *cliente* are used in the same way, in German and English the terms *Kunde* and *customer* are preferred. This is probably related to socio-linguistic and pragmatic factors because the words with the root *Konsum-/consum-* have a negative semantic prosody (see, e.g., Stubbs 1995; Smith and Nordquist 2012) or connotation in both English and German, but not in the Spanish language.

6.2. Collocation patterns first: FipsCo

For hybrid methods that focus first on linguistic and second on statistical information the main criterion for selecting a pair as a collocation candidate is the existence of a syntactic link between the two words. This is the case of the deep parser FipsCo; "Binary collocation candidates are identified from the parse structures built by Fips as the analysis of the text goes on" (Seretan 2011: 65). Therefore the identification of collocation candidates takes places after the parsing has been completed, but these two steps are not entirely separated. They alternate; as soon as a sentence has been analysed, the candidates are identified. According to Seretan (2011: 63) the parser is able to deal with a large range of syntactical constructions, i.e., passivization, relativization, interrogation, cleft constructions, enumeration, coordinated clauses, interposition of subordinate clauses, interposition of parenthesised clauses, apposition, etc.

Some of the most representative syntactic configurations used by FipsCo to extract collocations coincide with the collocation patterns by Hausmann (1984) presented in section 3, that is, "verb + noun (object)" or "noun (object) + verb", "adjective + noun", "noun (subject) + verb", "noun + (prepositional phrase) + noun", "adverb + adjective" und "verb + adjective" or "adverb + verb". This opens a wide range of possibilities allowing the translation trainer show his/her translation trainees different collocations from different perspectives.

Figure 2 shows how collocations can be filtered using Hausmann´s collocations patterns.
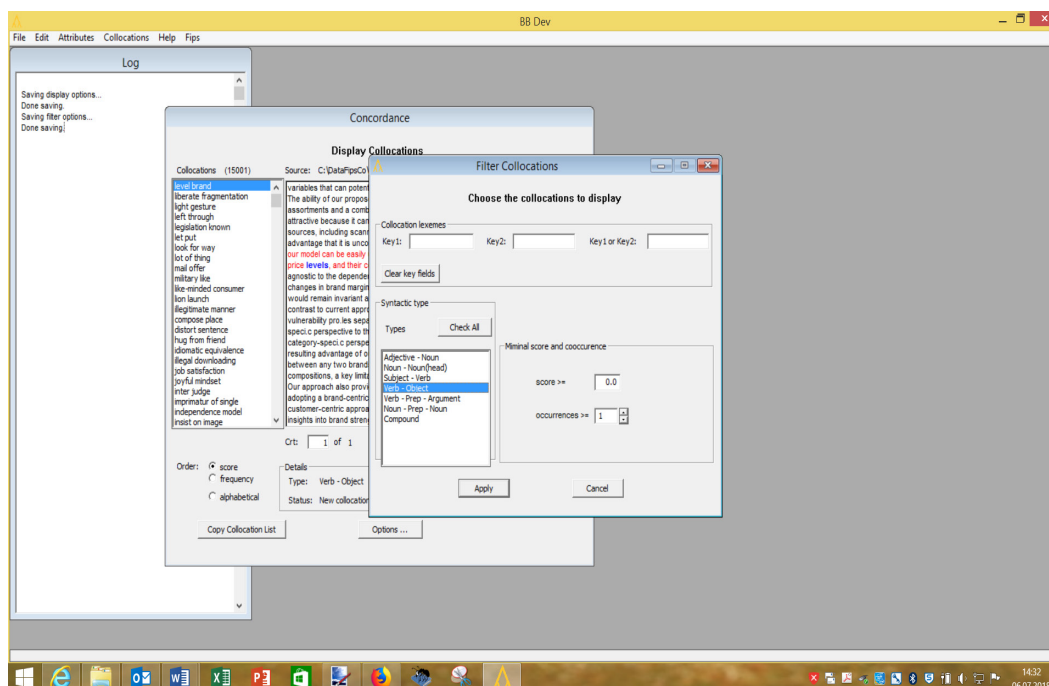
Figure 2. FipsCo screenshot that shows the filtering using collocations patterns

Table 6 shows the first 10 collocations of each language of the kind "verb + noun (object)/ noun (object) + verb".

English (frequency)	German (frequency)	Spanish (frequency)
Copy content (186)	Rolle spielen (70)	Percibir valor (183)
Have effect (128)	Frage stellen (40)	Percibir calidad (124)
Take place (81)	Werbung sehen (26)	Tener efecto (104)
Provide insight (70)	Tabelle sehen (24)	Llevar a cabo (93)
Have impact (69)	Anforderung erfüllen (23)	Percibir riesgo (81)
Play role (61)	Beitrag leisten (20)	Comprar producto (71)
Buy produce (58)	Effect haben (19)	Utilizar escala (70)
Create value (56)	Abbildung sehen (16)	Sentir este (58)
Post express (55)	Produkt bewerten (14)	Percibir utilidad (57)
Email article (51)	Produkt anbieten (13)	Dar lugar (56)

Table 6. First 10 English, German and Spanish collocations of the kind "verb + noun (object)/noun (object) + verb"

The most frequent collocations in the English corpus of the kind "verb + noun (object)" are general language collocations like: *take place*, *have impact* or *play role*. The collocation *buy produce* appears in the English corpus so frequently because in one of the research articles in the English corpus a survey was carried out where customers were asked if they *buy local produce*. Therefore sentences like *I buy*

local produce because it supports local producers appear over 50 times. The collocation *copy content* is a similar case. It is the most frequent collocation in the English Corpus of the kind "verb+noun (object)", but it has to do with the fact that every article or research paper has a paragraph about the copyright. It is therefore *have effect* the collocation that appears most often in the corpus in the running text.

As for the Spanish corpus while the most frequent collocations in Spanish seem to be specific of the language of Marketing like *percibir valor*, *percibir calidad*, *percibir riesgo* and so on, this is not the case because the collocation *percibir valor* appears only in the form *valor percibido*, i.e., *el valor percibido por el cliente...*, which means that it is a collocation of the type "noun + adjective". The same is applicable for the collocation *percibir calidad*, i.e., *la calidad percibida por el cliente...*, *percibir riesgo*, i.e. *para la medición del riesgo percibido*. Therefore, the most frequent collocation with the structure "verb + noun (object)" in the Spanish corpus is *tener efecto* (have effect) which is again a general language collocation. The collocation *sentir este* is obviously an error. The Spanish expression *En este sentido* appears in the Spanish corpus 58 times and means *in this respect*. FipsCo has interpreted the noun *sentido* as the verb *sentir* and the pronoun *este* as a noun. Examples from the corpus are: *en este sentido, la investigación desarrollada en el campo de la psicología...*, *en este sentido, los objetivos principales de este trabajo son*. There are other examples of errors committed during the parsing of the Spanish corpus that are not as frequent but still worth mentioning like *parir consumidor* (which means *to give birth to the consumer*) and *parir variable* (which means *to give birth to a variable*). The word combination *para el consumidor* appears in the Spanish corpus 36 times and means *for the consumer* (*En categorías de producto en las que el número de alternativas disponibles **para** el **consumidor** es elevado...*). FipsCo has interpreted the preposition *para* (*for*) as the verb *parir* (*to give birth*).

The most frequent collocations of the type "noun (object) + verb" in German are also general language collocations, like *Rolle spielen* (*to play a role*) and *Frage stellen* (*to ask a question*). The collocation *Werbung sehen* (*to see commercial*) is the first one specific to the language of Marketing. The collocations *Tabelle sehen* and *Abbildung sehen* appear only as *siehe Tabelle* (*see table*) and *siehe Abbildung* (*see figure*). These expressions are typical of academic writing and are used to allude to the tables and figures in the text.

Table 7 shows the first 10 collocations of each language of the kind "adjective + noun". FipsCo provides lemmatised collocations, therefore it has removed the endings for gender and number of the adjectives.

In the English corpus the most frequent collocations with the structure "adjective + noun" are typical for the field of E-Marketing like *social media* and *social network*, as well as collocations typical of traditional Marketing like *corporate brand* and *national brand* and collocations typical of academic discourse like *future research* and *previous research*. For the German corpus the collocations typical

of the academic discourse are the most frequent like *empirische Untersuchung*, *vorliegende Studie*, *empirische Studie* and *empirisches Studium*. To the same field can be ascribed the Spanish collocation *futura investigación*.

English (frequency)	German (frequency)	Spanish (frequency)
Social media (150)	Integriert Kommunikation	Nuevo producto (205)
Corporate brand (97)	Positiv Effect (71)	Siguiente hipótesis (59)
Future research (93)	Empirisch Untersuchung (70)	Nuevo tecnología (50)
High level (93)	Vorliegend Studie (69)	Futuro línea (49)
National brand (91)	Alt Mensch (66)	Mayor medida (47)
Social network (87)	Empirisch Studie (64)	Máximo verosimilitud (45)
Previous research (85)	Wahrnehmen Risiko (63)	Mismo modo (41)
Subjective norm (80)	Empirisch Studium (60)	Alto nivel (41)
Dependent variable (71)	Partitionieren Preis (58)	Futuro investigación (37)
Individual use (70)	Latent Variable (58)	Diverso autor (36)

Table 7. First 10 English, German and Spanish collocations of the kind "adjective + noun"

As opposed to the English and German language, where adjectives usually go before the nouns they modify, in Spanish adjectives usually come after the nouns they modify. FipsCo has the option to search for collocations of the kind "adjective + noun" but also of the kind "noun + adjective". In figure 3 the most frequent Spanish collocations of the type "noun + adjective" are shown: *sitio web* (*web site*), *efecto moderador* (*moderating effect*), *red social* (*social network*), *variable dependiente* (*dependent variable*), *efecto directo* (*direct effect*), *efecto positivo* (*positive effect*), *promoción monetario* (*monetary promotion*), *variable moderador* (*moderating variable*), *palabra clave* (*key word*) and *posición competitivo* (*competitive position*)[4]. The collocations *sitio web* and *red social* can also be seen as belonging to the field of E-Marketing. The main principle governing the choice of placing the adjective before or after the noun in Spanish language has to do with differentiation. Therefore, an adjective following a noun distinguishes that item from others that may have different qualities, i.e., *el coche rojo* (the red car) as opposed to el *coche blanco* (the white car), but placing an adjective before a noun implies that the quality expressed is naturally associated with that noun. In this second case, rather than describing the noun in order to differentiate it from others, the adjective merely attaches an unsurprising epithet to it.

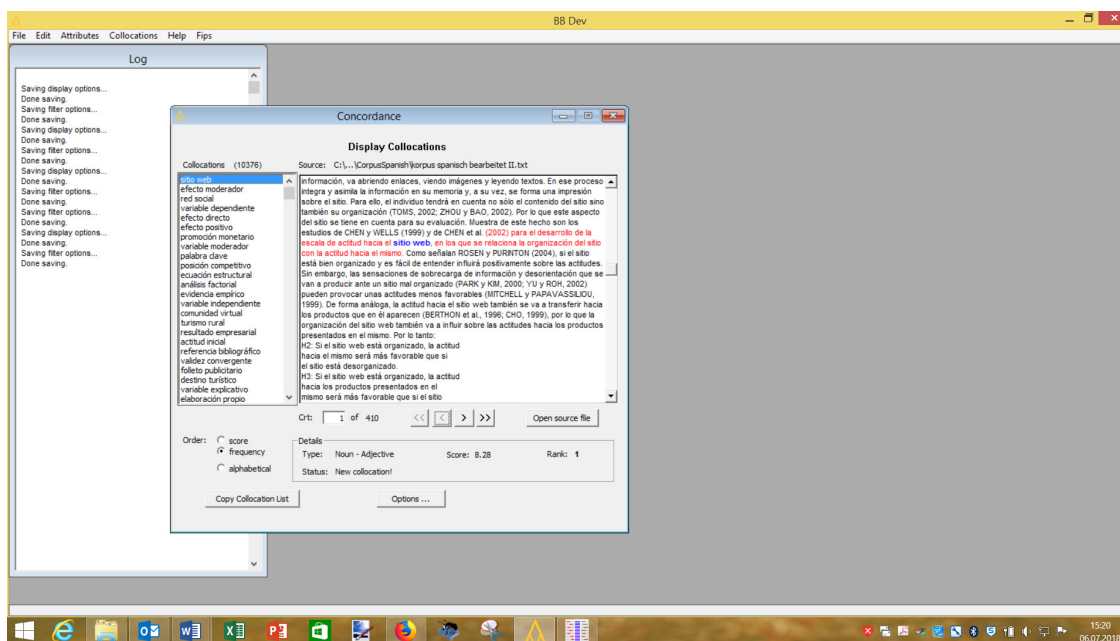**4**   Translations by the author of this article.

Figure 3. FipsCo screenshot showing the most frequent Spanish collocations of the type "noun + adjective"

Table 8 shows the first 10 collocations of each language of the kind "noun + noun". FipsCo has separated the German compounds in order to match the structure "noun + noun" collocation. For example, Zielgruppe into *ziel gruppe*, *Kundenbindung* into *kunde bindung*, *Kundenzufriedenheit* into *kunde zufriedenheit* and so on. It is very interesting to observe that the deep parser was able to identify German compounds as "noun + noun" collocations.

English (frequency)	German (frequency)	Spanish (frequency)
Market orientation (247)	Ziel gruppe (361)	Orientación de mercado (264)
Brand management (153)	Kunde bindung (230)	Agencia de viaje (208)
Consumer behaviour (138)	Marke name (187)	Categoría de product (197)
Product category (99)	Kunde zufriedenheit (179)	Boca a boca (157)
Relationship marketing (93)	Innovation kultur (175)	Imagen de marca (147)
Research interest (87)	Erfolg faktor (172)	Cuota de mercado (131)
Marketing strategy (82)	Zeit punkt (169)	Intención de compra (131)
Business school (76)	Kunde orientierung (160)	Precio de referencia (126)
Resource integration (75)	Einfluss faktor (131)	Marca de distribuidor (122)
Supply chain (69)	Auskunft person (124)	Comportamiento del consumidor (107)

Table 8. First 10 English, German and Spanish collocations of the kind "noun + noun"

The fields to which the most frequent collocations of this kind belong, for the German corpus it is the specialised field of Marketing. Not only *Kundenbindung* and *Kundenzufriedenheit* are typical of this field, but also *Zielgruppe*, *Markenname* and *Kundenorientierung*.

As for the English corpus the collocations belong to the semantic field of Marketing with the exception of *research interest*, which belongs to the academic discourse, and *business school*, which appears in the addresses of the authors of the journal articles that constitute the corpus and can therefore not be considered as part of the running text.

The ten most frequent collocations of the kind "noun + noun" in the Spanish corpus extracted by FipsCo are all accompanied by a preposition as in *orientación de mercado*, *agencia de viaje*, *boca a boca*, etc. The conclusion drawn in section 6.1 stating that the collocations in one language are often expressed by other types of word combinations in other languages is being taken into account by the deep parser FipsCo, because it extracted English collocations with the structure noun + noun, that are often expressed by compounds or possessive markers in the German language, and by prepositional phrases —either with or without article— in the Spanish language.

Table 8 shows that the third most frequent collocation with the structure "noun + noun" in the English corpus is *consumer behaviour*, which is specific to the LSP of Marketing and belongs to the group of terms from the field of business psychology (Author 2016: 102). On a closer look at this collocation it became obvious that for this specific case the word *consumer* is preferred over *customer*. With the base *loyalty* and *satisfaction* the preferred collocator is *customer* as shown in section 6.1, but with the base *behaviour* the preferred collocator is *consumer* as shown in Table 9. A search in Linguee has revealed that the word *behaviour* occurs more often with words with negative denotations, such as *punishable, sluggish, repressive, myopic, radical, fundamentalist, gamy,…* than with words with positive or neutral denotations, such as *autonomous, observed, conscious, differing, …*.

Base	Collocation	Frequency	Structure
Behavior	Customer behavior	56	Noun + noun
	Consumer behavior	228	Noun + noun

Table 9. Collocations with the base *behaviour*

The tenth most frequent collocation with the structure "noun + prep. + noun" in the Spanish corpus is *comportamiento del consumidor*, which is the Spanish equivalent for *consumer behaviour*. Again there is an obvious preference for the term *consumidor* versus *cliente* combined with *comportamiento*. As shown in section 6.1 in the Spanish language the word *consumidor* does not seem to be preferred in general language to the word *cliente*. They are both used indifferently with a slight tendency to use

consumidor in the LSP of marketing according to the searches in the above mentioned corpus[5]. Table 10 shows that in this case the singular is preferred over the plural.

Base	Collocation	Frequency	Structure
comportamiento	Comportamiento del cliente	2	Noun + prep. + noun
	Comportamiento del consumidor	107	Noun + prep. + noun
	Comportamiento de los clientes	2	Noun + prep. + noun
	Comportamiento de los consumidores	5	Noun + prep. + noun

Table 10. Collocations with the base *comportamiento*

In the German corpus the collocation *Konsumentenverhalten* does not belong to the group of the ten most frequently used collocations. Nevertheless, it is again evident that the co-occurrence of the base *Verhalten* with *Konsument-(en)* prevails over the co-occurrence with *Kunde(n)*. A search in Linguee has revealed that the word *Verhalten* occurs more often with words with negative denotations, such as *vertragswidrig, fahrlässig, illoyal, pflichtwidrig, ungewöhnlich,...* than with words with positive or neutral denotations, such as *kaufmänisch, rechtmäßig, ethisch,....*

Table 11 shows the frequency of the collocations "noun + noun" (which in German appear as copulative compounds) with the base *Verhalten* and the collocators *Kunde* and *Konsument*.

Base	Collocation	Frequency	Structure
Verhalten	Konsumentenverhalten	63	Copulative compound
	Kundenverhalten	9	Copulative compound
	Konsumverhalten	11	Copulative compound
	Verbraucherverhalten	2	Copulative compound

Table 11. Collocations with the base *Verhalten*

The content of these tables reveals the truth of Firth's dictum (1957: 11): "You shall know a word by the company it keeps". In the analysed German and English corpora the words *Konsument* and *consumer* seem to have negative connotations. Therefore they appear more frequently with words that have a tendency to cooccur with such words like *behaviour* and *Verhalten*. Nonetheless, in the Spanish corpus analysed the word *consumidor* occurs as often as *cliente*. Therefore this phenomenon is only visible for the German and English LSP of marketing.

---

[5] Searches in the CREA (Corpus de Referencia del Español Actual) have shown similar results.

7. CONCLUSIONS: COMPARISON AND USEFULNESS FOR THE TRAINING OF TRANSLATORS

7.1. Comparison of the hybrid methods used

The results of the collocation extraction by ConcGram (focusing on statistical information) and by FipsCo (focusing on linguistic information) are quite different. The collocation candidates displayed by ConcGram as the most frequent in the English and Spanish corpora are not the ones displayed by FipsCo. The German corpus is the only one that contains collocations displayed by ConcGram like *Kundenbindung* and *Kundenzufriedenheit*. In terms of the research alternatives and opportunities for the teacher of translation, the hybrid method that focuses first on linguistic patterns and second on statistics is more suitable.

Since ConcGram displays all co-occurrences that appear more frequently than expected for the size of the corpus there is usually a lot of noise in the extracted collocations candidates lists, and the translation trainer has to extract interesting collocations manually. As presented in section 6.2, with FipsCo the researcher or or trainer does not need to do any kind of manual search or filtering of collocation candidates and the criteria for the identification of collocations presented in section 4 do not have to be implemented. Every step is carried out by FipsCo; the parsing and the filtering of collocations candidates.

Some of the conclusions drawn after working with ConcGram had already been implemented in the deep parser FipsCo, like the fact that collocations can have different structures in different languages. FipsCo has extracted "noun + noun" collocations from the English corpus, but was able to extract compounds in German and split them into "noun + noun" collocations. For Spanish FipsCo extracts equivalent collocations with the structure "noun + prep. (art.) + noun".

FipsCo's usefulness for the training of translators are very broad compared to ConcGram. FipsCo allows the teacher to choose the collocation pattern he or she wants to analyse, which is not the case with ConcGram. While ConcGram only searches for collocations of the kind "noun + noun", FipsCo searches for collocations applying Hausmann`s collocation patterns.

To summarise, hybrid methods for collocation extraction that consider first linguistic criteria offer a much wider range of research opportunities than hybrid methods that focus on statistical information first.

7.2. Comparison of the collocations in the languages English, German and Spanish and the consequences for the teaching of translation

The comparison of the three languages English, German and Spanish has led to the conclusion that the collocations in one language are often expressed by other types of word combinations in the other languages.

These findings also provide evidence of the fact that collocations do not preserve their meaning across languages. As mentioned in section 6.1, while the compound *Kundenloyalität* occurs in the corpus only a few times, the compound *Kundenbindung* occurs 286 times in the same contexts as *customer loyalty* occurs in the English corpus. Therefore we can assume that *Kundenbindung* is the German equivalent for *customer loyalty* rather than *Kundenloyalität*, which is the literal translation. Literal translations lead very often to unnatural and awkward sounding formulations.

These findings have also shown that while in the Spanish language the terms *consumidor* and *cliente* are used indifferently, in German and English the terms *Kunde* and *customer* are preferred. This is related to sociolinguistic factors since the words with the root *Konsum-/consum-* have a negative semantic prosody in both English and German, but not in the Spanish language.

Table 12 is based on table 2 of this article. It is also based on Smadja's (1993) cross linguistic comparison of collocations and presents the results of the research with ConcGram.

			Literal translation	Correct translation
Free word combination	**German**	Produkte verkaufen	To sell products	To sell products
	Spanish	Vender productos	To sell products	To sell products
Collocations	**German**	Kundenbindung	Customer commitment	Customer loyalty
	Spanish	Lealtad del consumidor	Consumer loyalty	Customer loyalty

Table 12. Cross linguistic comparison of collocations in the LSP of marketing based on Smadja (1993)

Table 12 shows once more that literal translations of collocations often give origin to so-called anti-collocations.

As for the results of the analysis carried out in section 6.2 it can be said that the ten most frequent collocations with the structure "verb + noun (object)" belong to general language and the collocations with the structure "adjective + noun" and "noun + noun" belong to the LSP of Marketing and E-Marketing. For English, German and Spanish the findings were very similar.

Among the numerous fields of applications of this study, translation training must be particularly highlighted. For translators, especially when translating into the foreign language, collocations are a frequent error source, all the more when it comes to detecting false friends. The translation teacher has a base for explaining to the students that not only grammatical aspects are important for the use of a language, but that there are also uses which can neither be explained by grammatical nor semantical rules. Aspects which are related to the use of a language, which are pragmatic and can only be taught and learned in connection with the culture and the values of a society.

Translation trainers need corpus-based studies that explore the use of collocations in several languages in a systemic, rigorous and consistent way in order to be able to wake the awareness of translation students regarding the fact that collocations need a special treatment and can not be translated as if they were free word combinations.

Over the years many examples of outstanding scientific corpus-based works devoted to translation studies (cf. Baker 1995 and 1999, Corpas Pastor 2008, Hu 2016), to the use of collocations in LSP (cf. Torner and Bernal 2017) or the training of translators (cf. Bernardini 2004 and 2016, Vargas-Sierra 2014) can be found. Nonetheless there is a need to improve studies that focus on how to teach the translation of collocations.

REFERENCES

Arntz, Reiner; Heribert Picht and Klaus-Dirk Schmitz. 2014. *Einführung in die Terminologiearbeit*. Hildesheim: Georg Olms Verlag.

Baker, Mona. 1995. Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223–243.

Baker, Mona. 1999. The role of corpora in investigating the linguistic behaviour of professional translators. *International Journal of Corpus Linguistics*, 4(2), 281–298.

Barnbrook, Geoff. 1996. *Language and Computers: A practical Introduction to the Computer Analysis of Language*. Edinburgh: Edinburgh University Press.

Bartsch, Sabine. 2004. *Structural and functional properties of collocations in English, A corpus study of lexical and pragmatic constraints on lexical co-occurrence.* Tübingen: Narr.

Bausch, Karl-Heinz (eds.). 1976. *Fachsprachen*. Terminologie, Struktur, Normung. Berlin, Köln: Beuth Verlag GmbH.

Beier, Rudolf. 1980. *Englische Fachsprache*. Stuttgart, Berlin, Köln, Mainz: Kohlhammer.

Bernardini, Silvia. 2004. Corpus-aided language pedagogy for translation education. In K. Malmkjær (ed.). *Translation in undergraduate degree programmes*. Amsterdam and Philadelphia: John Benjamins Publishing, 97-111.

Bernardini, Silvia. 2016. Discovery learning in the language-for-translation classroom: corpora as learning aids. *Cadernos de Tradução*, 36(1), 14–35.

Birkenmaier, Willy; Mohl, Irene. 1991. *Russisch als Fachsprache*. Tübingen: Francke.

Bongard, Joachim. 2000. *Werbewirkungsforschung. Grundlagen-Probleme-Ansätze*. Hamburg/London: LIT Verlag Münster.

Cabré, María Teresa. 1999. La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos. Textos en castellá, catalá i frances. Barcelona: Universidad Pompeu Fabra, Institut Universitari de Lingüística Aplicada II.

Corpas Pastor, Gloria. 1996. *Manual de fraseología española*. Madrid: Gredos.

Corpas Pastor, Gloria. 2008. *Investigar con corpus en traducción: los retos de un nuevo paradigma.* Frankfurt: Peter Lang.

Cowie, Anthony P. 1981. The treatment of collocations and idioms in learner's dictionaries. *Applied Linguistics*, 2/3: 223-235.

Evert, Stefan. 2009. Corpora and collocations. In Anke Lüdeling and Merja Kytö eds. *Corpus linguistics: an international handbook*. Berlin: De Gruyter, 1212-1248.

Feix, Nereu. 1980. Die Terminologie der Wirtschaftswissenschaften (Volkswirtschaft) im Sprachvergleich: Spanisch und Deutsch. In José Rodríguez Richart, Gisela Thome und Wolfram Wilss ed. *Fachsprachenforschung und –lehre: Schwerpunkt Spanisch*. Tübingen: Narr, 81-87.

Firth, John Rupert. 1957. A Synopsis of Linguistic Theory 1930-55. *Studies in Linguistic Analysis*. Oxford: The Philological Society, 1-32.

Fluck, Hans-Rüdiger. 1996. *Fachsprachen: Einführung und Bibliographie*. München: Francke.

García-Page, Mario. 2008. *Introducción a la fraseología española*. Barcelona: Anthropos.

Hahn, Walter von. 1983. Fachkommunikation: Entwicklung, linguistische Konzepte, betriebliche Beispiele. Berlin/New York: de Gruyter.

Hausmann, Franz Josef. 1984. Wortschatzlernen ist Kollokationslernen. Zum Lehren und Lernen französischer Wortverbindungen. *Praxis des neusprachlichen Unterrichts,* 31, 395-406.

Hausmann, Franz Josef. 1989. Praktische Einführung in den Gebrauch des Student's Dictionary of Collocations. In Evelyn Benson, Morton Benson and Robert Ilson eds. Cornelsen. iv-xviii.

Hoffmann, Lothar. (1993) „Fachwissen und Fachkommunikation. Zur Dialektik von Systematik und Linearität in den Fachsprachen." In Bungarten, Theo (Hrsg.): *Fachsprachentheorie.* Bd. 2: Konzeptionen und theoretische Richtungen. Tostedt, S. 595-617.

Hu, Kaibao. 2016. *Introducing Corpus-based Translation Studies*. Berlin/Heidelberg: Springer.

Hundt, Markus. 1995. *Modellbildung in der Wirtschaftssprache: zur Geschichte der Institutionen- und Theoriefachsprachen der Wirtschaft.* Tübingen: Niemeyer.

Koike, Kazumi. 2003. Las unidades fraseológicas del español: Su distribución geográfica y variantes diatóticas. *Epos* XIX: 47-65.

Konovalova, Anastasia and Guadalupe Ruiz Yepes. 2016. Die Sprache des Marketings und ihre Übersetzung: morphologische und semantische Aspekte der Terminologie. *MonTI* 8: 95-123.

Krenn, Brigitte. 2000. *The usual suspects: data-oriented models for the identification and representation of lexical collocations.* Saarbrücken: DFKI and Universität des Saarlandes.

McKeown, Kathleen R. and Dragomir R. Radev. 2000. Collocations. In Robert Dale, Hermann Moisl and Harold Somers, eds. *A handbook of natural language processing*. New York: Marcel Dekker, 507-523.

Meffert, Heribert, Christoph Burmann and Manfred Kirchgeorg. 2015. *Marketing. Grundlangen marktorientierter Unternehmensführung. Konzepte – Instrumente – Praxisbeispiele*. Heidelberg: Springer.

Mel'čuk, Igor. 2001. *Communicative organization in natural language: The semantic-communicative structure of sentences*. Amsterdam and Philadelphia: John Benjamins.

Mel'čuk, Igor. 2006. Explanatory Combinatorial Dictionary. In Giandomenico Sica, ed. *Open problems in linguistics and lexicography*. Monza: Polimetrica, 225-355.

Pamies, Antonio and Jose Manuel Pazos. 2005. Extracción automática de colocaciones y modismos. In J.D. Luque Durán and A. Pamies Bertrán (Eds.), *La creatividad del lenguaje: colocaciones idiomáticas y fraseología*. Granada: Granada Lingvistica, 317-330.

Pearce, Darren. 2001. Synonymy in Collocation Extraction. In *NAACL 2001 Workshop: WordNet and Other Lexical Resources: Applications, Extensions and Customizations*. Carnegie Mellon University, Pittsburgh.

Pearce, Darren. 2002. A comparative evaluation of collocation extraction techniques. Proceedings of the Third International Conference on Language Resources and Evaluation (LREC-2002), Las Palmas, Canary Islands – Spain, 1530-1536.

Picht, Heribert. 1987. Fachsprachliche Phraseologie – Die terminologische Funktion von Verben. In Hans Czap and Christian Galinski, eds. *Terminology and Knowledge Engineering*. Frankfurt: INDEKS-Verlag, 21-34.

Picht, Heribert. 1988. Fachsprachliche Phraseologie. In Reiner Arntz ed. *Textlinguistik und Fachsprache*. Hildesheim: Georg Olms Verlag, 187-196.

Reder, Anna. 2006. *Kollokationen in der Wortschatzarbeit*. Wien: Praesens

Seretan, V. and Wehrli, Eric. 2006. Accurate collocation extraction using a multilingual parser. In Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL 2006), Sydney, 952-960.

Seretan, Violeta. 2011. *Syntax-based collocation extraction.* Heidelberg: Springer.

Sinclair, John. 1991. *Corpus, concordance, collocation.* Oxford : Oxford University Press.

Smadja, Frank. 1993. Retrieving collocations from text: Xtract. *Computational Linguistics,* 19/1: 143-177.

Smith, K. Aaron and Dawn Nordquist. 2012. A critical and historical investigation into semantic prosody. *Journal of Historical Pragmatics* 13/2: 291-312.

Schmitt, Peter A. (1985) Anglizismen in den Fachsprachen: eine pragmatische Studie am Beispiel der Kerntechnik. Heidelberg: Winter.

Schmitt, Peter A. (1990) Die Berufspraxis der Übersetzer. Eine Umfrageanalyse. Berichtssonderheft des Bundes-verbandes der Dolmetscher und Übersetzer. Bonn: BDÜ.

Stubbs, Michael. 1995. Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2/1: 23-55.

Torner Castells, Sergi and Elisenda Bernal Gallen (eds.). 2017. *Collocations and other lexical combinations in Spa-nish. Theoretical and Applied approaches.* London and New York: Routledge.

Vargas-Sierra, Chelo. 2014. Innovación didáctica en traducción especializada: sobre la enseñanza virtual de tra-ducción de páginas web de contenido económico. In Gallego-Hernández, Daniel (ed.). *Traducción econó-mica: entre profesión, formación y recursos documentales, VERTERE*. Monográficos de la Revista Hermeneus: 110-130.

Weller, Marion und Ulrich Heid. 2010. Corpus-derived data on German multiword expressions for lexicography. Proceedings of the Euralex International Congress, 2010 Leeuwarden [CDROM].