Meta

Journal des traducteurs
Translators' Journal

ΜΕΤΑ

Interference and the Translation of Phraseological Units in a Parallel and Multilingual Corpus

Zuriñe Sanz-Villar

Résumé de l'article

Dans cet article, nous rapportons les résultats de l'analyse de traductions d'unités phraséologiques (UP) d'un corpus de textes traduits de l'allemand en basque, en nous intéressant notamment à la forme que les interférences (Toury 1995/2012) prennent dans la traduction littéraire de ces deux langues. Nous montrerons ainsi que l'on trouve des interférences « typiques » de la langue source dans les textes traduits et que d'autres interférences découlent d'autres langues et d'autres textes connus du traducteur, en l'occurrence, dans notre étude, des interférences de l'espagnol. Afin de réaliser une telle analyse, nous avons constitué un corpus numérique, parallèle et multilingue. Ce dernier contient des textes sources (TS) en allemand, des textes cibles en basque, ainsi que leurs versions intermédiaires, en cas de traduction indirecte.

Interference and the Translation of Phraseological Units in a Parallel and Multilingual Corpus

ZURIÑE SANZ-VILLAR
University of the Basque Country, Vitoria-Gasteiz, Spain
zurine.sanz@ehu.eus

RÉSUMÉ

Dans cet article, nous rapportons les résultats de l'analyse de traductions d'unités phraséologiques (UP) d'un corpus de textes traduits de l'allemand en basque, en nous intéressant notamment à la forme que les interférences (Toury 1995/2012) prennent dans la traduction littéraire de ces deux langues. Nous montrerons ainsi que l'on trouve des interférences «typiques» de la langue source dans les textes traduits et que d'autres interférences découlent d'autres langues et d'autres textes connus du traducteur, en l'occurrence, dans notre étude, des interférences de l'espagnol. Afin de réaliser une telle analyse, nous avons constitué un corpus numérique, parallèle et multilingue. Ce dernier contient des textes sources (TS) en allemand, des textes cibles en basque, ainsi que leurs versions intermédiaires, en cas de traduction indirecte.

ABSTRACT

This paper will report on findings obtained from a corpus-based translation analysis of phraseological units on texts translated from German into Basque, with special emphasis placed on the shape that interference (Toury 1995/2012) takes in German-into-Basque literary translations. It will be shown not only that "typical" source language interference can be found in translated texts, but also that interference arises from other texts and/ or languages present in the translator's mind, most commonly in this study; Spanish. For the purpose of carrying out such an analysis, a digitized, parallel and multilingual corpus was compiled consisting of German source texts (STs), Basque target texts (TTs) and their corresponding intermediary versions in cases of indirect translation.

RESUMEN

El objetivo de este artículo es dar a conocer los resultados obtenidos de un análisis basado en corpus sobre la traducción de unidades fraseológicas (UF) en textos traducidos del alemán al vasco, y se hará especial hincapié en la interferencia (Toury 1995/2012) que pueda darse en dichas traducciones literarias. Los resultados muestran que en los textos traducidos no solo encontramos la interferencia «típica» ejercida por la lengua original, sino también otro tipo de interferencia que surge de otros textos y/o idiomas —en el caso del presente estudio, sobre todo del español— que están presentes en la mente del traductor. Para llevar a cabo el análisis, se ha creado un corpus digitalizado, paralelo y multilingüe compuesto por textos originales alemanes (TOs), textos meta en vasco (TMs) y las correspondientes versiones intermedias en el caso de las traducciones indirectas.

MOTS-CLÉS/KEYWORDS

interférence, phraséologie, corpus, *Descriptive Translation Studies*, basque
interference, phraseology, corpora, Descriptive Translation Studies, Basque
interferencia, fraseología, corpus, estudios descriptivos de traducción, vasco

1. Introduction

The main goal of the present paper is to present some of the findings obtained during the completion of my PhD thesis (Sanz-Villar 2015) on the corpus-based translation analysis of phraseological units (PUs) in the language combination German-Basque.

Translation activity in the Basque Country has played an important role in the development of Basque, a minority language which coexists unequally with a major language: either that of Spanish in the Basque Country and Navarre, or with French in the provinces situated within the French Pyrénées Atlantiques department, namely Labourd, Basse Navarre and Soule. It is argued here that this imbalanced coexistence affects both the process of translation and its resultant translated products.

In the specific case of German-into-Basque translations, one form of the dominant language's influence is related to the number of *indirect translations*; i.e. texts that have not been translated from the original text, the German source text (ST), but from an intermediary text or texts, most commonly here, a Spanish version. Similarly, when translating directly – that is, when translations are made drawing mainly on the German ST – it is also not uncommon for the translator to consult other translations, given that often Spanish, French or English translations are available prior to the Basque version's publication. Thus, it is very interesting to examine not only the relationship between ST(s) and target texts (TT), but also to what extent traces of interference can be observed in their resulting Basque translations.

For this purpose, the AleuskaPhraseo corpus, which allows the study of ST(s) interference on translations, was built from scratch. It is a parallel and multilingual corpus consisting of literary texts that also includes intermediary texts in cases where the corresponding ST has been translated indirectly. While the catalog which contains German-into-Basque translations was being compiled and updated, the publishers and translators of the texts were contacted in order to determine the translation mode used for them – that is, if the TTs were translated directly from the ST(s) or indirectly through an intermediary text.[1]

2. Translation Studies and Interference

Baker (1993: 243-245) proposed a list of candidates for translation universals. Not only was interference not included in this list, but a clear distinction was even made between the proposed translation universals and interference, the former being considered as features of translated texts and not "the result of interference from specific linguistic systems" (Baker 1993: 243).

Later, in order to explain regular translation behavior, Toury (1995/2012) proposed both norms and laws – the law of growing standardization and the law of interference –, which are probabilistic in nature and serve as heuristic tools. According to the law of interference, translators tend to accept phenomena pertaining to the structure of the source text and pass them to the target text.

This phenomenon has become the object of investigation in a variety of studies (Mauranen 2004; Kenny 2005; Lefer and Vogeleer 2013, to mention but a few), and recently the interest in analysis of interference has grown, probably due to the proliferation of corpus-based studies using not only comparable but also parallel corpora (Marco and van Lawick 2015).

The studies included in the volume edited by Lefer and Vogeleer (2013) focus on the analysis of interference and normalization. One of the conclusions drawn in the introduction of the volume regarding interference is that it "appears as the most remarkable and inherent property differentiating translated language from original (non-translated) language" (2013: 16).

From a different computational and quantitative perspective, Volansky (2012) sought to distinguish between translated and non-translated language. For that purpose, he employed text classification techniques as a "tool to study the makeup of translated texts in a general way" (Volansky 2012: 10). Thirty-two features were selected and analyzed to determine if they were relevant in distinguishing translated texts from originals (Volansky 2012: 8). In terms of interference, four features (POS n-grams, character n-grams, contextual function words and positional token frequency) were examined, and in a similar fashion to Lefer and Vogeleer (2013), Volansky concludes that interference "is the most robust phenomenon typifying translations" (2012: 30).

Beyond linguistic factors that may serve to identify interference, other factors, such as cognitive machinery and socio-cultural context play an important role in the translation process. Despite, "the fact that translators *are* a (somewhat special) kind of bilinguals is rarely taken into account in translation studies" (Lefer and Vogeleer 2013: 8), however, it is clear that having a *multilingual brain* influences the translation process. Translators that have translated from German into Basque are, at least, bilingual, and apart from the non-native languages the translator may know, will have a native command of both Basque and Spanish (or French). Typically, the dominant language exerts a greater influence upon Basque than the other way around, and, as stated in Zubillaga (2013) and Sanz-Villar (2015), the influence of the Spanish language will not only be present in translations made indirectly – that is, when translating through an intermediary Spanish translation –, but also in direct translations from German.

Thus, not only will cases of *textual interference* be found (where interference is created by the text, be that the German original or the Spanish intermediary version), but also cases of *cognitive interference* (where the Basque language's diglossic status[2] and the translator's multilingual brain are behind the interference). A final type of interference that needs to be considered is *instrumental interference,* which relates to the bilingual tools that translators usually employ (dictionaries, databases, corpora, etc.).[3] To date there is just one German-Basque dictionary (Martínez 1996),[4] which comprises around 32,000 entries, meaning that in all likelihood, translators will also consult tools that combine German and other major languages, such as Spanish, French, English, etc. It is argued here that the use of these indirect tools may have an impact on translation. The creation of these two concepts, *cognitive interference* and *instrumental interference*, was intended to complement and expand Toury's law of interference.

With regards to socio-cultural context as a factor that influences interference, Toury mentions that attitudes towards interference may vary, by place and culture, with respect to differences in prestige associated with the to-be-translated language and its culture. According to Toury (2012: 314), when translating from a highly prestigious or major language, the tendency to accept interference will be higher, especially when the target language is regarded as weak.

As Basque is a minority language with a short literary tradition, cases of interference are expected to be found in translation analysis, and the acceptance of interference may be reinforced by the fact that the majority of German source authors included in the AleuskaPhraseo corpus are canonical literary authors. However, it is also possible that attitudes towards interference may change when translating indirectly through Spanish translations, because, in an attempt to conceal interference from the Spanish language and thus hide the indirect character of the translation, a tendency to avoid interference and create genuine Basque texts that differ from the Spanish version may be enhanced.

3. Phraseology and Interference

Many researchers (Baldwin and Kim 2010) agree on the quantitative relevance of PUs: "The number of MWEs [multiword expression] is estimated to be of the same order of magnitude as the number of simplex words in a speaker's lexicon." However, there is little consensus either on the terminology used to designate these MWEs (Moon 1998; Burger, Dobrovolskij *et al.* 2007: 1) or on the definition of those multiword units; i.e. there are no clear criteria that are recognized by all experts (lexicologist, lexicographers, etc.) to define PUs (Schemann 2011: 19). In Corpas' definition from 1996, PUs are described as fixed word combinations of at least two words that are characterized by the frequency of use and frequency of co-occurrence of their PU's constituents, by institutionalization, idiomaticity and gradation.

Regarding the polylexical character of PUs, there has been a debate on the inclusion of a new category – the so called **Einwortphraseologismen**, **Einwortphraseme**, **Einwortidiome** ('One-word PUs') – to the phraseological field (Szczęk 2004; van Lawick 2013). These units are compounds that have been excluded from phraseology for not being polylexemic items, but they do meet some of the criteria commonly ascribed to PUs, such as idiomaticity and fixedness. Some examples found in our corpus (see Section 4) represent compounds (**händereibend** or **Händchenhalten**) that originated from PUs. In this respect, van Lawick (2013: 135) and Moon (1998) refer to the "blurring of the boundaries between single-word and multi-word items" (Moon 1998: 8).

Fixedness means that, in comparison with other word combinations created ad hoc, PUs are institutionalized word strings wherein language users use the same lexical items in the same order over and over again. It is therefore regarded as "the key feature" (Fiedler 2007: 19) which characterizes PUs. However, PUs are not as fixed as was previously thought. Indeed, with the arrival of corpus-based studies, it has been suggested that variation is a more widespread phenomenon than was traditionally believed (van Lawick 2013: 133). Thus, as can be read in Burger (2007: 31), fixedness is a relative feature. As Mellado (2004) points out, one only needs to take a look at dictionaries to notice the variable character of PUs, because different lexicographic tools will often provide the user with different forms of the same PU.

As far as idiomaticity is concerned, it can be understood in either a narrow or wide sense. On the one hand, from the former perspective, the term idiomaticity is used to designate semantic irregularity; i.e. when the phraseological meaning is not the same as the collective meaning of the PU's constituents, that is, when the PU is not compositional (Mellado 2004: 42). On the other hand, the latter approach is more

widespread in contemporary research on the subject (Urizar 2012) because idiomaticity is not only understood in the aforementioned semantic sense, but also, as Baldwin and Kim (2010) indicate, in the lexical, syntactic, statistic and pragmatic sense. According to the authors, "*idiomaticity* refers to markedness or deviation from the basic properties of the component lexemes" (Baldwin and Kim 2010: 4) at the multiple levels mentioned above. Other authors such as Moon (1998) also mention syntactic or pragmatic non-compositionality: "Institutionalized strings which are grammatically ill formed or which contain lexis unique to the combination may also be considered non-compositional. Other cases involve what may be termed pragmatic non-compositionality. The string is decodable compositionally, but the unit has a special discoursal function" (1998: 8).

Depending on the PU, the presence or relevance of the above features will vary. As can be read in Moon (1998: 9), there are degrees of institutionalization, fixedness and non-compositionality and "[t]his means that it is difficult to identify cleanly discrete categories of FEI [fixed expressions]." Burger (2007), for instance, classifies PUs that are below the sentence level according to their degree of idiomaticity into collocations, partial idioms and idioms, but he points out that the boundaries between the three types are fuzzy: "Es lassen sich dabei drei hauptsächliche Typen unterscheiden, deren Abgrenzung – entsprechend dem graduellen Character des Kriteriums Idiomatizität – nicht strikt sein kann, zwischen denen vielmehr fließende Übergänge bestehen" (2007: 37).

These are some general features that serve to characterize PUs across languages, but each language may have its own individual peculiarities. As for Basque, due to the extralinguistic characteristics mentioned above – the fact that Basque is a minority language that is still in a process of normalization –, the presence of phraseological calques may be greater than in other more established languages. Alberdi (2010) emphasizes the relevance of using calques – i.e. the transfer of the meanings of words from other languages using elements of the target language – as a resource:

> […] calques are a highly productive resource or mechanism that has often been, and must be, used in the creation of neologisms and new terminology in any language (cf. Gómez Capuz, 2005: 39) and particularly in Basque and languages in similar minority status […]. (Alberdi 2010: 32)

Therefore, it is assumed that in the present corpus, examples of PUs that have been translated literally from either German (in the case of direct translations) or Spanish (when translating both directly and indirectly) will be found. These cases will serve as indicators of interference, because "[…] the direct transfer of idioms can be viewed as cases of interference" (Lefer and Vogeleer 2013: 9).

For the purpose of labeling this phenomenon, that is, "the direct transfer of idioms," the translation technique *PU-direct copy* will be used. With regards to the different translation options, it is necessary to mention that the starting point for creating the list of translation techniques was Marco's (2013) proposal which, itself is based on Delabastita's (1996) translation methods.

As argued before, not only could textual interference but also cognitive interference play a role in German-into-Basque translations. Thus, the *PU-indirect copy* technique will be employed when fingerprints of another text, which is not the German original, are found in the Basque translations. Examples of direct and indi-

rect copies will be presented in Section 6. Aside from these, the following translation options are available when performing translation analysis:

– *PU-similar PU*: when the source text PU and the target text PU share the same phraseological meaning, lexical structure and image, both PUs will be considered as similar;
– *PU-different PU*: when there is a discrepancy with regards to the factors just mentioned (phraseological meaning, lexical structure and image) between both PUs, they will be regarded as different PUs;
– *PU-no PU*: in this case, the target word combination does not constitute a PU;
– *PU-another rhetorical resource*: there is no PU in the target text, and instead, another rhetorical resource (for instance, repetition) is used;
– *PU-Ø*: the counterpart of the source text PU does not feature in the target text;
– *No PU-PU*: there is a PU in the target text, but no PU in the source text; this translation option makes the target text more phraseological than the source text;
– *No PU-indirect copy*: the counterpart of an indirect copy in the TT (that is, a word-for-word copy of another text which is not the assumed original) is a free word combination in the ST.[5]

The examples given in Table 1 are extracted from the corpus and intended to clearly illustrate the different types of translation options found during the analysis.

TABLE 1

Examples of the different translation options extracted from the corpus

Source Text	Target Text	Translation option
1. „Oder wollt ihr den Kaiser <u>mit leeren Händen</u> willkommen heißen? (BUSde[6])	-Edo <u>esku hutsik</u> eman behar al diogu ongietorria Kaiserrari? (BUSeu[7])	*PU-similar PU*
2. Der fahrende Schüler steckte den Beutel ein, <u>gab</u> dem Bürgermeister <u>die Hand</u> und versicherte: (BUSde, see note 5)	Estudiante ibiltariak boltsatxoa gorde, alkateari <u>eskerrak eman</u> eta ziurtatu zuen: (BUSeu, see note 6)	*PU-different PU*
3. „Du läßt die Finger von dem, Herbert, der gehört dir nicht. <u>Gibst mir die Hand drauf?</u>“ (BAde[8])	“Ez hadi arrimatu ere egin, Herbert, ez duk hire arazoa. <u>Hitzemaidak.</u> (BAde[9])	*PU-no PU*
4. [...] kommen <u>Tag für Tag</u> tausend Zeitungen, Zeitschriften, Reden, öffentliche und geheime Sitzungen [...]. (DSWde[10])	[...] <u>egunero egunero</u> milaka egunkari, aldizkari, hitzaldi, topaketa ireki eta ezkutu zetozkion [...]. (DSWde[11])	*PU-another rhetorical resource*
5. Ich bin ein Gewohnheitsmensch <u>durch und durch</u> und liebe keine Neuerungen. (DAde[12])	Ohitura-gizona naiz eta berrikuntzak ez ditut batere maite. (DAeu[13])	*PU-Ø*
6. Meggie hörte die Stimme, auch ohne dass sie an der Wand lauschen musste. (THde[14])	Meggiek <u>argi eta garbi</u> entzun zuen ahots hura, belarria hormaren kontra jarri gabe ere. (THeu[15])	*No PU-PU*

In the first example, the German PU **mit leeren Händen** and the Basque PU <u>esku hutsik</u> share the same phraseological meaning, lexical structure and, therefore, image. In the second case, extracted from the same book, while in the German text the student shakes hands with the mayor (**jemandem die Hand geben**), in the Basque version the student thanks (<u>eskerrak eman</u>) the mayor. The use of a different PU in the TT may indicate, in the case of this children's literature text, an adaptation to the target culture. In the ST of the third example, the German PU **die Hand darauf/ dadrauf/auf das/ein Versprechen/… geben** is employed, which means to promise

something to someone with a handshake. The Basque counterpart represents a single lexical unit, a verb (<u>hitzeman</u>), which means to promise. The next example was classified as *another rhetorical resource* because the word <u>egunero</u> ('every day') is repeated to emphasize its meaning. In the fifth example the PU **durch und durch** ('through and through') has been omitted in the TT. In the Basque translation of the sixth example, the addition of the PU <u>argi eta garbi</u> ('clearly') gives emphasis to the way Meggie heard the voice, i.e., she heard the voice clearly.

4. Methodology: descriptive and corpus-based

Many studies that have analyzed the translation of PUs have done so from a prescriptive perspective (for more details on this matter see Sanz-Villar 2016), and as mentioned by Marco (2009: 843) "the studies mentioned are seldom empirical, and when they are, they move within the narrow limits of manual analysis." The present research, however, is situated within corpus-based Descriptive Translation Studies (DTS) (Toury 1995/2012), in a similar fashion to the studies carried out by some of the members of the COVALT research group, who have analyzed the translation of PUs into Catalan based on the COVALT (Corpus Valencià de Literatura Traduïda) corpus (Bracho 2013).

Following the empirical methodology proposed by DTS, the first step consisted of the creation of a catalog, the Aleuska catalog,[16] containing all literary works that have ever been translated from German into Basque. Then, the criteria for the selection of the works that would constitute the corpus under analysis were defined. First, from the different types of text that are included in the catalog (adult literature, children's literature, poetry, essay and theatre), texts labeled as *adult literature* (AL) and *children's literature* (CL) were selected. This is mainly because narrative texts represent 71% of the total Aleuska catalog. Second, translations prior to 1980 were excluded, mainly due to the fact that the number of German-into-Basque translations started increasing from that year on. The third criterion was related to the translation mode: since direct and indirect translations are both represented in the catalog, both types were included in the corpus. Fourth, author diversity was ensured when selecting the texts, because the aim was not to analyze the style of a specific translator or author. Finally, results of other studies were also considered: for instance, as it has been demonstrated that the books of Erich Kästner are especially PU-loaded (Richter-Vapaatalo 2007), it was made certain that books from this author were included in the corpus.

Once the texts were selected, they were digitized and aligned. Table 2 summarizes the main features of the AleuskaPhraseo corpus. The table is divided into data concerning AL and CL. For each of the two text types, it is possible to see how many direct and indirect translations were selected, the works of how many source or target authors (or translators) were included and the total number of words for each language included. Since in the Aleuska catalog (Sanz-Villar 2015: 151) indirect translations are much more widespread in CL than in AL, the number of indirect texts included in the corpus is greater in CL. All in all, the corpus consisted of around 3.5 million words.

Table 2

Features of the AleuskaPhraseo corpus

	Translation mode		Authors		Number of words		
	Direct	Indirect	Source author	Target author	German	Basque	Spanish
AL	19	5	17	15	1,120,534	935,530	198,274
CL	15	9	13	16	593,871	512,204	166,860
Total	34	14	30	28	1,714,405	1,447,734	365,134

For the simultaneous alignment of multiple texts, a program called *TAligner* was used[17] and then further developed by the computer technician Iñaki Albisua within the aforementioned TRALIMA-ITZULIK research group. The appearance of the program's interface is shown in Figures 1, 2 and 3.

The steps required for the compilation of a corpus are clearly defined in the upper part of the interface: cleaning (*limpiar*), tagging (*etiquetar*) and aligning (*alinear*). Since tagging differs depending on the structure of the texts, on the left side of the interface a selection can be made between narrative texts, theatre or poetry.[18] With narrative texts, tagging involves providing each text with so-called metatextual information or metadata, followed by their division into paragraphs and sentences. Figure 1 shows what a tagged text looks like (*etiquetado*) in the *TAligner* program.

Figure 1

Appearance of a tagged text in the *TAligner* program

Once the XML files are loaded for alignment, the user manually makes the adjustments necessary to create 1 to 1 correspondences between the texts to be aligned. As can be seen in Figure 2, one original German text and two translations (both the Spanish and the Basque version) have been loaded. This is because the Basque text was indirectly translated through the Spanish translation. Figure 2 shows how the aligned texts should look after performing adjustment.

In order for the researcher to be able to make queries of the corpus, the aligned texts have to be uploaded into a database. Once this process is repeated for all the

Figure 2

Three texts aligned with the *TAligner* program

Figure 3

The result of random query in the *TAligner* program

texts that constitute the corpus, searches can be conducted. Figure 3 represents the result of a random query.

At the top of the interface, the search engines' options can be seen. In this case, a very simple query has been carried out: to find the German word **Detektive** and its equivalents in the corpus. Since the word has been found in a text that was translated indirectly, the search results include the German sentence where the word is to be found, together with the sentences preceding and following it. Also returned are the equivalent sentences in Spanish and Basque. The column "code" is necessary for the researcher to know to which text of the corpus each result corresponds.

After the compilation and description of the Aleuska catalog and the creation of the AleuskaPhraseo corpus (or, in other words, after conducting preliminary analysis), a two-fold analysis was carried out: the texts of the corpus were examined at a macrotextual and at a microtextual level. The macrotextual part included the analysis of the paratextual elements (Genette 1987) located within and outside the texts, as well as the translation analysis of those paratexts. Interesting results were obtained from this analysis, mainly regarding the relationship of the ST(s) and the target text. Indeed, examining this relationship was of great relevance throughout the entire study, because in order to conduct the translation analysis, it was essential to know from which ST(s) the translations originated. But the present paper will focus on the microtextual part, where the main goal was to examine the translation of certain PUs.

One important decision that needed to be made concerned the type of PU that would serve as object of study. First, a wordlist was created using *AntConc* (Anthony 2014),[19] and then, a lemma list of the most frequent nouns was created manually; i.e., inflected forms of nouns corresponding to the same lemma were grouped together. Soon it will be possible to create a lemma list automatically, since the corpus is being lemmatized and tagged at the part-of-speech level. As can be observed in Table 3, both German and Basque body parts (such as **Hand**, **Auge**, **Kopf** for German, and buru, esku, begi, aurpegi for Basque) were in the top-ten list of the most frequent nouns of the corpus. As a result of this it was decided that somatic PUs – the ones that have at least one word referring to a part of the human or animal body – would be analyzed first; more specifically (and due to their high frequency), somatic PUs containing the word **Hand** in German and esku in Basque were extracted.[20] After the extraction, different dictionaries (such as Schemann's Deutsche Idiomatik[21] or Duden[22] for German, and Elhuyar[23] for Basque) and corpora (DeReKo[24] in the case of German and Euskal Testuen Corpusa[25] for Basque) were consulted to verify their status as PUs. As Milizia and Spinzi argue, "frequent words have strong phraseological tendencies, and the more frequent a word is the more likely it will appear in multi-word units" (2008: 328). That is why frequency was regarded as an important factor when defining the object of study. In the future, we intend to submit somatic PUs consisting of other parts of the body (as well as other types of PUs) to translation analysis with the aim of contrasting the findings of the present analysis.

TABLE 3
The most frequent nouns German and Basque in the AleuskaPhraseo corpus

Lemma (DE)	Frequency	Lemma (EU)	Frequency
Hand [hand]	2,664	buru [head]	4,366
Herr [mister]	2,387	ama [mother]	3,861
Mann [man]	2,177	etxe [house]	3,209
Mensch [human]	2,132	esku [hand]	3,068
Frau [woman]	1,980	hitz [word]	2,930
Auge [eye]	1,878	gizon [man]	2,775
Haus [house]	1,786	urte [year]	2,730
Kopf [head]	1,598	gauza [thing]	2,314
Mutter [mother]	1,563	begi [eye]	2,216
Leben [life]	1,562	aurpegi [face]	1,633

In order to run the queries in the *TAligner* program, first the words beginning with **Hand** and **Händ** in German and <u>esku</u> in Basque were requested, and then a manual selection was made; i.e. only the results representing PUs were saved and analyzed.

5. Results

As can be concluded from the above, queries were carried out in two directions: first the STs were taken as starting point (ST->TT analysis) and then, following the target-oriented approach, the TTs served as a basis to conduct the queries (TT->ST). In the following section, the results obtained from both analyses will be presented.

5.1. Results of the ST->TT analysis

The total number of occurrences of the lemma **Hand** in the German texts was 2,664, and from those, 302 were identified as somatic PUs. Figure 4 shows the distribution (in percentages) of the translation options across the four subcorpora: Children's Literature translated DIrectly, Adult Literature translated DIrectly, Children's Literature translated INDIrectly and Adult Literature translated INDIrectly.

Figure 4

Distribution of the translation options across the four subcorpora

	CL DI %	AL DI %	CL INDI %	AL INDI %	Total %
PU⇒similar PU	15.52	31.14	7.32	38.89	25.83
PU⇒different PU	29.31	32.93	58.54	30.56	35.43
PU⇒no PU	36.21	23.35	12.20	19.44	23.84
PU⇒direct copy	8.62	9.58	14.63	5.56	9.60
PU⇒indirect copy	10.34	2.40	7.32	2.78	4.64
PU⇒another resource	0.00	0.60	0.00	0.00	0.33
PU⇒∅	0.00	0.00	0.00	2.78	0.33
Total	100.00	100.00	100.00	100.00	100.00

According to these figures, an effort was made to maintain the phraseological character of the STs in the TTs, considering that in 61.26% of the cases, a percentage that results from the sum of the first two options (25.83% and 35.43%), an ST PU was translated with another PU (similar or different). The next most common outcome (23.84%) was represented by those cases in which the counterpart of an ST PU was a free word combination. The percentages of direct and indirect copies, the fourth (9.60%) and the fifth (4.64%) options, were lower (14.24%), but compared to the results of other studies (Marco 2013), the tendency to use word-for-word copies in the Basque TTs is quite high. In other words, if direct and indirect copies serve as an indicator of interference, it can be argued that interference in the present case is a more frequent phenomenon than it is in other cases - where copies only account for 0.39% of the cases (Marco 2013: 174). Figure 5 provides a clearer picture of the differences and similarities between CL and AL texts.

As can be inferred from Figure 5, while the preservation of the phraseological character of the STs is greater in AL translations, the number of copies (especially of the indirect kind) is higher in CL TTs. The greater number of PUs in AL translations may be connected to the prestige of this text type: the higher the prestige, the greater the effort on the translator's part to create more elaborate texts (with more PUs). As for the number of copies, it can be argued that it is more common to find fingerprints of the ST (or another text/language) in CL TTs than in AL translations.
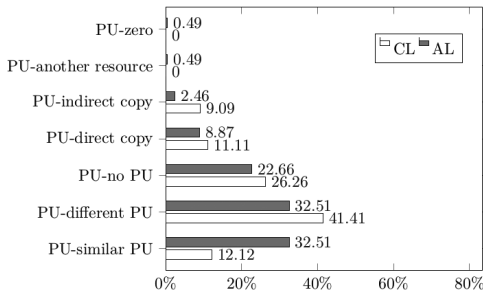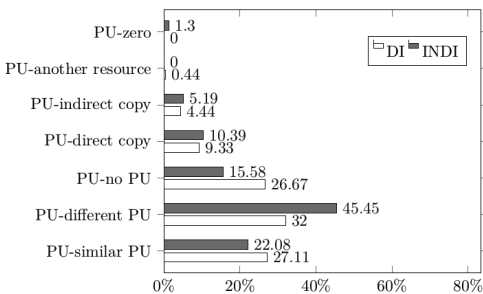
FIGURE 5

Comparison between CL and AL translations

FIGURE 6

Comparison between direct and indirect translations

The comparison between direct and indirect translations in Figure 6 shows that the percentage of PU usage (similar or different) is larger when translating indirectly: 59.11% in direct translations (27.11% and 32%) opposed to 67.53% in indirect translations (22.08% and 45.45%). The number of copies is also greater in texts translated indirectly.

5.2. Results of the TT->ST analysis

The lemma <u>esku</u> appears 3,068 times in the Basque TTs of the AleuskaPhraseo corpus. Of those, 326 were identified as Basque somatic PUs; i.e. more somatic PUs containing the lemma **hand** were found in the TTs. These are the translation options that were found: PU->PU,[26] no PU->PU, PU->indirect copy and No PU->indirect copy.

Figure 7 indicates that there is a significant number of cases in which PUs have been used in the Basque TTs where there was no phraseological stimulus in the STs, and that this happens mostly in AL texts translated indirectly. Also worth highlighting is the high percentage of indirect copies in CL texts translated in an indirect manner.

Figure 8 shows that the tendencies observed previously are maintained: more PUs were found in AL translations compared to CL and more copies were identified in CL translations. Thus, more phraseological texts are produced when translating AL texts, and interference is more frequent in CL translations.

FIGURE 7

**Distribution of the translation options across the four subcorpora**

	CL DI %	AL DI %	CL INDI %	AL INDI %	Total %
PU⇒PU	23.53	28.95	37.50	9.09	26.54
No PU⇒PU	70.59	69.30	45.83	90.91	69.19
PU⇒indirect copy	3.92	1.75	4.17	0.00	2.37
No PU⇒indirect copy	1.96	0.00	12.50	0.00	1.90
Total	100.00	100.00	100.00	100.00	100.00

FIGURE 8

Comparison between CL and AL translations
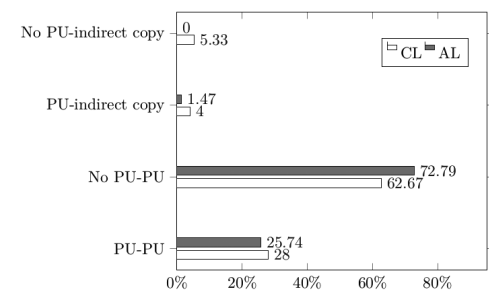
From Figure 9 it can be seen that more indirect copies were found in indirect translations as compared to direct translations (8.69% to 3.03% respectively).

FIGURE 9

Comparison between direct and indirect translations
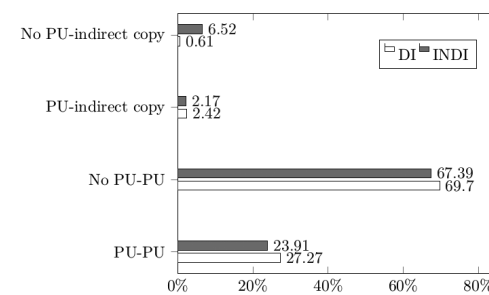
6. Discussion

Sometimes, when copying word-for-word ST PUs in translations, the literal character of a target text's solutions led to the creation of unconventional word combinations in the Basque texts. For instance, in the translation of the first example in Table 4 (ELeu), Meike and her friend leave the stage eskua eskutan or hand in hand, the Basque word combination is a word-for-word copy of the German PU **Hand in Hand**. This is very unusual, as demonstrated by the fact that no occurrence of eskua eskutan has been found in any other Basque monolingual corpus. The Basque word combination of the second example (eskutik gidatu), which is the literal counterpart of the

German PU **an der Hand führen** (to lead somebody by the hand), appears just once in a Basque reference corpus (see note 21) of around 269 million words. These examples show that, sometimes, interference from strict adherence to the ST results in unusual usage of the language in the TT.

Table 4

Example of direct copies in direct translations

Source Text	Target Text
1. „War es dir peinlich" wollte Meike wissen, nachdem der Vorhang zum letzten Mal gefallen war und sie Hand in Hand von der Bühne gingen. (ELde[27])	-Lotsatu egin zara? -jakin nahi izan zuen Meikek, oihala azken aldiz erori eta eskua eskutan jokalekutik alde egiten zutela. (ELeu[28])
2. Kinder muss man so lange an der Hand führen, bis sie den rechten Weg genau kennen!" (DAKde[29])	Haurrak eskutik gidatu behar dira harik eta bide zuzena berezteko gai diren arte. (DAKeu[30])

With regards to direct copies, another phenomenon, as exemplified in the cases of two German somatic PUs which can be found in the examples of Table 5, is also worth mentioning. The examples **für jmdn/etwas die/seine Hand/Hände ins Feuer legen** (literally meaning 'to put the hand in the fire for somebody,' and phraseologically, 'to vouch for somebody or something') and **sich (vor Vergnügen / Schadenfreude / … Kälte / … / zufrieden / …) die Hände reiben** (to rub one's hands), each appear three and twelve times respectively in the STs. Interestingly, different Basque target authors have consistently translated each one using the same target word combinations: eskua(k) sutan jarri and eskuak igurtzi, as illustrated in the examples corresponding to CL texts of Table 5. Accordingly, they have been classified as direct copies, because they are word-for-word copies of the German PUs, and are not codified in dictionaries.[31] The interesting thing about the aforementioned two word combinations is that, in Spanish, similar equivalent PUs exist (*poner la mano en el fuego* and *frotarse las manos*), and although different PUs in Basque (or many other translation options) could have been used in these cases, translators have opted to use word combinations that are not codified in dictionaries but are copies of those PUs frequently used both in German and Spanish. These cases illustrate not only that the assumed ST influences the translation, but also that, consciously or unconsciously, a third language/text may have an impact on the TTs. These examples raise the question as to whether they should be considered as direct or indirect translations.

Table 5

Examples of both German and Spanish influence in direct translations

Source Text	Target Text
1. Dafür lege ich die Hand ins Feuer. (BUSde, see note 5)	Eskua sutan jarriko nuke! (BUSeu, see note 6)
2. „Na, bei dem ganzen Kuchen!" Sie rieb sich die Hände. (SLHde[32])	"Horrenbeste tarta eta abarrekin!". Neskatoak eskuak igurtzi zituen. (SLHeu[33])

When analyzing indirect copies, it is important to distinguish translations assumed to be direct from indirect ones. When indirect copies are observed when examining the former (the assumed direct translations), it is usually difficult to

determine if it is a case of textual interference (where another text that is not the German one is the reason for the interference), of instrumental interference (where the indirect tools used by the translator are the reason for the interference), of cognitive interference (where the translator's command of several languages is the reason for the interference), or a combination of different types of interference. The example in Table 6 may help to clarify this issue.

In the example of the CL text from Table 6, the German PU **eine feste Hand brauchen** is used, which literally means 'to need a tough hand,' and the figurative meaning would be 'to act firmly towards somebody in a certain situation.' The use of the word combination <u>esku gogorra behar</u> in Basque can be classified as a word-for-word copy of the Spanish PU *(tener) mano dura* (which has the same figurative meaning as the German PU **eine feste Hand brauchen**), since there is a 1 to 1 correspondence of the lexical elements. In other words, the lexical components of the Basque word combination resemble more the Spanish PU than the German one. This is an example of the influence that the Spanish language/text can exert over the translation, even in an assumed direct translation.

Table 6

Example of an indirect copy in an assumed direct translation

Source Text	Target Text
„Solche großen Hunde <u>brauchen</u> eine <u>feste Hand</u>!" (SLHde, see note 25)	"Horrelako txakur handiek <u>esku gogorra</u> behar dute!" (SLHeu, see note 26)

The same phenomenon has been observed when extracting somatic PUs containing the lemma **esku** from Basque translations. In much the same way as above, in assumed direct translations, the type of interference (textual, cognitive, etc.) remains difficult to be determined. Irrespective of the source of interference, both of the indirect-copy examples below only make sense if we keep in mind that the Basque word combinations - **eskuko hatzekin kontatu** ('to tick off something on one's fingers') and **Jainkoaren eskutik utzitako desertua** ('a god-forsaken wasteland') - have very similar counterparts in Spanish (*contar algo con los dedos de una mano* and *dejado de la mano de Dios*). In other words, there is a greater similarity between the Spanish and Basque word combinations, than between the German and Basque ones.

Table 7

Other examples of indirect copies in assumed direct translation

Source Text	Target Text
1. Wat ick nötig habe, kann ick mir jeden Tag <u>an die Finger abzählen</u>. (BAde, see note 7)	Nik zer behar dudan, <u>eskuko hatzekin konta zezakeat egunero</u>. (BAeu, see note 8)
1. „Wo soll denn in dieser <u>gottverlassenen</u> Einöde ein Dorf sein? (THde, see note 13)	–Nola egongo da, ba, <u>Jainkoaren eskutik utzitako desertu</u> honetan herri bat? (THeu, see note 14)

As for the indirect copies in indirect translations, it is important to point out that, whilst originally they were translated from an intermediary text, for the purpose of this analysis, the relationship between the German and Basque translations were compared. Therefore, most of the time, indirect copies in indirect translations are word-for-word copies – that is to say, direct copies – of the intermediary text's PUs.

In the example of Table 8 for instance, the Basque word combination <u>eskuak burura eraman</u> (literally 'to take the hands to the head') is a word-for-word translation of the Spanish PU *llevarse las manos a la cabeza* (and not of the German PU **die Hände über dem Kopf zusammenschlagen**), but it was classified as indirect copy because the original German text and the Basque TT were the texts that were compared.

TABLE 8

Example of an indirect copy in an indirect translation

Source Text	Intermediary Text	Target Text
Jens schlug die Hände über dem Kopf zusammen und Bernhard hielt ihn zurück. (BLAde[34])	Jens se llevó las manos a la cabeza y Bernhard intentó calmarlo. (BLAes[35])	Jensek eskuak burura eraman zituen eta Bernhard saiatu zen hau lasaitzen. (BLAeu[36])

As stated before, in the case of indirect translations, indirect copies are almost always influenced by the Spanish intermediary texts; that is to say, the copies in the Basque TTs are word-for-word copies of Spanish PUs. There is, however, one exception that needs to be mentioned at this point. In Table 9, the German PU **mit vollen Händen Geld ausgeben**, which means 'to spend one's money left, right and center' – has been translated with a similar PU into Spanish (*gastar dinero a manos llenas*). What is found in the Basque translation, though, is a word-for-word copy of another Spanish PU, namely *echar algo por la ventana*. According to the Spanish dictionary RAE: *Desperdiciarlo o malgastarlo* ('to waste something,' money in this case). Moreover, the phraseological character of the Basque word combination (which is not very common in Basque) is enhanced through the metalinguistic explanation <u>esaera den bezala</u> ('as the saying goes'). Although BMeu has been classified as an assumed indirect translation, if it is actually a direct translation, then the example in Table 9 would represent a case of cognitive interference. However, if it was instead translated indirectly through the Spanish version, the type of interference found here would become more complex to categorize.

TABLE 9

Example of a more complex type of interference

Source Text	Intermediary Text	Target Text
Ich fing an, Geld auszugeben – <u>mit vollen Händen</u>, wie man sagt. (BMde[37])	Comencé a <u>gastar dinero a manos llenas</u>, como suele decirse. (BMes[38])	Hasi nintzen <u>dirua leihotik botatzen</u>, esaera den bezala. (BMeu[39])

When conducting queries based on the Basque translations, cases with PUs that are not very typical in Basque (in the sense that they have a Spanish "touch") were found in indirect translations. These findings are in contradiction with the hypothesis outlined above concerning authors' intentions, because although these are indirect translations, there is no attempt on the translator's part to conceal interference from the Spanish language in order to hide the indirect character of the translations. For the purpose of finding an explanation for these solutions, another factor needs to be taken into account. The mentioned examples were found in texts that were published in the late 80s and early 90s. In that regard, it could tentatively be argued that the tendency, at least in indirect translations, to create more or less acceptable (in the sense of Toury) translations has changed over time. At first, when the need to

create Basque texts was urgent, creating texts that would adhere to the norms of the target language may not have been the priority. The absence of direct copies in indirect translations from the early 20th century of the AleuskaPhraseo corpus may indicate that over time an awareness of creating more authentic Basque texts may has grown. The next figure shows three of these examples.

Table 10

Example of the influence of the Spanish versions on the Basque TTs translated indirectly

Source Text	Intermediary Text	Target Text
1. Verschwindest du freiwillig, oder muß ich <u>nachhelfen</u>?" (DKGde[40])	¿te esfumas por las buenas o he de <u>echarte una mano</u>? (DKGes[41])	onean aldegingo duzu ala <u>esku bat bota</u> behar dizut? (DKGeu[42])
2. „Man hat Ihnen also <u>geholfen</u>?" fragte der Uhu schließlich. "Man hat, wie Sie sehen, sagte das kleine Gespenst. (DKGde, see note 35)	-¿Le han ayudado a usted? -preguntó el búho por último. -Me han <u>echado una mano</u> como puede usted ver -dijo el fantasmita. (DKGes, see note 36)	–Lagundu al dizue? -galdetu zuen hontzak azkenean. –<u>Esku bat bota</u> didate, ikus dezakezuenez -esan zion fantasmatxoak. (DKGeu, see note 37)
3. Du bist wohl <u>von allen guten Vampiren verlassen</u>!" (DKVZde[43])	¡Tú <u>estás dejado de la mano de los vampiros buenos</u>! (DKVZes[44])	Zu banpiro onen <u>eskuetatik utzita zaude</u>! (DKVZeu[45])

In the first two examples, found in the same text, the verbs **nachhelfen** and **helfen** have been used in the German originals. In the Spanish version, both have been translated with the Spanish PU *echar una mano a alguien* (literally, 'to throw somebody a hand'), and in the Basque translation a word-for-word copy of the Spanish PU has been found (<u>esku bat bota</u>). Thus, no sign of avoiding or hiding the indirect character of the translation can be seen here. The third example seems to be a modification of the German PU **von allen guten Geistern verlassen sein** ("to have taken leave of one's senses"). In the Spanish translation there is also a modified PU (the PU *dejado de la mano de Dios*, meaning 'god-forsaken,' has been modified), but it is a different PU, in the sense that the German and Spanish PUs have different meanings. The modified Spanish PU is the one that has been directly translated into Basque. As a result, in this case the Spanish PU's influence is reflected in the Basque translation, meaning that the German original and the Basque translation differ at the semantic level.

Other examples, however, corroborate the proposed hypothesis, and in terms of the translation process of indirect translations, some examples show that there is a tendency to deviate from the Spanish intermediary texts, and to use PUs that are very typical of the Basque language. For example, of the 22 cases in which a variant of "to give somebody one's hand" was used in the German STs, similar PUs were employed in the Spanish translations, but when translating the Spanish PUs into Basque, a different PU was always used. Moreover, in 17 cases (out of those 22) the word <u>bosteko</u> ('five,' meaning 'hand') was employed within the Basque PUs, thereby enhancing the typical character of the Basque PUs. The translator is aware of the fact that s/he is translating indirectly, and in an attempt to compensate for this indirectness, s/he may employ more typical or authentic language combinations in the TTs.

7. Conclusion

Within Toury's (1995/2012) DTS, the use of corpora is of great relevance for the explanation, description and prediction of translational behavior in terms of translational norms and laws. The creation of a parallel and multilingual corpus – together with the tool needed to create it, that is, the *TAligner* program – was, in this case, essential for the analysis of interference in German-into-Basque translations. Even more interesting would have been to create a comparable corpus of Basque original (non-translated) texts, but that was an impossible task due to time constraints. This could, however, be regarded as a future step. For the time being, other Basque corpora, available on the web, have filled the gap.

The very fact that a multilingual corpus was created at all also needs to be emphasized. Although the aim of the study was to analyze German-into-Basque translations, the compilation of a bilingual corpus would have made no sense in the present case. Updating the Aleuska catalog made it clear that indirect translations needed to be taken into account, and that was why, when compiling the corpus, intermediary texts were added to the corpus when, according to publishing houses or translators, texts other than the German had been used as STs. The importance of considering not only linguistic, but extra-linguistic factors, over the course of the whole research study, was given great importance.

Importantly also, when delimiting the object of study, namely PUs, several factors were considered. Not only do they represent interesting language units to be analyzed from a translational perspective, but they can also be viewed as units especially vulnerable to interference from other languages. In the present study, it has not only been shown that calquing PUs from other languages plays an important role, but that this direct transfer of PUs can also serve as a tool to measure interference. The results obtained from the translation options *direct copy* and *indirect copy* were also taken into account when analyzing interference, respecting the translation laws proposed by Toury.

With regards to direct copies, relatively high percentages were found when examining the translation of somatic PUs in the German-Basque language combination, especially when compared to the results of other studies (Marco 2013). Sometimes, unusual word combinations were observed as a result of the literal character of the translation option, and at other times the fact that similar PUs existed in both major languages, German and Spanish, encouraged the direct borrowing of these PUs. It can thus be concluded that when translating from German into Basque, textual interference is, to a certain extent, accepted. Since the percentages are higher when translating texts for children, it can be argued that fingerprints of the ST are more accepted when translating CL texts. Thus, the source language's status, the sociolinguistic situation of the target language and the text type may all influence the acceptance of interference.

Another variable suggested by Toury is that of cognition, which is strongly connected to the sociolinguistic environment in which the language is situated. In the present study, the impact of this variable is reflected in the translation option *indirect copy*. It seems that the text type influences the acceptance of indirect copies, since higher percentages were found in CL translations. What indirect copies show is that when performing assumedly direct translations from German into Basque, another

text or language (which is not the German ST) can influence the translation process. It is possible that this type of interference could be textual, because translators may have had access to other translations when translating. However, the type of interference could also be instrumental, in the sense that the indirect tools used by the translators may have influenced the translation. According to the third kind of interference, i.e. cognitive interference, the translator's command of another language could also have an impact on the TT. In any case, it is often difficult to determine the exact type of interference present. In conclusion, it can be said that when translating from a prestigious language A into a minority language B, if language B coexists unequally with a dominant language C, then, according to different variables, different types of interference from language C into language B can occur to different degrees.

Where indirect copies were found in indirect as opposed to direct translations, they mostly represented cases of textual interference. As for attitudes towards interference, it can be seen that in some cases, due to the historical context in which they were created, interference from a Spanish text was not hidden, but at other times there was a tendency to deviate from the Spanish intermediary version and create more typical Basque texts.

It is true that corpus data will not allow us to see into the translator's mind and this makes it difficult to gain an insight into the translation process. But very interesting results were obtained regarding the identification of different types of interference thanks to data obtained from the corpus. In the near future, not only the results, but also the corpus itself will be very useful, for instance, in a translation-teaching environment.

NOTES

1. For details on the catalog, see Section 4.
2. As explained by Amorrortu (2003: 46), while "Ferguson (1959) described the concept of diglossia in the context of internal variation within a language (…) scholars in minority language situations adopted the concept of diglossia to refer to the imbalance in the functional allocation of two languages."
3. These concepts were also used in publications written by members of the TRALIMA-ITZULIK research group from the University of the Basque Country, to which I belong (see for instance Zubillaga 2013; Sanz-Villar 2015; Barambones, Manterola *et al.* 2015).
4. Martínez, Elena, ed. (1996): *Euskara-alemana hiztegia. Deutsch-Baskisch Wörterbuch* [Basque-German dictionary]. Donostia: Elkar.
5. Examples of this translation option can be consulted in Table 10.
6. Preussler, Otfried (1958): *Bei uns in Schilda* [Here in Schilda] Stuttgart: Thienemann.
7. Preussler, Otfried (1958/1987): *Markako eroak* [The heroes of Marka]. (Translated by Pello Zabaleta) Madrid: SM.
8. Döblin, Alfred (1929): *Berlin Alexanderplatz* [Berlin's Alexanderplatz]. Berlin: Fischer.
9. Döblin, Alfred (1929/2000): *Berlin Alexanderplatz* [Berlin's Alexanderplatz]. (Translated by Antton Garikano) Euba: Ibaizabal.
10. Hesse, Hermann (1927): *Der Steppenwolf* [The wolf of the steppes]. Berlin: Fischer.
11. Hesse, Hermann (1927/1986): *Estepako otsoa* [Wolf of the steppes]. (Translated by Pello Zabaleta) Donostia: Elkar.
12. Werfel, Franz (1928): *Der Abituriententag* [Graduation day]. Wien: P. Zsolnay.

13. Werfel, Franz (1928/2008): *Batxilerren eguna* [Bachelor's day]. (Translated by Juan Luis Agirre) Iruñea: Igela.
14. Funke, Cornelia (2003): *Tintenherz* [Inkheart]. Hamburg: Dressler.
15. Funke, Cornelia (2003/2009): *Tinta-bihotz* [Heart of ink]. (Translated by Naroa Zubillaga) Donostia: Elkar.
16. Professor Ibon Uribarri was the one who started compiling the catalog in 2003.
17. It was first created within the TRACE research project: <http://www.ehu.eus/trace/inicio.php>.
18. Due to the lack of research studies regarding poetry, this last option has not been further developed yet.
19. Anthony, Laurence (2014): *AntConc* [computer software]. Version 3.4.3. Tokyo: Waseda University.
20. Apart from somatic PUs, binomials were extracted from the corpus and analyzed (Sanz-Villar 2016).
21. Schemann, Hans (2011): *Deutsche Idiomatik: Wörterbuch der deutschen Redewendungen im Kontext* [German Idiomatics: Dictionary of German idioms in context]. Berlin/Boston: Walter de Gruyter.
22. *Duden: Redewendungen und sprichwörtliche Redensarten. Wörterbuch der deutschen Idiomatik* [Duden: phrases and proverbial sayings. Dictionary of German Idiomatics] (1992): Mannheim/Leipzig/Zürich: Dudenverlag.
23. The Elhuyar dictionary can be consulted here: <http://hiztegiak.elhuyar.eus/>.
24. The DeReKo corpus can be consulted here: <https://cosmas2.ids-mannheim.de/cosmas2-web/>.
25. The Euskal Testuen Corpusa (ETC, Basque Texts Corpus) can be consulted here: <http://www.ehu.eus/etc/>.
26. Since the translation option PU->similar PU has already been analyzed previously, it has not been included here.
27. Arold, Marliese (1996): *Einfach nur Liebe* [Just love]. Bindlach: Loewe Verlag.
28. Arold, Marliese (1996/2004): *Maitasuna besterik ez* [Only love]. (Translated by Pello Zabaleta) Donostia: Elkar.
29. Nöstlinger, Christine (1982): *Das Austauschkind* [Exchange child]. Weinheim: Beltz & Gelberg.
30. Nöstlinger, Christine (1982/1991): *Ingeles bat etxean* [An English child at home]. (Translated by Xabier Mendiguren) Lizarra: Elkar.
31. As for their occurrence in the Basque reference corpus ETC (last consulted in April 2014), eskua sutan jarri is used 45 times: mostly in press texts (22 times) and in the subcorpus that consists of the dialogues of a popular Basque television series (17), and less in literary texts (5 times in translated texts and once in non-translated texts). The second word combination (eskuak igurtzi) is more frequent, as 56 occurrences were found, and it is used more frequently in press texts (29 occurrences) than in literary texts (27 occurrences), although the margin is narrower. It should be borne in mind here that since the subcorpus consisting of press texts is larger in ETC, it is to be expected that more occurrences should be found there.
32. Sommer-Bodenburg, Angela (1991): *Schokolowski. Lustig ist das Hundeleben* [Schokolowski. Funny is the dog life]. München: Bertelsmann.
33. Sommer-Bodenburg, Angela (1996): *Txokoloski. Dibertigarria da txakurren bizitza* [Txokoloski. The life of dogs is fun]. (Translated by Edurne Azkue) Bilbao: Desclée de Brouwer.
34. Härtling, Peter (1979): *Ben liebt Anna* [Ben loves Anna]. Weinheim/Basel: Beltz & Gelberg.
35. Härtling, Peter (1979/1982): *Ben quiere a Ana* [Ben loves Anna]. (Translated by Víctor Canicio) Madrid: Alfaguara.
36. Härtling, Peter (1979/1990): *Benek Ana maite du* [Ben loves Anna]. (Translated by David Urbistondo) Baiona: Elkar.
37. Roth, Joseph (1936): *Beichte eines Mörders* [Confession of a murderer]. Amsterdam: Allert de Lange.
38. Roth, Joseph (1936/1992): *Confesión de un asesino*. (Translated by Juan José del Solar) Barcelona: Anagrama.
39. Roth, Joseph (1936/2003): *Hiltzaile baten aitormena* [Confession of a murderer]. (Translated by Matías Múgica) Iruñea: Igela.
40. Preussler, Otfried (1966): *Das kleine Gespenst* [The little ghost]. Stuttgart: Thienemann.
41. Preussler, Otfried (1966/1969): *El pequeño fantasma*. (Translated by Carmen Maluenda) Barcelona: Bruguera.
42. Preussler, Otfried (1966/1984): *Fantasmatxoa* [The little ghost]. (Translated by Xabier Gartzia) Donostia: Elkar.

43. Sommer-Bodenburg, Angela (1980): *Der kleine Vampir zieht um* [The little vampire is moving]. Reinbek bei Hamburg: Rowohlt.

44. Sommer-Bodenburg, Angela (1980/1987): *El pequeño vampiro se cambia de casa*. (Translated by José Miguel Rodríguez) Madrid: Alfaguara.

45. Sommer-Bodenburg, Angela (1980/1990): *Banpiro txikia etxez aldatzen da* [The little vampire goes home]. (Translated by Juan Legarreta) Bilbao: Desclée de Brouwer.

REFERENCES

Alberdi, Xabier (2010): A typology of calques. The calquing mechanism in contemporary Basque. *ELUA*. 24:13-35.

Amorrortu, Estibaliz (2003): *Basque Sociolinguistics: Language, Society and Culture*. Reno: Center for Basque Studies.

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.

Baldwin, Timothy and Kim, Su Nam (2010): Multiword expressions. *In*: Nitin Indurkhya and Fred J. Damerau, eds. *Handbook of Natural Language Processing*. Boca Raton: CRC Press, 267-292.

Barambones, Josu, Manterola, Elizabete, Sanz-villar, Zuriñe, *et al.* (2015): Itzulpen ikasketak eta euskara. Zenbait ekarpen Itzulpen Ikasketa Deskribatzaileei, eta haratago [Translation Studies and Basque. Contributions to Descriptive Translation Studies, and beyond]. *In*: *In*: José Jorge Amigo Extremera, ed. *Traducimos desde el Sur. Actas del VI Congreso Internacional de la Asociación Ibérica de Estudios de Traducción e Interpretación*. Las Palmas de Gran Canaria: ULPGC.

Bracho, Llum (2013): El corpus COVALT: un observatori de fraseologia traduïda [The COVALT corpus: an observatory of translated phraseology]. Aachen: Shaker.

Burger, Harald (2007): *Phraseologie. Eine Einführung am Beispiel des Deutschen* [Phraseology. An introduction using the example of German]. Berlin: Erich Schmidt.

Burger, Harald, Dobrovol'skij, Dimitrij, Kühn, Peter *et al.* (2007): Phraseologie: Objektbereich, Terminologie und Forschungsschwerpunkte [Phraseology: subject area, terminology and research topics]. *In*: Harald Burger, Dimitrij Dobrovol'skij, Peter Kühn, *et al.*, eds. *Phraseologie: Ein internationales Handbuch der zeitgenössischen Forschung / Phraseology: An International Handbook of Contemporary Research*. Berlin/New York: Walter de Gruyter, 1-10.

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

Delabastita, Dirk (1996): Introduction. *The Translator*. 2(2):127-139.

Fiedler, Sabine (2007): *English Phraseology: A Coursebook*. Tübingen: Narr.

Genette, Gérard (1987): *Seuil*. Paris: Éditions du Seuil.

Kenny, Dorothy (2005): Parallel corpora and translation studies: old questions, new perspectives? Reporting *that* in Gepcolt: a case study. *In*: Geoff Barnbrook, Pernilla Danielsson, and Michaela Mahlberg, eds. *Meaningful Texts: The Extraction of Semantic Information from Monolingual and Multilingual Corpora*. London/New York: Continuum, 154-165.

Lefer, Marie-Aude and Vogeleer, Svetlana (2013): Interference and Normalization in genre-controlled multilingual corpora: Introduction. *In*: Marie-Aude Lefer and Svetlana Vogeleer, eds. *Interference and Normalization in Genre-controlled Multilingual Corpora*. Amsterdam: John Benjamins, 1-21.

Marco, Josep (2009): Normalisation and the Translation of Phraseology in the COVALT corpus. *Meta*. 54(4):842-856.

Marco, Josep (2013): La traducció de les unitats fraseològiques de base somàtica en el subcorpus anglès-català [The translation of somatic phraseological units in the English-Catalan subcorpus]. *In*: Llum Bracho, ed. *El corpus COVALT: un observatori de fraseologia traduïda* [The COVALT corpus: an observatory of translated phraseology]. Aachen: Shaker, 163-215.

MARCO, Josep and VAN LAWICK, Heike (2015): Enhancing translator trainee's awareness of source text interference through use of comparable corpora. *In*: Agnieszka LEŃKO-SZYMAŃSKA and Alex BOULTON, eds. *Multiple Affordances of Language Corpora for Data-driven Learning*. Amsterdam/Philadelphia: John Benjamins, 255-244.

MAURANEN, Anna (2004): Corpora, universals and interference. *In*: Anna MAURANEN and Pekka KUJAMÄKI, eds. *Translation Universals: Do They Exist?* Amsterdam/Philadelphia: John Benjamins, 65-82.

MELLADO, Carmen (2004): *Fraseologismos somáticos del alemán*. Frankfurt am Main: Peter Lang.

MILIZIA, Denise and SPINZI, Cinzia (2008): The "terroridiom" principle between spoken and written discourse. *International Journal of Corpus Linguistics*. 13:322-350.

MOON, Rosamund (1998): *Fixed Expressions and Idioms in English. A corpus-based approach*. Oxford: Clarendon Press.

RICHTER-VAPAATALO, Ulrike (2007): *Da hatte das Pferd die Nüstern voll. Gebrauch und Funktion von Phraseologie im Kinderbuch. Untersuchungen zu Erich Kästner und anderen Autoren* [*Da hatte das Pferd die Nüstern voll*. The use and function of phraseology in children literature. Analyzing Erich Kästner and other authors]. Frankfurt am Main: Peter Lang.

SANZ-VILLAR, Zuriñe (2015): *Unitate fraseologikoen itzulpena: alemana-euskara. Literatur testuen corpusean oinarritutako analisia* [The translation of phraseological units: German-Basque. A corpus-based analysis]. Doctoral dissertation, unpublished. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea.

SANZ-VILLAR, Zuriñe (2016): German-into-Basque/Spanish translation analysis of binomials in a parallel and multilingual corpus. *In*: Gloria CORPAS, ed. *Computerised and Corpus-based Approaches to Phraseology: Monolingual and Multilingual Perspectives*. Geneva: Editions Tradulex, 80-92.

SZCZĘK, Joanna (2004): Einwortphraseologismen und ihr Verhältnis zur Phraseologie (am Beispiel des Deutschen und des Polnischen) [One-word phraseological units and their relationship to phraseology in German and Polish]. *In*: Csaba FÖLDES and Jan WIRRER, eds. *Phraseologismen als Gegenstand sprach- und kulturwissenschaftlicher Forschung. Akten der Europäischen Gesellschaft für Phraseologie (EUROPHRAS) und des Westfälischen Arbeitskreises „Phraseologie/Parömiologie"* [Phraseologisms as an Object of Linguistic and Cultural Studies. Proceedings of the European Society of Phraseology (EUROPHRAS) and of the research group Westfälische Arbeitskreis "Phraseologie / Parömiologie"]. Baltmannsweiler: Schneider, 75-85.

TOURY, Gideon (1995/2012): *Descriptive Translation Studies and Beyond*. Amsterdam: John Benjamins.

URIZAR, Ruben (2012): *Euskal lokuzioen tratamendu konputazionala* [Computational processing of Basque idioms]. Doctoral dissertation, unpublished. San Sebastián/Donostia: Universidad del País Vasco/Euskal Herriko Unibertsitatea.

VAN LAWICK, Heike (2013): Apunts sobre fraseologia i per a l'anàlisi traductològica [Notes on phraseology and translation analysis]. *In*: Llum BRACHO, ed. *El corpus COVALT: un observatori de fraseologia traduïda* [The COVALT corpus: an observatory of translated phraseology]. Aachen: Shaker, 129-161.

VOLANSKY, Vered (2012): *On the Features of Translationese*. Master's thesis, unpublished. Haifa: University of Haifa.

ZUBILLAGA, Naroa (2013): *Alemanetik euskaratutako haur- eta gazte-literatura: zuzeneko nahiz zeharkako itzulpenen azterketa corpus baten bidez* [Children's literature translated from German into Basque: a corpus-based analysis of direct and indirect translations]. Doctoral dissertation, unpublished. Vitoria-Gasteiz: Universidad del País Vasco/Euskal Herriko Unibertsitatea.