

PROBABILISTIC GRAMMAR IN TRANSLATION

Silvana Maria de JESUS
Adriana Silvina PAGANO
(Universidade Federal de Minas Gerais/CAPES)

ABSTRACT: *Adopting a systemic-functional approach to translation studies, this paper focus on probabilistic grammar (Halliday, 1991, Matthiessen, 2001), developing the concept of conditioned probabilities (Catford, 1965) through the investigation of translation equivalence in a fictional multilingual parallel corpus.*

KEY-WORDS: *Translation studies, probabilistic grammar, translation equivalence, translation techniques, multilingual parallel corpus.*

1. Introduction

This article reports on an ongoing Ph.D. dissertation within a research project – CORDIAL (Corpus of Discourse for the Analysis of Language and Literature), developed by NET (Núcleo de Estudos da Tradução) at Faculdade de Letras, Universidade Federal de Minas Gerais - which adopts a systemic-functional approach to translation studies drawing on evidence obtained through corpora, concordancing and statistical software.

More specifically, it explores insights from the interface between systemic functional grammar (SFG) and translation studies that has been pursued since the 1960s. Such interface can be said to begin with Catford's (1965) development of *A linguistic theory of translation* based on Halliday's theory of grammar (1964). Halliday himself goes into some translation issues, placing translation within comparative descriptive linguistics (1964:112) and in his essay "Towards a theory of good translation" (2001) he discusses the concept of equivalence based on three parameters: stratification, metafunction and rank.

Drawing on Halliday's (1991) proposal of probabilistic grammar to find out specific text characteristics within a given register and Matthiessen's (2001) proposal of mapping out the resources of source and target language in order to contextualize the environments of translation, this work develops Catford's concept of *conditioned probabilities* (1965:31) in equivalent relations between original and translated items. The corpus used is composed by three novels – an original, *Point counter point*

(1928) by Aldous Huxley, its Brazilian-Portuguese translation, *Contraponto* (1934), by Érico Veríssimo, and its Spanish translation, *Contrapunto* (1940) by Lino Nóvas Castro, making up a multilingual parallel corpus within the fictional register. Using SFG for qualitative analysis the data is also examined with concordancing and statistical software for quantitative analysis based on a probabilistic approach.

The analysis starts with the lexical item SAYⁱ and examines it in a broader context, mapping out the functional roles it plays and relating them probabilistically with SAY's prototypical equivalent DIZER and DECIR in Portuguese and Spanish, respectively. The roles its usual counterparts – DIZER and DECIR – play in their own systems are studied as well, attempting an interpretation of the contexts in which equivalence is possible and more frequent between them. A descriptive and contrastive analysis of the functions of these lexical items is carried out, exploring the relations of equivalence among the three languages.

This paper is organised in three sections. Section one brings the theoretical basis for a probabilistic grammatical analysis in translation, mainly based on Catford (1965), Matthiessen (2001), Halliday (1964, 1991, 2001, 2005), Halliday and Matthiessen (2004), Teich (2001, 2003) and Nesbitt & Plum (1988). Section two develops data analysis from a sample of the corpus; and then, section three points out some final remarks.

2. A theoretical basis for probabilistic grammar in translation

Translation equivalence has always been a main issue in translation studies. From a probabilistic perspective based on systemic functional theory, Catford (1965:27) makes a distinction between translation equivalence as an empirical phenomenon and the study of the conditions of translation equivalence. Both approaches can be used to investigate equivalence in a fictional corpus; specific items of the source text can be analysed in terms of the way they are retextualized in the translated text. Investigating them in wider contexts (the clause as a semantic unit) it is possible to describe its *conditioned probabilities* of equivalence. These aspects are related to what Halliday (2005:60) calls *instantiation* and *realization*.

Catford does not fully develop the concept of conditioned probability, relating it basically to the idea of collocations and Machine Translation. Adding Matthiessen's (2001) perspective of mapping out the

environments of translation from the widest to the narrowest contexts it is possible to explore this concept from a probabilistic perspective.

In “The environments of translation”, Matthiessen (2001) says that linguistics and translation studies are closely related areas; although the translation phenomenon is not reduced to its linguistic aspect, “the process of translation works in terms of the resources of the systems involved in the translation” (p.72), and clearly, the linguistic system is the main one; lexicogrammar, semantics and context are strata in systemic theory which instantiate texts from general patterns in a given system.

From this point of view, Matthiessen takes SFG to discuss some of the basic issues in translation, such as *translation equivalence*, *translation shift* and the concept of translation itself. Based on an analysis of the *Process to translate* in the LOB corpus, the author states that translation can be seen in two ways: i) in a “material model” or “transfer” in which translation relates to “transforming a semiotic entity from one language into another”; or ii) in a “relational model” or “signification” in which translation is seen as “an assignment of meaning – as mapping from one language to another” (p.47).

Taking the second view, of translation as an assignment of meaning, Matthiessen suggests a model where translation has to be contextualised “from the most global environments of language (in context) to the most local ones” (p.48), considering that it is necessary “to map out” the resources of both source and target language in order to analyse some basic concepts in translation. In this approach, Matthiessen defines translation as “a process of construing: experience construed as meaning in one system is (re)construed as meaning in another” (p.51).

Matthiessen (2001:84), for instance, explains how existential clauses in English and German have different “systemic values”. This is one of the possibilities of analysis, i.e., investigating how existential clauses are construed lexicogrammatically in these languages; from this perspective, when the analysis takes functional categories, the interpretation of the clause is taken from above. However, it is also possible to start an investigation from below, that is, from a narrower context – that of lexicogrammar – and move upwards, multi-contextualising the analysis by including the semantic level. As the author states, the lower levels are also important because they are resources for “meaning-making power” (p.81).

Although comparative and contrastive linguistics are not the focus of translation studies, “translators do of course use a certain type of comparative description” (Matthiessen, 2001:88), and contrastive linguistics, especially from parallel corpora, can provide interesting insights into translation and into translation equivalence.

From the perspective of contrastive-linguistic description, in “Towards a model for the description of cross-linguistic divergence and commonality in translation”, Teich (2001) also applies SFG to “theoretically contextualise some of the basic notions of translation theory” (p.191). Interested in describing translation from a contrastive-linguistic approach, Teich intends to fill the gap between theoretical and applied translation studies. Arguing that contrastive-linguistic knowledge is part of the “skill set” necessary to solve translation problems, Teich assumes that uncovering and describing contrastive-linguistic knowledge is “a major component of theorising translation” (p.193). Following the hallidayan model, Teich states that it is important to have “knowledge about commonalities and divergences in the instantiation of language systems in text and according to context”. The author suggests a model that will lead to contextualise “translational concepts, such as translation strategy, translation procedures, translation type, etc”. Teich is interested in investigating the options available to translators and what motivates their choices.

Teich (2003) further develops the model suggested above in *Cross-linguistic variation in system and text: a methodology for the investigation of translations and comparable texts* where she compares English and German parallel and comparable corpora, describing some properties of English and German systems and texts. Teich (p.29) criticizes Baker’s (1995) and Biber’s (1995) works on the basis of their lack of “an underlying theory and model of language” and establishes her model on two grounds: a socio-semiotic theory of language, i.e. the hallidayan model, and a corpus-based methodology for quantitative analysis.

In “The linguistic sciences and language teaching”, Halliday (1964:112) explains the importance of comparative/contrastive linguistics to describe and compare the patterns of two or more languages. He says:

“(…) language events are not random: they follow recognizable patterns. The patterns can be thought of as predictions which first distinguish between what is possible and what is

impossible, and then, within what is possible, show what is more likely and what is less likely.” Halliday (1964:137)

He also recognizes that “the line between ‘impossible’ and ‘very unlikely’ is often difficult to draw.” (idem). Thus, as Halliday says, language events follow “recognizable patterns” and, according to him, while quantitative lexical patterns tend to be recognized and accepted, quantitative grammar patterns do not. However, in systemic theory, grammar and lexis are not two separated phenomena, but “they are the same phenomenon looked at from different ends” (Halliday, 2005:60).

Halliday (2005:61) uses a probabilistic interpretation to explain how patterns of a text relate to patterns in the linguistic system. This is what he calls the “Hamlet factor”, i.e., if we have a “highly valued instance”, this may cause a change in the system. The issue at stake here is that while these patterns may or may not cause changes in the linguistic system, they do model translation equivalence. Given a pattern of the original text, the translator has to choose either to keep or change it, but before this choice it is necessary for the translator to realize that pattern.

As Halliday (2005:59) points out, some people raise objections such as that the patterns of a text cannot be interpreted as patterns of the system, to what he answers that “it requires a lot of observations to approximate to a quantitative profile of the grammar of a language”. In the same way, in order to present useful data to translation studies large-scale corpus of text must be investigated.

Here are some important principles of probabilistic interpretation. Halliday (2005:48) considers that grammatical patterns seem to be “bimodal” and tend to be “equiprobable” (0.5:0.5) or “skew” (0.9:0.1), that is, a grammatical system can present two options that are equally probable to occur – e.g. *number*, for which one can choose between singular and plural (0.5:0.5 equiprobable) – in this case there is no marked choice; or a grammatical system may have one option that is more frequent – e.g. *polarity*, for which one can choose between positive or negative, though negative is the marked choice, that is, positive is more frequent (0.9:0.1 skew). As he explains:

“This seemed to account for our sense that some systems have an ‘unmarked term’, whereas others have not. An unmarked term is a default condition: that which is selected unless there is good reason for selecting some other term. It is not *defined* by frequency, but it is likely to correspond to the more probable

term in a system whose probabilities are skew". Halliday (1991:35)

An example of this kind of approach is given by Nesbitt and Plum (1988) in a study of taxis and logical relations in the complex clauses of English. Nesbitt and Plum (1988:8) propose that grammatical patterns should be explained qualitatively – a choice between option A or B (*either this or that*) - associated with a quantitative analysis - that establishes if A occurs more or less frequently than B (*more like this / less like that*).

This paper attempts to develop a probabilistic interpretation to the qualitative and quantitative analysis of translation equivalence, aiming to investigate the conditioned probabilities of equivalence between the lexical item SAY and its prototypical equivalents in Portuguese and Spanish, DIZER and DECIR, respectively. Considering, as Halliday (1991:31) puts it, that lexicogrammar is a “unified phenomenon, a single level of ‘wording’, of which lexis is the ‘most delicate’ resolution”, the analysis of SAY in the semantic and grammatical environments where it occurs multi-contextualise translation equivalence and brings contrastive-linguistic knowledge that is necessary to translators. The next section will explain this point with an illustration based on corpus analysis.

3. Descriptive and probabilistic analysis using corpora

The first four chapters of each novel in the corpus were considered, i.e. *Point counter point* (the original in British English), *Contraponto* (the translation to Brazilian Portuguese) and *Contrapunto* (the translation to Spanish). Using *WordSmith Tools* we find some of the statistical data for this sample of the corpus.

	ST	TT1(Portuguese)	TT2(Spanish)
Types	3.992	4.961	4.954
Tokens	19.805	21.014	20.811
Occurrences of the lexical item	98 (0.49%) SAY	88 (0.41%) DIZER	116 (0.55%) DECIR

Table 1 – Statistical data of the corpus sample

Then, each occurrence of SAY, DIZER and DECIR was taken and analysed. In this sample, there are 98 occurrences of SAY, 88 DIZER and 116 DECIR. Each concordance line was analysed in order to

describe the role of the lexical item in the context of the clause and compared with the parallel corpus in order to investigate the relations of equivalence.

In order to understand the environments where these lexical items can occur, one should keep in mind that, according to the systemic theory a semantic unit can be realized in different ways, i.e., by distinct grammatical units. In an example given by Halliday & Matthiessen (2004:592), the semantic domain of modality “is construed in more than one place in the grammar; for example, it is construed by clauses such as *I suppose* and *it is possible*, by verbal groups with finite modal operators such as *may* and by adverbial groups with modal adverbs such as *perhaps*”. As the authors explain, from a grammatical point of view they are all different, but from a semantic one they are all agnates (p. 597). The concept of agnation is important here because “explaining something consists not of stating how it is structured but in showing how it is related to other things: its pattern of systemic relationships, or **agnateness**” (p.31).

It is also possible to have “realignments” or “re-mappings” in the “realizational relationship between semantic units and grammatical ones”. That is,

“a participant is realized by a nominal group, a process by a verbal group and a circumstance by an adverbial group or a prepositional phrase. But once these couplings between the two strata of the content plane have been established, ‘cross-couplings’ become theoretically possible”. (Halliday & Matthiessen, 2004:592)

These two possibilities are metaphorical or uncongruent modes of expression, contrasting with the common congruent mode of expression. From a probabilistic interpretation, we can consider that a semantic unit has its typical, i.e., most frequent, grammatical unit of realization, but language can be expanded by doing these “realignments”, which create agnate forms of expression of a semantic unit.

As it was said, lexis and grammar are two poles of the same phenomenon, so it is not difficult to perceive how the lexical items here investigated fit into these grammatical units: SAY, DIZER and DECIR are typically part of verbal groups (grammatical unit) that usually realize verbal Process (semantic unit) in a verbal projecting clause (semantic contextual environment), but they can also realize other semantic units, and, the point in this research is that the functional role these items play

in different semantic units (e.g. Process, Modality) conditions the probabilities of their translation equivalence. Because the most frequent role these items play is that of a verbal Process in a projecting clause, the logical aspect is also considered, taking into account if the clause where the lexical item appears is projecting or not.

Let us now describe each environment where these items can occur and look at their equivalent relations. The methodology is based on corpus linguistics and statistical data was processed using *WordSmith Tools* and *SPSS* (Statistical Package for Social Science).

According to the data analysed, SAY can be part of clauses realizing experiential or interpersonal meanings. It can be part of an element in the transitivity system or part of an interpersonal grammatical metaphor. SAY is mainly a Process realizing experiential meaning (98%), with few occurrences as an interpersonal grammatical metaphor (2%). It is not easy to draw a clear distinction between congruent and incongruent meaning, and we could analyse grammatical metaphors twice, “first as a separate clause and then as a Mood Adjunct within the clause it expresses an opinion about” (Butt et al, 2000:116), but for the purpose of the present discussion, a grammatical metaphor is considered only from its interpersonal meaning.

From an experiential point of view, SAY can be a verbal Process (94%) or a mental one (6%)ⁱⁱ. Within the transitivity system it can occur in four different environments: i) a main Process in projecting verbal clauses; ii) a main Process in non-projecting verbal clauses; iii) a Process in a clause that functions as a Participant; and iv) a Process in a clause functioning as Qualifier of a nominal element in the clause complex. As it was said, the most frequent environment of SAY is in a projecting verbal clause, where it is the main Process and projects a quote or a report. The possibilities found in the corpus are:

Experiential meaning

- 84.7% Process in projecting verbal clauses
- 8.2% Process in verbal clauses functioning as Qualifier
- 3.1% Process in verbal clauses functioning as Participant
- 2.0% Process in non-projecting verbal clauses

Interpersonal meaning (grammatical metaphor)

2.0% Process in clauses functioning as interpersonal grammatical metaphor

These are some examples from the corpus.

Example 1 - Environments of SAY from a perspective of experiential meaning

Process in projecting verbal clauses

'No, I won't be late,' **said** Walter, unhappily and guiltily certain that he would be.

Process in verbal clauses functioning as Qualifier

He realized that she had not taken the slightest interest in what he had been **saying**, had not even troubled to pay the least attention.

Process in verbal clauses functioning as Participant

I seem to remember my wife having **said**... Quite informal.

Process in non-projecting verbal clauses

And when one met a potential voter, what on earth one to **say**?

Example 2 - Environments of SAY from a perspective of interpersonal meaning (grammatical metaphor)

Interpersonal meaning - Process in clauses functioning as interpersonal grammatical metaphor

'I must say,' Polly handsomely admitted, 'he doesn't look it.'

Although DIZER and DECIR are prototypical equivalents of SAY, these items present some different environments of manifestation. DIZER is mainly a Process realizing experiential meaning (90%), but it also occurs as interpersonal grammatical metaphor (10%). DIZER is mainly a verbal Process (92%) as well, but it also occurs as mental (4%) and relational (4%). According to the data analysed, it can be part of clauses realizing experiential, interpersonal or textual meanings. It can be part of an element in the transitivity system, part of an interpersonal grammatical metaphor and part of a sentence that works as a conjunctive Adjunct. Within the transitivity system it can occur in five different environments: i) a main Process in a projecting verbal clause; ii) a main Process in a

non-projecting verbal clause; iii) a Process in a clause that functions as Participant; iv) a Process in a clause functioning as Circumstance; and a Process in a clause functioning as Qualifier of a nominal element in the clause complex. As it was said, the most frequent environment of DIZER is in a projecting verbal clause, where it is the main Process and projects a quote or a report just like SAY. The possibilities found in the corpus are:

Experiential meaning

63.6% Process in projecting verbal clauses

12.5% Process in verbal clauses functioning as Qualifier

7.9% Process in non-projecting verbal clauses

3.4% Process in verbal clauses functioning as Participant

2.3% Process in verbal clauses functioning as Circumstance

Interpersonal meaning (grammatical metaphor)

9.1% Process in clauses functioning as interpersonal grammatical metaphor

Textual meaning (grammatical metaphor)

1.1% Process in clauses functioning as conjunctive Adjunct

These are some examples from the corpus.

Example 3 - Environments of DIZER from the perspective of experiential meaning

Process in projecting verbal clauses

- Tu és tão boa - **dissera**-lhe.

Process in verbal clauses functioning as Qualifier

- Se é tudo o que me querias **dizer**, julgo que então posso...

Process in non-projecting verbal clauses

(por que não lhe podia **dizer** a crua, a brutal verdade sem rebuços? no fim das contas, ela já sabia);

Process in verbal clauses functioning as Participant

- Lembro-me de ter ouvido minha mulher **dizer**... Algo muito sem cerimônia...

Process in verbal clauses functioning as Circumstance

A mãe caminhava devagar, sem nada **dizer**.

Example 4 - Environments of DIZER from the perspective of interpersonal and textual meaning (grammatical metaphor)

Interpersonal meaning - Process in clauses functioning as interpersonal grammatical metaphor
 Admitia tudo, **por assim dizer**: que o pobre Eric tivesse morrido prematuramente, que tivesse sofrido em sua doença, relutado em deixar a vida - a música admitia tudo.

Textual meaning - Process in clauses functioning as conjunctive Adjunct
Dito isto passou adiante, deixando as duas moças a se entreolharem, atarantadas e vermelhas.

DECIR is also mainly a Process realizing experiential meaning (92%), with some occurrences as interpersonal grammatical metaphor (8%). DECIR is mainly a verbal Process (95%) as well, but it also occurs as mental (2%) and relational (3%).

According to the data analysed, it can be part of clauses realizing experiential, interpersonal or textual meanings. It can be part of an element in the transitivity system, part of an interpersonal grammatical metaphor and part of a sentence that works as a conjunctive Adjunct. Within the transitivity system it can occur in five different environments: i) a main Process in a projecting verbal clause; ii) a main Process in a non-projecting verbal clause; iii) a Process in a clause that functions as Participant; iv) a Process in a clause functioning as Circumstance; and a Process in a clause functioning as Qualifier of a nominal element in the clause complex. As it was said, the most frequent environment of DECIR is in a projecting verbal clause, where it is a main Process and projects a quote or a report just like SAY and DIZER. The possibilities found in the corpus are:

Experiential meaning

- 75.9% Process in verbal projecting clauses
- 6.9% Process in verbal non-projecting clauses
- 4.3% Process in verbal clauses functioning as Qualifier
- 3.4% Process in verbal clauses functioning as Participant
- 1.7% Process in verbal clauses functioning as Circumstance

Interpersonal meaning (grammatical metaphor)

6% Process in clauses functioning as interpersonal grammatical metaphor

Textual meaning

1.7% Process in clauses functioning as conjunctive Adjunct

These are some examples from the corpus.

Example 5 - Environments of DECIR from the perspective of experiential meaning

Process in projecting verbal clauses

-No; no volveré tarde **-dijo** Walter, con la culpable y desdichada certeza de que lo haría.

Process in non-projecting verbal clauses

¿Por qué no le había **dicho**, sin rodeos, la verdad cruda y brutal?

Process in verbal clauses functioning as Qualifier

Se dio cuenta de que ella no había prestado el menor interés a lo que le había **dicho**, que no se había tomado siquiera el trabajo de poner la menor atención.

Process in verbal clauses functioning as Participant

Creo recordar que me ha **dicho** mi mujer... Algo como en familia...

Process in verbal clauses functioning as Circumstance

- Pero no se habrá usted ofendido por lo que yo haya **dicho**, ¿verdad? - preguntó Lady Edward con grandes muestras de inquietud y contrición.

Example 6 - Environments of DECIR from the perspective of interpersonal and textual meaning (grammatical metaphor)

Interpersonal meaning - Process in clauses functioning as interpersonal grammatical metaphor

La música lo admitía todo, **por así decir**: la prematura muerte de Eric, el sufrimiento de su enfermedad, su apego a la vida: lo admitía todo.

Textual meaning - Process in clauses functioning as conjunctive Adjunct

El arte nos da la sensación, la idea y el sentimiento absolutamente puros: químicamente puros, **quiero decir** - había agregado riendo-, no moralmente.

From this data, obtained through SPSS tools and summarised in Table 2 below, we can see that although the three items have great similarity in functional roles, they also have some differences.

	SAY		DIZER		DECIR	
Occurrences	98	100%	88	100%	116	100%
1. Experiential	96	98%	79	90%	107	92%
2. Grammatical Metaphor	2	2%	9	10%	9	8%
1.1						
Verbal Process	90	94%	73	92%	102	95%
Mental Process	6	6%	3	4%	2	2%
Relational Process	0	-	3	4%	3	3%
1.2						
Projecting verbal clause	83	86,5%	56	70,9%	88	82,2%
Non-projecting verbal clause	2	2,1%	7	8,9%	8	7,5%
clause as Qualifier	8	8,3%	11	13,9%	5	4,7%
clause as Participant	3	3,1%	3	3,8%	4	3,7%
clause as Circumstance	0	-	2	2,5%	2	1,9%
2.1						
Interpersonal metaphor	2	100%	8	89%	7	78%
Conjunctive Adjunct	0	-	1	11%	2	22%

Table 2 – Contextual environments of SAY, DIZER and DECIR

SAY is most frequent as Process in projecting verbal clauses; it does not occur as a relational Process as the other two; it does not occur in clauses functioning as Circumstance or in clauses functioning as conjunctive Adjunct; it has few occurrences as Process in non-projecting verbal clauses; it also has the lowest frequency as interpersonal grammatical metaphor.

DIZER, on the other hand, is less frequent than SAY and DECIR as Process in projecting verbal clauses; it occurs as a mental and relational Process; it occurs as Process in clauses functioning as Circumstance and in clauses functioning as conjunctive Adjunct; it has the highest frequency as Process in clauses functioning as Qualifier and as interpersonal grammatical metaphor.

And, finally, DECIR occurs less as Process in projecting verbal clauses than SAY but more than DIZER; it occurs as a mental and relational Process; it occurs as Process in clauses functioning as Circumstance and in clauses functioning as interpersonal grammatical metaphor; it has the

highest frequency as Process in clauses functioning as conjunctive Adjunct.

Now, it is possible to analyse the equivalent relations among these items considering their different contextual environments. Based on the concept of translation techniques (Molina & Hurtado-Albir, 2002), the equivalent relations were classified into five categories: established equivalent, other verbal Process, non-verbal Process, omission and amplification. Considering that SAY, DIZER and DECIR are prototypical equivalents among the three languages the translator can choose to translate SAY into the established equivalent. However, it can also be translated by another verbal Process or still by a non-verbal Process; it is also possible for an occurrence of SAY in the original text to be omitted or for an item (DIZER/DECIR) in the translation to “introduce details that are not formulated in the ST” (Molina & Hurtado-Albir, 2002:510), i.e., it is an amplification.

It is important here to say that equivalent relations are **not** taken here as a one way road, i.e., at the same time we consider how SAY is translated into Portuguese and Spanish, we also consider how DIZER and DECIR, its typical equivalents, relate to SAY and which other lexical items are involved in both directions (from original to translations and from translations to original). We also are considering here two kinds of equivalence: “word-rank translation equivalence” and “conditioned probabilities” modelled by “higher-ranking environment” (Matthiessen, 2001:76); the first is a step in order to analyse the second.

Within the 98 occurrences of SAY and considering the translation into Spanish, 88.8% is translated as the established equivalent, 9.2% as other verbal Process and 2.0% is omitted. Considering the translation into Portuguese, 60.2% is translated as the established equivalent, 20.4% as other verbal Process, 16.3% is omitted and 3.1% is translated as non-verbal Process.

Within the 88 occurrences of DIZER, 67% retextualize the established equivalent, 15.9% relates to other verbal Process, 9.1% relates to non-verbal Process and 8% is amplification.

Within the 116 occurrences of DECIR, 75% retextualize the established equivalent, 15.5% relates to other verbal Process, 7.8% relates to non-verbal Process and 1.7% is amplification.

	SAY - Portuguese		SAY – Spanish		DIZER – English		DECIR – English	
1	59	60.2%	87	88.8%	59	67%	87	75%
2	20	20.4%	09	9.2%	14	15.9%	18	15.5%
3	03	3.1%	-	-	8	9.1%	09	7.8%
4	16	16.3%	02	2.0%	-	-	-	-
5	-	-	-	-	07	8.0%	02	1.7%

Table 3 – Summary of word-rank translation equivalence

1-Established equivalent, 2-Other verbal Process, 3-Non-verbal Process, 4-Omission, 5-Amplification

We can call this “word-rank translation equivalence” and after this first step, we can focus on the “conditioned probabilities”. The point is to try to describe how the environments where these items occur and the meaning they realize condition the kind of translation equivalence they present.

As we can see in Table 4 below, related to the contextual environments that condition equivalence of SAY into Portuguese, when SAY occurs as a Process in a projecting verbal clause it is 57.8% of the times translated as its established equivalent, with 19.3% translated as other verbal Process, 3.6% as non-verbal Process and 19.3% being omitted. The other environments have different equivalent relations; as a Process in non-projecting verbal clauses, SAY is translated 50% as its established equivalent and 50% as other verbal Process; in clauses functioning as Participant it is translated 100% as its established equivalent; and in clauses functioning as Qualifier, 75% as its established equivalent and 25% as other verbal Process; as interpersonal grammatical metaphor it is translated 50% as its established equivalent and 50% as other verbal Process.

Contextual environments * Translation English/Portuguese Crosstabulation

			Translation English/Portuguese				Total
			equivalent	verbal	non-verbal	omission	
Contextual environments	non-projecting	Count	1	1			2
		% within Contextual environments	50,0%	50,0%			100,0%
	projecting	Count	48	16	3	16	83
		% within Contextual environments	57,8%	19,3%	3,6%	19,3%	100,0%
	participant	Count	3				3
		% within Contextual environments	100,0%				100,0%
	interpersonal	Count	1	1			2
		% within Contextual environments	50,0%	50,0%			100,0%
	qualifier	Count	6	2			8
		% within Contextual environments	75,0%	25,0%			100,0%
Total		Count	59	20	3	16	98
		% within Contextual environments	60,2%	20,4%	3,1%	16,3%	100,0%

Table 4 – Contextual environments conditioning equivalence of SAY into Portuguese

According to Table 5 below, which presents the data related to contextual environments that condition equivalence of SAY into Spanish, when SAY occurs as a Process in a projecting verbal clause it is 91.6% translated as its established equivalent, with only 6.0% translated as other verbal Process and 2.4% being omitted. The other environments have different equivalent relations; as a Process in non-projecting verbal clause, SAY is translated 50% as its established equivalent and 50% as other verbal Process; in clauses functioning as Participant it is translated 100% as its established equivalent; and in clauses functioning as Qualifier 75% as its established equivalent and 25% as other verbal Process; as interpersonal grammatical metaphor it is translated 50% as its established equivalent and 50% as other verbal Process.

Contextual environments * Translation English/Spanish Crosstabulation

			Translation English/Spanish			Total
			equivalent	verbal	omission	
Contextual environments	non-projecting	Count	1	1		2
		% within Contextual environments	50,0%	50,0%		100,0%
	projecting	Count	76	5	2	83
		% within Contextual environments	91,6%	6,0%	2,4%	100,0%
	participant	Count	3			3
		% within Contextual environments	100,0%			100,0%
	interpersonal	Count	1	1		2
		% within Contextual environments	50,0%	50,0%		100,0%
	qualifier	Count	6	2		8
		% within Contextual environments	75,0%	25,0%		100,0%
Total	Count	87	9	2	98	
	% within Contextual environments	88,8%	9,2%	2,0%	100,0%	

Table 5 - Contextual environments conditioning equivalence of SAY into Spanish

In relation to DIZER, Table 6 below shows that when DIZER occurs as a Process in a projecting verbal clause it is 85.7% translated as its established equivalent, 12.5% is translated as other verbal Process, and 1.8% as non-verbal Process. The other environments have different equivalent relations; as a Process in non-projecting verbal clauses, DIZER is translated 14.3% as its established equivalent, 28.6% as other verbal Process, 42.9% as non-verbal Process, and 14.3% is amplification; in clauses functioning as Participant it is translated 100% as its established equivalent; and in clauses functioning as Qualifier 54.5% as its established equivalent, 18.2% as other verbal Process, and 27.3% as non-verbal Process; as interpersonal grammatical metaphor it is translated 12.5% as its established equivalent, 25% as other verbal Process, 50% as non-verbal Process, and 12.5% is amplification; and the only occurrence as conjunctive Adjunct is amplification.

Contextual environments * Translation English/Portuguese Crosstabulation

			Translation English/Portuguese				Total
			equivalent	verbal	non-verbal	amplification	
Contextual environments	non-projecting	Count	1	2	3	1	7
		% within Contextual environments	14,3%	28,6%	42,9%	14,3%	100,0%
	projecting	Count	48	7	1		56
		% within Contextual environments	85,7%	12,5%	1,8%		100,0%
	participant	Count	3				3
		% within Contextual environments	100,0%				100,0%
	circumstance	Count		1		1	2
		% within Contextual environments		50,0%		50,0%	100,0%
interpersonal	Count	1	2	4	1	8	
	% within Contextual environments	12,5%	25,0%	50,0%	12,5%	100,0%	
conjunctive	Count				1	1	
	% within Contextual environments				100,0%	100,0%	
qualifier	Count	6	2		3	11	
	% within Contextual environments	54,5%	18,2%		27,3%	100,0%	
Total	Count	59	14	8	7	88	
	% within Contextual environments	67,0%	15,9%	9,1%	8,0%	100,0%	

Table 6 – Contextual environments conditioning equivalence of DIZER into English

About DECIR, Table 7 below shows that when DECIR occurs as a Process in a projecting verbal clause it is 87.5% translated as its established equivalent, 10.2% is translated as other verbal Process, 1.1% as non-verbal Process and 1.1% is amplification. The other environments have different equivalent relations; as a Process in non-projecting verbal clauses, DIZER is translated 12.5% as its established equivalent, 37.5% as other verbal Process, 37.5% as non-verbal Process, and 12.5% is amplification; in clauses functioning as Participant it is translated 50% as its established equivalent and 50% as non-verbal Process; and in clauses functioning as Qualifier 100% as its established equivalent; as interpersonal grammatical metaphor it is translated 14.3% as its established equivalent, 71.4% as other verbal Process, and 14.3% as non-verbal Process; and as conjunctive Adjunct it is 100% non-verbal Process.

Contextual environments * Translation English/Spanish Crosstabulation

			Translation English/Spanish				Total
			equivalent	verbal	non-verbal	amplification	
Contextual environments	non-projecting	Count	1	3	3	1	8
		% within Contextual environments	12,5%	37,5%	37,5%	12,5%	100,0%
	projecting	Count	77	9	1	1	88
		% within Contextual environments	87,5%	10,2%	1,1%	1,1%	100,0%
	participant	Count	2		2		4
		% within Contextual environments	50,0%		50,0%		100,0%
	circumstance	Count	1	1			2
		% within Contextual environments	50,0%	50,0%			100,0%
interpersonal	Count	1	5	1		7	
	% within Contextual environments	14,3%	71,4%	14,3%		100,0%	
conjunctive	Count			2		2	
	% within Contextual environments			100,0%		100,0%	
qualifier	Count	5				5	
	% within Contextual environments	100,0%				100,0%	
Total	Count	87	18	9	2	116	
	% within Contextual environments	75,0%	15,5%	7,8%	1,7%	100,0%	

Table 7 – Contextual environments conditioning equivalence of DECIR into English

These data shows that when the kind of equivalence SAY, DIZER and DECIR present is considered within the contextual environments where they realize meaning some patterns can be observed changing the probabilities presented by word-rank equivalence (see Table 3).

The established equivalent is the most frequent option when the lexical items occur as Process in projecting verbal clauses; in fact, in general, it seems that environments that realize experiential meaning tend to have the established equivalent as the most frequent option while environments that realizes grammatical metaphors the most frequent is non-equivalent options.

4. Final remarks

This paper analysed the lexical items SAY, DIZER and DECIR, multi-contextualising their functional roles in English, Portuguese and Spanish, in a cross-linguistic approach that checks the possibilities of equivalence among these languages related to these specific items in a specific register.

Although the word-rank translation equivalence shows that the established equivalent is the most frequent equivalent among the items, high-rank analysis presents a different perspective, namely, that the established equivalent is more frequent when it realizes experiential meanings and that it is less frequent when it realizes metaphorical ones.

These preliminary findings will be further investigated with a larger corpus in order to prove this hypothesis and develop a probabilistic interpretation of translation equivalence based on systemic theory.

This type of investigation may be useful for the developing of translation aids, such as dictionaries which present wider contexts of equivalence, and it gives insights into important theoretical issues for translation training with practical applications in Machine Translation.

REFERENCES

- BUTT, David et. al. 2000. *Using functional grammar: an explorer's guide*. 2^a ed. Sydney: Macquarie University.
- CATFORD, F. C. 1965. *A linguistic theory of translation*. London: Oxford University Press. (Language and Language Learning).
- JESUS, Silvana Maria de & Janaina Minelli de OLIVEIRA. 2003. Brasil-Canadá: diálogos entre mulheres negras. In: DINIZ, Dilma Castelo Branco (Org.). *Brasil-Canadá: confrontos literários e culturais*. Belo Horizonte: Faculdade de Letras/ABECAN/NEC/UFMG: 123-146.
- JESUS, Silvana Maria de. 2004. *Representação do discurso e tradução: padrões de textualização em corpora paralelo e comparável*. Dissertação (Mestrado em Letras/Linguística Aplicada). Belo Horizonte: Faculdade de Letras, UFMG/PosLin.
- HALLIDAY, M.A.K. et al. 1964. *The linguistic sciences and language teaching*. London: Longmans. (Longmans' Linguistics Library).
- HALLIDAY, M. A. K. 1991. Corpus studies and probabilistic grammar. In: AIJMER, Karin., ALTENBERG, Bengt. (Ed.). *English corpus linguistics: studies in honour of Jan Svartvik*. London e New York: Longman: 30-43.
- _____. 2001. Towards a theory of good translation. In: STEINER, E., YALLOP, C. (Ed.). *Exploring translation and multilingual text production: beyond content*. Berlin, New York: Mouton de Gruyter: 13-18.
- _____. 2005. *Computational and quantitative studies*. Jonathan J. Webster (Ed.). London, New York: Continuum: 42-189. (Collected works of M. A. K. Halliday, v. 6).
- HALLIDAY, M. A. K. & MATTHIESSEN, Christian M. I. M. 2004. *An introduction to functional grammar*. 3rd edition, rev. ampl. London: Arnold.
- MATTHIESSEN, Christian M.I.M. 1993. Register in the round: diversity in a unified theory of register analysis (Part VI. A unified theory of register analysis). In: GHADESSY, Mohsen (Ed.) *Register analysis: theory and practice*. London: Pinter Publishers: 220-292. (Open Linguistics Series).
- _____. 1999. The system of transitivity: an exploratory study of text-based profiles. *Functions of language*, 6.1: 1-51, John Benjamins.
- _____. 2001. The environments of translation. In: STEINER, E., YALLOP, C. (Ed.). *Exploring translation and multilingual text production: beyond content*. Berlin, New York: Mouton de Gruyter: 41-124.
- MOLINA, Lucía, HURTADO-ALBIR, Amparo. 2002. Translation techniques revisited: a dynamic and functionalist approach. *Meta*, XLVII, 4: 498-512. Disponível em: < <http://www.erudit.org/erudit/meta/v43n04/index.html> >. Acesso em: 23 de novembro de 2004.
- NESBITT, Christopher; PLUM, Guenter. 1988. Probabilities in a systemic-functional grammar: the clause complex in English. In: FAWCETT, Robin, YOUNG, David J. (Ed). *New developments in systemic linguistics*. London, New York: Printer Publishers: 6-38. (Theory and application, v. 2).

- TEICH, Elke. 2001. Towards a model for the description of cross-linguistic divergence and commonality in translation. In: STEINER, E., YALLOP, C. (Ed.). *Exploring translation and multilingual text production: beyond content*. Berlin, New York: Mouton de Gruyter: 191-227.
- _____. 2003. *Cross-linguistic variation in system and text: a methodology for the investigation of translations and comparable texts*. Berlin: Mouton de Gruyter. (Text, translation, computational processing, 5).

ⁱ SAY, DIZER and DECIR are used as lemma, i.e., they include all the forms all of these verbs appearing in the corpus.

ⁱⁱ For the purpose of this paper, the occurrences of SAY, DIZER and DECIR as non-verbal process will not be considered separately.