THE NATURE OF FIXED LANGUAGE
IN THE SUBTITLING OF A DOCUMENTARY FILM

Cláudia Susana Nunes Martins

Escola Superior de Educação – Instituto Politécnico de Bragança

Universitat Rovira i Virgili

A dissertation submitted in partial fulfillment of the requirements

for the Diploma of Advanced Studies

Doctoral program in Translation and Intercultural Studies

Universitat Rovira i Virgili, Spain

Supervisors:

Professor Josélia Neves, Instituto Politécnico de Leiria, Portugal

Professor Belinda Maia, Universidade do Porto, Portugal

2007

Table of Contents

Abstract

This research is based on the assumption that audiovisual translation (AVT) performs a social and cultural function over its viewers, thus still justifying studies that focus on linguistic issues. This is particularly striking for the reading literacy in countries such as Portugal, that are traditionally subtitling countries, even if dubbing and voice-over are also occasionally used, particularly in the case of documentaries.

The purpose for this research is to analyze the examples of restricted lexical occurrence, i.e. set phrases or semantic phrasemes (Mel'čuk 1995), depending on the terminology chosen, within the two versions of the AVT – subtitling and voice-over – of a documentary film called "The Real Da Vinci Code".

In order to achieve this purpose, we chose to follow a case study methodology that allowed us to center our attention on and single out a specific feature of language – set phrases – used in this type of audiovisual texts – documentary films. Nonetheless, we will later attempt to combine this methodology with corpus studies at the level of the doctorate thesis.

The analysis conducted intended to identify set phrases in English and their respective Portuguese translations and to classify them in terms of the terminology used by Mel'čuk (1995), as well as to combine this with the identification of possible metaphors that could be on their basis. Then, from their interpretation and translation analysis, it was possible to confirm our initial hypotheses, which were that set phrases are as usual in general language as in specialized texts within the audiovisual context and the informative genre, namely documentary films, and that occasionally set phrases are omitted by translators either because they miss to identify them or because of constraints related to translation or to the target language. As a final point, we managed to prove that everything that is on screen at the time of the utterance of the set phrase is of influence to their strengthening.

Keywords: audiovisual translation; subtitling; dubbing and voice-over; word combinations; free and restricted lexical co-occurrence; setness or frozenness; idiomaticity; collocations; idiomatic expressions.

List of tables

List of figures

Figure 1. Printscreen of "on the heart of (any quest)".

Figure 2. Printscreen of "think-tank".

Figure 3. Printscreen of "(make one's) legwork".

Figure 4. Printscreen of "a spit in the ocean".

Figure 5. Printscreen of "one hell of (a journey)".

Figure 6. Printscreen of "follow the footsteps".

Acknowledgements

The moment to acknowledge people's help at the end of any piece of work is always complicated, due to the effect that any small gesture, positive or negative, can hold on a choice or a decision made.

Nevertheless, I wish to thank:

- Josélia Neves for accepting my crazy and dazzling dream and helping me shape it into a decent piece of research, and for being constantly demanding;
- Belinda Maia for her continuous belief in me and my abilities, her endless support and her offer of new and stimulating research paths;
- Anthony Pym for having me in his doctorate programme;
- And my family for their everlasting patience and encouragement.

1. Introduction

Translation Studies (TS) has been based on the idea that translation is a process of language transfer, which initially excluded AVT from its realm (an issue elicited below). However, with research conducted by scholars throughout the last decades, such as Gambier, Gottlieb or Díaz-Cintas, among many others, it came to be regarded as part of this field of knowledge.

Luyke (1991: 153-155) argues that, although AVT is a form of translation, it differs from other types of translation in some important aspects. To begin with, the message that is to be transferred from one language into other(s) is expressed by everything within the screen, namely the image, the acting, the sound and the language, which means that the linguistic transfer will only replace the message and thus cannot change any of those other fulcral components. Secondly, audiovisual language transfer is unable to use resources like those from other translation forms, such as "explanatory footnotes, asterisks or asides" (Luyke 1991: 154), but, at the same time, it must be complete in such a way as to be understood, being that it must delete things from the original. Thirdly, the text transferred is shorter than the original, "a mere fraction of the original dialog" (Luyke 1991: 154), forcing translators to drastically abridge the text. Finally, audiovisual translators must integrate editorial skills concerning ommisions or additions of information and condensation of the original.

Even though AVT involves a considerable degree of adaptation of the original texts to the audiovisual means of communication, it has to be accepted as translation. In Gambier (2003), AVT is described as an example of *transadaptation*, though this name has been highly questioned.

Transadaptation means to go beyond the dichotomy between literal and free translation, translation and adaptation, among others, and takes audience into account. This concept means that translating in the audiovisual context has nothing to do with word-for-word transfer, but comprehends:

> a set of strategies that might include summarizing, paraphrasing, etc. (…) [as well as] taking into consideration the genre, the film-maker's style, the needs and expectations of viewers (…) and the multimodality of audiovisual communication. (Gambier in Gambier 2003: 178)

According to Orero (2004: vii), another question to elaborate on is:

> the unsettled terminology of AVT (…) A step further would be to agree on a generic name to define the multiple and different modes of translation when the audio (radio), the audio and the visual (screen), or the written, the audio and the visual (multimedia) channels are the source text. (Orero 2004: vii)

This terminological unsettleness that Orero (2001: vii) refers to is quite clear in the abundance of terms used to refer to the audiovisual field: *traducción subordinada* or constrained translation (Titford 1982; Mayoral 1984; Rabadán 1991; Díaz-Cintas 1998; Lorenzo & Pereira 2000 and 2001); film translation (Snell-Hornby 1988); film and TV translation (Delabastita 1989); screen translation (Mason 1989); media translation (Eguíluz 1994); film communication (Lecuona 1994); *traducción fílmica* (Díaz-Cintas 1997), audiovisual translation (Luyken 1991; Dries 1995; Shuttleworth & Cowie 1997; Baker 1998); (multi)media translation (Gambier & Gottlieb 2001). Other authors have developed this issue of AVT's terminological inconsistency, for instance Chaume (2003), as shall be seen in chapter 3.

AVT is then regarded as the designation encompassing "all translations – or multisemiotic transfer – for production or postproduction in any media or format, and also new areas of media accessibility" (Orero 2004: viii), such as subtitling for the deaf and hard-of-hearing or audiodescription, among others.

For Gambier (in Gambier 2003: 171), AVT is the term that brought "to the foremost the multisemiotic dimension of all broadcast programmes", surpassing designations such as 'language transfer', which focus too much on language, 'versioning', 'screen translation', that covers products distributed by means of a screen, and 'multimedia translation' that may lead to confusion because of implications related to the theater, comics, films, TV, cinema, video and on-line and off-line products and services. However, AVT has come to include all these outputs on the most varied formats, as has been briefly mentioned above.

Because of its complexity and multisemiotic wealth, research in AVT may "draw on a variety of (…) methodologies – from polysystem theory, psycholinguistics, cultural studies, critical discourse analysis, relevance theory, as well as functional approaches to translation" (Gambier in Gambier 2003: 183), attempting to tackle

concepts such as those of text, meaning, norms, equivalence, manipulation or acceptability.

Therefore, our research has made the assumption that AVT is a case of transadaptation as its starting point and chose to study a documentary film in its two available audiovisual versions – subtitling and voice-over – in Portugal, so as to analyse examples of set phrases found in the original sound track and its two translated versions.

The fact that we chose to conduct research on a linguistic item (as this one of phraseologies in general) within AVT has given rise to a great deal of criticism, echoed in the words of Gambier that considers it "a misuse of time" if we compare it to other much more interesting and useful research. Nonetheless, we must not forget that in a traditionally subtitling country such as Portugal, since the *Estado Novo*, the Portuguese dictatorship, "the power of the written word" is of the utmost importance. It is even more striking if we think of the importance of subtitling for the development of reading habits and literacy in a population with an unusual reading deficit. We should also mention that the Portuguese illiteracy rate is around 10-12%. Finally, there is also a limited number of research done in AVT in Portugal if we have other European countries' numbers as a *tertium comparationis*.

To achieve these purposes, the present minor dissertation has been divided into several chapters. Firstly, we will explain that this research was included in a product-oriented approach at the level of AVT and in the field of linguistics, namely lexicography and phraseology. Although it was our initial intention to deal with AVT from a process-oriented approach, it was not possible to have access to the process in which the material for our pilot study went through, nor did we manage to contact the translators/subtitlers involved in the same process. It is still a point to be taken into account at the level of the doctorate thesis.

Secondly, we shall divide this research into chapters: after clarifying our theoretical framework, a section for dealing with the issues involved in the discussion of word combinations, with special emphasis on set phrases or restricted lexical co-occurrence (that is collocations and idiomatic expressions), will follow; then another chapter to approach AVT, specifically subtitling, on the one hand, and dubbing and voice-over, on the other; afterwards, a part to deal with one of the functional TS approaches we wish to develop – the theory of text-types and their relation to genres, specifically a brief characterization of audiovisual genres and of documentary films.

Finally, the research methodology will be elicited and the pilot study presented, with some of the results found discussed in the light of its theoretical framework.

2. Theoretical Framework

The theoretical approaches that will sustain our research are two-fold and will enable the intertwining of fields not frequently linked, namely fixed language and AVT, both in their subtitled and voiced-over versions.

These topics have been chosen not only because of the social importance of subtitling in the development of viewers' literacy, especially in a subtitling country such as Portugal, but also to test the extent to which the manifestations of fixed language (free and restricted word combinations) are different from one language to another and are either maintained or dropped when subtitling documentary films, as opposed to voiced-over versions.

Nevertheless, there are a number of previous research studies that have combined AVT and linguistic aspects, specifically those pertaining to the field of free and restricted word combinations. One of the first most-known ones we should refer to is the PhD thesis of Gottlieb (1997) – "Subtitles, Translation and Idioms". To name a few more recent pieces of research, those conducted by Schröter (2005) – "Shun the pun, rescue the rhyme? The dubbing and subtitling of language-play in film" that analyzes language-play in the dubbing and subtitling of 18 original films and their 99 various target versions; by Araújo (2004) – "To be or not to be natural: clichés of emotion in Screen translation", which focuses and classifies 250 clichés of emotion in 5 dubbed and subtitled films in Brazilian Portuguese; by Fuentes (2001) – "La recepción del humor audiovisual traducido: estudio comparativo de fragmentos de las versions doblada y subtitulada al español de la película 'Duck Soup', de los Hermanos Marx"; or by Gómez (1994) – "Calcos sintácticos, fraseológicos y pragmáticos en los doblajes del inglés al español".

At this moment, having presented our working theme, it is of the utmost importance to present a concise chronological overview of several TS theories, so as to later explain the theoretical framework we have chosen to adopt in the course of our research.

One of the first translation approaches Hermans (1999: 55) accounts for is that of Nida's science of translating (1964), within the equivalence theory, who by means a three-step technique intended to compare translations with their sources and different translations with their original. This one takes the source text into consideration and is regarded by Hermans (idem) as "too rudimentary".

Within this same model dealing with equivalence, Munday (2001) refers to Jakobson, Newmark and Koller. Equally, this author regards Vinay and Darbelnet (and their comparative stylistic analysis), Catford (and his translation shifts), Levý (and his decision-making model) and van Leuven-Zwart as examples of the translation shift approach.

From these, Levý (cit. Hermans 1999: 73) emphasizes, in 1967, the act of translating as a decision-making process, in which the translator chooses the best option among several alternatives, a decision that influences other decisions to be made and determines the shape of the final product. In response to this model, Popovič (1970 cit. Hermans 1999: 74) discusses the idea that translation is "a confrontation of two sets of linguistic and discursive norms and conventions", those of the source and the target texts and cultures, an issue to be recovered later by other authors. If certain choices are taken in a regular way throughout a large number of texts, then these decisions will become norms, an topic to be discussed further below.

Between the 1980 and 1995 works by Toury, van Leuven-Zwart's (cit. Hermans 1999: 58-59) comparative-descriptive model from 1984 should be mentioned since she carried out a piece of research that could be valuable for ours, though it deals with literature. Its main objective is to "describe and catalogue 'shifts' in translation, and to deduce from these the translator's underlying strategy or norm. (…) [trying] to provide (…) 'tendencies' in certain translations." This is a concept studied by Catford in his "Translation Shifts" (in Venuti 2000: 141-147), in which he distinguishes between formal correspondent and textual correspondent and mentions level and category shifts (these including structural, class, unit/rank and intra-system shifts).

By applying Vinay & Darbelnet and Levý's categories to the descriptive analysis of translation, van Leuven-Zwart attempeted to "both systematize comparison and to build in a discourse framework above the sentence level" (Munday 2001: 63), as well as by analyzing translators' translational trends which provide information on the norms they apply. The comparison between ST and TT in their microstructural shifts demands the selection of samples that are to be divided into "comprehensible textual

units called 'transemes' " (Munday 2001: 64). As a result, a *transeme* is this model's basic unit, that can be divided into "the state of affairs transeme and the satellite transeme" (van Leuven-Zwart cit. Hermans 1999: 58). These transemes result from semantic, stylistic or pragmatic shifts identified in source and target texts and these are analyzed in terms of similarities and dissimilarities.

According to van Leuven-Zwart (cit. Hermans 1999: 59-60), the application of this transeme approach to translation is done in two stages: one is the identification of similarities which allows for the construction of the *architranseme*; the other being the comparision of each transeme to the architranseme. From this comparison, four possible relationships can be determined: synonymic if both transemes are synonymous with the architranseme, thus no shift being found; hyponymic if one transeme is synonymous and the other hyponymic; contrastive if both transemes are hyponymic; and, finally, if no relationship is to be found between the transemes. The cases of hyponymy (x is a form/class/mode of y) are categories of micro-structural shifts that are referred to as modulation, assuming either the form of specification or of generalization. When no similarities are identified and no relationship established, we have mutation. In addition, this comparative model can also be applied to macro-structure because semantic shifts can lead to changes in the ideology transmitted by the discourse.

Nevertheless, some criticism has been made to this model (Stegeman 1991; Linn 1993; Hermans 1999), especially due to the fact that it involves a certain degree of interpretation, not properly worked on, it does not touch important issues, such as those of cohesion and genres, and it provides an idea of transeme that seems to go beyond culture, in an almost idealistic perspective that there is direct access to the concepts of what the words convey. However, van Leuven-Zwart's model has been put into action by several of her students.

Moreover, in 1991, Stegeman (cit Hermans 1999: 63-64) presents his reader-response criticism, focusing on the significance of the reader's role, since readers react differently to translations and their originals. A text only comes to life when a reader reacts to it, because it functions as a stimulus to communication. In his approach, Stegeman tried to ascertain the concept of equivalence and he claims that "equivalence is obtained when no significant difference can be observed in the way source-language readers react to a source text and target-language readers react to the corresponding target text" (Hermans 1999: 63). Although he covered micro-structural, macro-structural and paratextual features of the texts, Hermans (1999: 64) concludes that he

has only dealt with the "psychology of reading in artificial laboratory situations" and little with the description and analysis of translation.

The contextual model (Lambert & Lefevere 1977; Lambert & van Gorp 1985 cit. Hermans 1999: 63-65) draws on the need to understand translation as a cultural phenomenon, rather than a mere comparison of texts, including genres or translators' status, among other aspects. Thus, translation should be understood as a process involving two communication systems and, according to Lambert & Van Gorp (cit. Hermans 1999: 66-68) their aim was to study "translational norms, models, behaviour and systems" by means of a checklist of preliminary data, macro-level, micro-level and systematic context, emphasizing "the link between the individual case study and the wider theoretical framework" (Munday 2001: 121).

However, Hermans (1999: 68) views Lambert & van Gorp's models as too general and even if they focus on "multidimentionality and flexibility (…) [they remain] essentially binary". The attempt to be general leads to oversimplification and to neat divisions between source and target languages and cultures. On the other hand, Delabastita's scheme (1989 cit. Hermans 1999: 68) was more comprehensive and used for audiovidual texts by combining rhetoric with norms and providing checklists specifically for subtitling and dubbing.

At the end of the presentation of these models and schemes, Hermans (1999: 71) summarises that these are only additional theories that must be complemented by researchers' questions and attention focus. In their attempt to be helpful, they failed and "[t]he failures have taught us the utopia of neutral description, of fixing stable units for translation, of neat divisions, of excluding interpretation, of studying translation in a vaccum." (Hermans 1999: 71)

From the viewpoint of Munday (2000: 111-117), Toury is placed at the heart of system theories, though in 1980 he would only present his attempt to undertake his "(In) Search for a Theory of Translation" discussing the key concept of *tertium comparationis*, which was later replaced by the concepts of *replaced and replacing segments* in his "Descriptive Translation Studies and Beyond" (1995).

Nevertheless, Munday (2001: 109) does not disregard polysystem theory, which was developed by Even-Zohar in the 1970s and mainly applied to literature. Any piece of work is not be studied on its own, because it maintains relations with all the systems that are comprehended under the name of *polysystem*, in which:

there is an ongoing dynamic and mutation and struggle for the primary position (…) [this is] a dynamic process of evolution (…) between innovatory and conservative systems [which] are in a constant state of flux and competition. (Munday 2001: 109)

Within this flux, translation can be either primary or secondary: if it occupies the former position, then translation is active in the center of the polysystem and may lead to new models and new forces; while if part of the latter, translation will be in a peripheral situation, described as conservative, "preserving conventional forms and conforming to (…) norms of the target system" (Munday 2001: 110). The position that translation holds in the polysystem will naturally and indubitably influence translators' strategies, as we hope to demonstrate.

Shuttleworth & Cowie (1997: 176) summarize the overall definition of polysystem in the following way:

> The polysystem is conceived as a heterogeneous, hierarchized conglomerate (or system) of systems which interact to bring about an ongoing, dynamic process of evolution within the polysystem as a whole. (Shuttleworth & Cowie 1997: 176)

In Toury's work (1995), unquestionably inserted in the descriptive paradigm, norms appear "as the first level of abstraction and the first step towards an explanation of choices and decisions which translators make" (Hermans 1999: 79). This descriptive approach is, according to Toury (1995: 1), of paramount importance for "[d]escribing, explaining and predicting phenomena pertaining to its object level", allowing also for the realization of "more significant studies" and "the elaboration of *applications* of the discipline". Consequently, from the point of view of Toury (1995: 3), the biggest advantage in a descriptive method is the fact that:

> What is missing (…) is but a systematic branch proceeding from clear assumptions and armed with a methodology and research techniques made as explicit as possible and justified within Translation Studies itself. Only a branch of this kind can ensure that the findings of individual studies will be intersubjectively testable and comparable, and the studies themselves replicable, at least in principle, thus facilitating an ordered accumulation of knowledge. (Toury 1995: 3)

In order to justify his descriptive approach, Toury (1995: 9-10) refers to Holmes and his TS map, according to which TS is divided into Pure and Applied TS, being that

the former include Theoretical and Descriptive Studies and the latter discuss issues concerned with translator training, translation aids and translation criticism. Thus Descriptive Translation Studies include product-oriented, process-oriented and function-oriented studies, as will be explained further, and Toury (1995: 11) believes it is essential to have a grasp of "the interdependencies of all three aspects if we are ever to gain true insight into the intricacies of translational phenomena", a position which is highly arguable, as we will try to explain below.

Another point to mention is that TS, according to this descriptive approach (Toury 1995: 15), should deal with three types of questions: the subjects that translation talks about – the theory of translation; what it involves and the circumstances and reasons for this choice – comprehensive description and explanations; and what is likely to involve – the possibility of formulating predictions and reaching more elaborate theory, i.e. theory followed by description and explanation would lead to more theory.

It is in this context that Toury (1995: 16) states the value of "coherent laws" formulated in the course of descriptive studies research based on regularities of behaviour and supposed to reveal the relations established between all variables relevant to translation.

Toury's approach (cit. Hermans 1999: 75) is thus essentially behaviorist towards translation. The mentioned regularities found in translation can either be attributed to translators' conduct or to external contrainsts, both influencing translators' options. The explanation of these constraints resides in what Toury (1995) presents as *norms*, i.e. "performance instructions", since performance consists of translators' option, whereas competence refers to their set of options, emphasizing Levý's decision-making perspective.

Hence, Toury (1995: 53-54) considers *translation norms* to bear a social and cultural function in the context of the target language and culture. If translation is subject to a large number of constraints, such as the source text, differences between languages and texts within the translation process and the constitution of translators' mental apparatus, then these constraints could be categorised in "relatively absolute rules" and "pure idiosyncrasies", between which we can find a large number of other factors that are to be called norms.

There are various definitions of 'norms' according to subject areas or the authors discussing them, for example:

[Norms are] general values or ideas shared by a community – as to what is right or wrong, adequate or inadequate – [turned] into performance instructions appropriate for and applicable to particular situations specifying what is prescribed and forbidden as well as what is tolerated and permitted in a certain behavioural dimension. (Toury 1995: 55)

Toury (1995: 56-57) also conveys the idea that norms are acquired throughout the socialization process and allow people's conduct to be evaluated, leading or not to sanctions, and, in the end, regulate their behaviour. As a result, translation should also be seen as a norm-oriented activity that involves at least two languages and thus two sets of different norms. In this process, there is an effort in subscribing one set of norms (usually of the source culture) to the other set of norms (of the target culture), which may involve incompabilities between the two, leading to a hanging balance between adequacy and acceptability. A consequence, already regarded as a universal of translation, is the inevitable occurrence of translation shifts (as briefly referred to above).

Subsequently, Toury (1195: 58-59) speaks of two kinds of translation norms that affect the entire process of translating: preliminary norms and operational norms. The former relate to translation policy (e.g. the choice of text-types) and directness of translation (the issue of direct and indirect translation), whereas the latter are aimed at the decisions made throughout the process of translation, encompassing matricial norms (those that raise questions such as: is the target text complete when compared to the source text; is it in the right place?; what kind of ommissions, additions or changes – i.e. textual segmentation – were carried out?) and textual-linguistic norms (comprehending the selection of material for the target text and the replacement of the original material).

On the other hand, also described as descriptivist, Chesterman (Munday 2001: 118) states that "all norms exert a prescriptive pressure". He then presents another set of norms that include Toury's initial and operational norms, which are product or expectancy norms and professional norms. The first ones are related to readers' expectations towards what translation will be like and are determined by translation tradition in the target culture, discourse conventions and economic and ideological considerations. Expectancy norms lead to "evaluative judgements about translations (…) and are sometimes validated by a norm-authority of some kind" (Munday 2001: 118-119), although there may be a difference in the acceptance emanated from this authority and from society. They are regarded as "constitutive norms [because] if

translators abide by them, then their products will be classified as (genuine, proper, legitimate) translations" (Hermans 1999: 78).

On the other hand, *professional norms* that regulate the translation process are dependent on the expectancy norms and include the accountabilty norm (of an ethical nature that tackles with "professional standards of integrity and thoroughness" (Munday 2001: 119) and with translators' responsibility over their translations), the communication norm (a social norm, since translators as communication experts wish to ensure understanding from the parties involved) and the relation norm (a linguistic norm that deals with the relation between source text and target text).

Chesterman (cit. Hermans 1999: 77) discusses social, ethical and technical norms in action within translation: the first regulate interpersonal coordination; the second represent the need for translators to "uphold the values of clarity, truth, trust and understanding" (Hermans 1999: 77); the last ones are divided into product and process norms that are also related to these four values.

Furthermore, Nord's constitutive and regulative conventions are considered by Hermans (1999: 79) as clearer. The first set of conventions determines the ideia of translation that a particular community holds, the sum of which will bring about the general concept of translation, whereas the last set of conventions oversee the ways of handling with translation problems below the text level. Both Nord and Chesterman's norms are a step further on Toury's norms because not only do they bring other perspectives into scene, but "they [also] allow us to conceptualize a domain of translation, or a translation tradition (…) and to think about ways of describing its boundaries" (Hermans 1999: 79).

Although norms are seen in the descriptive paradigm "as the first level of abstraction and the first step towards an explanation of choices and decisions which translators make" (Hermans 1999: 79), which is the course of action we intend to take in this research, they are not only constraints on translators, but they also function as templates, ways of solving the problems they encounter in their work. Norms demonstrate regularities, patterns of behaviour – they are "a psychological and social entity" (Hermans 1999: 80). They mediate between the individual – intentions, choices and actions – and the collective – beliefs, values and preferences; they allow for stability and the prediction of behaviour by reducing uncertainty, by regulating. Because translation is a communication act, transational norms and conventions "guide and facilitate decision-making" (Hermans 1999: 80).

Several scholars (for example, Lewis 1975; Fokkema 1989; Hermans 1999) distinguish between norms and conventions, since these last ones may become norms if they turn out to be successful, whereas norms are "stronger, prescriptive versions of social conventions" (Hermans 1999: 81), they are directive. When a translation follows a set of translational norms and conventions of a certain society, it means that it "conform[s] to the relevant correctness notion (…), being that this [c]orrectness in translation is relative – linguistically, socially, politically, ideologically" (Hermans 1999: 85).

Consequently, one may ask what the point of studying norms and conventions in translation is and how it is possible to achieve this within the descriptive paradigm. Again Hermans (1999: 85) summarises the most important aspects to consider in conducting such a study: first of all translation is a social and communicative practice; then, if norms are not observable (since norms in action are different from their formulation), one must make use of "likely sources" according to Nord (1991) and Toury's (1995) terminology. These are as follows: translations themselves; various translations of the same text; bibliographies of translations; paratexts and metatexts; "statements and comments by translators, editors, publishers, readers and collectives such as translator's associations" (Hermans 1999: 85); reviews and criticisms; theroretical and programmatic statements; the activity of schools of translation; textbooks in translation training; lawsuits and copyright law; distribution information; and translation prizes.

Recalling Holmes and Toury's division of TS and trying to make the connection between TS and AVT, Mayoral Asensio (in Duro 2001: 25-37) should be mentioned when he states that AVT is susceptible of being studied through *product-centered approaches* or *process-centered approaches*, two of the three branches already explained above.

Concerning the first ones, Mayoral Asensio (in Duro 2001: 25-30) mentions semiotic or semiological studies that focus on the language of images disconnected from verbal language and analyzes the role played by the different non-verbal narrative elements. In addition, there are those that look at culture within the translation product, that is to say according to the theories of manipulation, polysystem, post-colonialism or the translator' visibility – translation is definitely a cultural product. Apart from these, there is also a growing concern about the normalization of procedures and style, because of the need for homogenization of AVT procedures to ensure basic quality standards,

since experience shows that there is a variety of 'incorrect' forms that are introduced into general language. At last, it is worth referring to sociological and historical aspects, represented by research conducted, for example, by Luyken *et al*. (1991), on the one hand, and Ballester (1999 and 2001), Gottlieb (1997) or Díaz-Cintas (1997 and 2000), on the other.

As for the second type of approaches, these comprehend communicative studies, those that pay attention to the linguistic effects of synchronization (the narrowing of meanings and the elimination of those that are redundant or secondary; the synthesized ot telegraphic style for which the comparison between languages is essential) or to psycholinguistic issues in subtitling or dubbing (type of spectators, action or film projection contingencies, kinesic aspects, the speed of dubbers' diction or of the subtitles, their repercussions on listeners or viewers, the format of the subtitles, among others), and professional studies that intend to discuss and solve specific problems in the professional activity, such as the description of different processes of AVT, the localization of multimedia products, the translation of linguistic variation (e.g. idiolects, sociolects or dialects) or humorous or cultural aspects (Mayoral Asensio in Duro 2001: 30-33).

Finally, presented as a final course of research by Mayoral Asensio (in Duro 2001: 37) or Toury's third functional approach, it is also worth mentioning the instrumental application of AVT to different fields, such as subtitling used for the teaching of a foreign language and the development of viewers' reading literacy in their own language or even subtitling for the deaf and hard-of-hearing.

As far as fixed language is concerned, on the other side of our research, we will start from the general standpoint of Linguistics, (specifically that of Semantics) that deals with the canon of compositionally (distinguishing the linguistic items – free and restricted word combinations – we wish to study), and gradually enter the approaches of Lexicography and Phraseology that deal with the issue of phraseology and idiomaticity.

The question of idiomaticity in TS stands out because it relates to "a translation strategy which aims for a TT which reads as naturally as possible" (Shuttleworth & Cowie 1997: 72), because meaning is retained above form. This is what Beekman & Callow (1974) and Larson (1984) refer as Idiomatic Translation or the Idiomatic Approach. In order to achieve this, it is necessary to translate by means of "careful linguistic reformulation and paraphrase (…), [as well as to pay] close attention to the need to make explicit to the target readers information which (…) was generally

available to the source audience and thus only implicitly contained in ST" (Shuttleworth & Cowie 1997: 73).

Consequently, by studying idiomaticity in linguistic terms and its manifestations in different levels of phraseology, we intend to ascertain the degree of overall idiomaticity in the target texts of the documentary that functions as our pilot study.

Among the linguistic theories that we selected, we will review a number of perspectives, such as that of Saussure, Mel'čuk, Zgusta and Lyons, to name only a few, so as to reach a general understanding of the complexity of the topic of fixed language. The authors that have been mentioned present their own designations, which can differ from the others, though they may share some aspects. These can be, for example, summarised in a list of criteria like the one put forward by Zgusta (1971), attempting to determine which word combinations are free and which are restricted.

To sum up, our research will be placed within the paradigm of Descriptive TS, attempting to approach AVT through the comparison of the target text to the source text, hoping to find regularities and patterns in audiovisual translators' outputs (in subtitled and voiced-over versions) and reach the formulation of some norms or, at the very least, some conventions. In order to shed light on part of the constraints on translators when doing subtitling and later revoicing, a summary of the few written Portuguese standards for subtitling and their constrast with international standards shall be a significant part of future research.

Moreover, we will try to demonstrate how we achieved the combination of these two areas of knowledge by approaching AVT not only from a product-centered perspective, stressing the ideas of the relationships between source and target cultures and the importance of normalization of standards and norms in AVT , but also from a process-centered standpoint (if feasible), especially the effects of synchronization on the end-product in linguistic terms, and eventually from a professional point of view, so as to highlight the influence it has over viewers' reading literacy.

3. Fixed language

3.1. Compositionality

An aspect that must be taken into account when discussing language, either figurative or literal, is *the canon of compositionality*, according to which phrases and sentences result from the sum of separate units of meaning. The global meaning of a linguistic expression is the result of the sum of the meanings of each of the elements that make up this same expression, thus its sense being compositional.

Contrary to this type of language, figurative language is viewed as special because of the impossibility of imposing a process of compositionality upon the words that integrate their expressions. Proverbs and idioms are examples of a non-compositional meaning (Hoffman & Honeck 1980: 8-9). Nevertheless, Iriarte Sanromán (2001: 126) mentions that even morphologically regular derivates can bring about problems in the understanding of their meaning, such as *adorável* or *considerável* that do not match the immediate sense of 'that you can adore' or 'that you can consider', but rather demand the user to go beyond the mere sum of the meanings of its morphemes or affixes.

As a means of illustrating the unpredictability of any natural language, the German philosopher and logicist Frege (cit. Fromkin 2000: 374-375) mentioned that:

> It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the thought. (Frege cit. Fromkin 2000: 374-375)

For Frege, the understanding of a sentence comes from the comprehension of its parts and their combination within the structure of a sentence, allowing us to recognize the meaning of the familiar elements and the usual ways of combining them in sentences that have never been read or heard. Therefore, the principle of *semantic compositionality* consists of the process of progressive construction of meanings from the morpheme to the sentence itself and of the relationship that these meanings establish among themselves.

However, this principle of compositionality is not always respected in the linguistic constructions realized by speakers, as in the case of idiomatic expressions, since their interpretation is not dependent on the meaning of their parts. These expressions are usually non-compositional because their constituent parts are not real semantic elements or are not relevant for the global meaning of the expression or their meaning cannot be inferred through a compositional process (Curse 2000: 74).

To sum up, Hudson (1999: 273-276) presents another approach to compositionality, which includes *linear compositionality*, *non-linear compositionality* and *non-compositionality*. The first type of compositionality is obtained when the meaning of sentences matches the sum of the meanings of their parts; the second refers to the cases in which the elements of the phrases are discontinuous, i.e. these are separated ones from the others by other words or phrases (e.g. "A guy, who is at the door, wants to speak to you", in which the relative pronoun violates the compositional meaning of the sentence because it interferes in the information to convey and makes its understanding more difficult); the last is typically represented by figurative expressions, whose meaning can not be reached through the sum of the signifiers of their parts.

3.2. Word combinations

They are made up of smaller elements
(phonemes, graphemes, syllables, morphemes),
and they are embedded in larger co-texts
(phrases, clauses, sentences, paragraphs) which in
turn are part of the wider extralinguistic context
of speech acts and genres. (Hartmann 1983: 8)

According to Vilela (1995: 9-11), the lexical knowledge of a language implies not only the knowledge of morphemes, simple and compound words and their respective meanings, but also of a number of fixed or set phrases, whose meaning can not be inferred from the meaning of their constituents, because it is non-compositional. Their importance derives from the fact that they are extremely common in any language and they reflect its wealth, which may be paradoxically more metaphorical than literal.

Generally speaking, these expressions tend to reach a certain degree of frozenness, not normally allowing other combinations and preventing the order of their parts from being changed. Additionally, they frequently break the combination rules of

semantic proprieties, possessing syntactic and semantic features of their own. They must then be understood as whole units with specific meanings and restrictions to their occurrence. (Fromkin & Rodman 1993: 197-198)

According to what Saussure (1992: 214) defended, all units of a language depend either on what is around them in speech (co-text) or on the parts that make them up, thus stressing the importance of syntagmatic relationships in the organization of a language. These relationships will inevitably lead to problems of polysemy, to different significances of a word being interpreted according to the senses it acquires when combined with other words. Therefore, the existence of a certain degree of frozenness or setness in language is generally accepted, which is a common characteristic of phraseological expressions, not necessarily implying idiomaticity. In the words of Iriarte Sanromán (2001: 25), the more or less fixed expressions of a language are generally known as phrases and they correspond to specific sequences learnt by heart, lexicalized phrases or lexical combinatory patterns.

Consequently, *lexical combinatorics* means that this range of combinations is understood by speakers as words, which include everything in the spectrum from collocations to idiomatic expressions. Lexical restrictions are much more economical and easier to collect for any type of work, but a more or less fixed combination is not necessarily idiomatic. (Iriarte Sanromán 2001: 28)

"PEOPLE SPEAK IN SET PHRASES – rather than in separate words; hence the importance of set phrases" (Mel'čuk 1998: 1). Thus, *co-occurrence* or *lexical co-occurrence* is the capacity for lexical units to combine themselves into phrases, syntactic and lexical expressions, and to convey a certain meaning, based on the principle of structuralist linguistics according to which linguistic units never work as separate phenomena, but rather establish a relation of interdependence within a whole which is called a structure (Iriarte Sanromán 2001: 117).

This co-occurrence may be *free* or *restricted*: it is free when this combination is done following the grammatical rules of a language (free phrases), whereas it is restricted when the combination happens with two or three lexemes in accordance with semantic and syntactic rules and some kind of purely lexical restriction (set phrases or phrasemes).

As for Mel'čuk (1995: 175), a *free phrase* A + B consists of two lexemes, which are regularly constructed so as to express a certain concept, it can be replaced by another synonymic lexical expression and its signifier is understood out of the sum of

the signifiers of A and B. Conversely, *non-free* or *restricted lexical combinations* can be divided into pragmatic phrasemes or pragmathemes and semantic phrasemes, the latter also encompassing complete phrasemes or idiomatic expressions, semi-phrasemes or collocations and quasi-phrasemes.

Zgusta (1971: 142-151) also discusses the issue of *set combinations* (referred to as multiword lexical units), presenting a number of criteria for distinguishing them from free combinations: substitution is impossible; addition of other words is frequently impossible; the meaning of the whole is not derivative from the meaning of the single constituent parts; a synonym or near-synonym may exist, consisting of only one word; a small group of expressions may be related and have an analogous status; a one-word equivalent in a foreign language may suggest that it is a multiword lexical word; they may show special formal and grammatical properties, like the absence of articles. Nevertheless, even conforming to most of these criteria, there are some combinations of words that are not set phrases, because they do not perform the same syntactic and onomasiological function as a morphologically simple unit at the syntagmatic and paradigmatic levels.

In the perspective of this author (Zgusta 1971: 154-155), this last criterion is of crucial importance to distinguish between multiword lexical units (mostly idiomatic expressions) from other set groups of words, such as proverbs, sayings, quotations, similar fossilized or petrified expressions. Moreover, there are different degrees of *setness* or of restrictions that can be extremely useful when comparing examples like 'light burden' and 'light supper', i.e. in the latter, the combinatory possibilities are more restricted. Thus, the more severe the restrictions imposed on word combinations are, the more "set" these combinations are.

Returning to Mel'čuk's (1995: 181) concept of *semantic phraseme*, this is the combination of two or more lexemes, which signified is the regular sum of the signifieds of its lexemes, while its signifier is different from the sum of their meanings. The meaning in a semantic phraseme is freely chosen; it is not imposed by the situation, contrary to what occurs with pragmatic phrasemes. The lexical selection of its meaning is partially or totally limited, even if it may be a regular construction in morphological and syntactic terms.

Within this type of phrasemes, we can have *complete phrasemes* or *idiomatic expressions*. According to Alonso Ramos (1993: 182), they are semantically non-compositional (their meaning is not equal to the sum of the meanings of their

constituents) and coherent (their elements are required ones by the others); they resist to formal variation; they can be ambiguous and bring about problems when they are to be dealt with by linguistic models.

Apart from these, there are also *semi-phrasemes* or *collocations*. For Mel'čuk (1995: 46), their meaning is X, which does not match the sum of the meanings of their elements; they equal the signifier of A plus a lexeme B that expresses C, not freely selected; outside the combination of AB, B would not mean C. Because of this, they are not free combinations of lexemes, but frequent, probable, preferential or usual combinations of lexemes (namely [noun + adjective] or [verb + noun]), as well as apparently free combinations created according to the rules of a language where some type of lexical restriction determined by these rules is to be found. However, frequency should not be the only nor the most important criterion for the identification of collocations, others should be considered.

The third type consists of *quasi-phrasemes* which preserve the meaning of the lexemes that make them up, plus an additional sense that is not deduced from the sum of their elements, thus creating a lexicalized whole such as the idiomatic expressions. (Mel'čuk 1995: 46)

Finally, *pragmatic phrasemes* or *pragmathemes* are structures whose meaning is not freely built from a specific conceptual representation, though they may be regular, i.e. their meaning cannot be replaced by any other meaning. (Mel'čuk 1995: 179)

The most important type of pragmatic phrasemes are *routine formulae*, also known as conversational formulae or those used to realize speech acts, which are viewed as units for habitual and stereotypical social interaction that accomplish specific functions in ritualized situations. These formulae include: *discourse formulae*, such as those for opening and closing conversations or for turn-taking; and *psycho-social formulae* that comprehend attitudinal-expressive, attitudinal-commissive, attitudinal-directive, assertive, ritual and miscellaneous formulae. Even *proverbs* can be understood as pragmathemes according to certain authors, such as Corpas Pastor (1995: 354-378).

Co-occurrence can co-occur with another designation – (*lexical*) *solidarity* – which according to Coseriu (1977 and 1979) and also Vilela (1979 and 1994), consists of the relationship among lexemes of different fields, in which one is partially or totally included in the other, as a seme that limits the ability of their combination. Consequently, lexical solidarity is more restricted than collocation, whilst lexical

solidarities are a type of collocation. Analogous concepts are 'entourage' and 'contorno' according to Rey-Debove (1971), Seco (1987) and Corpas Pastor (1995) that shall be discussed further in the context of the thesis.

This discussion of fixed language, lexical solidarity and co-occurrence has taken linguists from various schools to set forth a myriad of designations for free lexical combinations and non-free lexical combinations, which are summarized in the table below.

Table 1. Summary of different proposals for free and non-free word combinations.

Authors	Free lexical combinations	Non-free lexical combinations
Coseriu (1977)	Discourse technique	Repeated discourse
Pottier (1978) and Fonseca (1981)	Simple *lexia*	Complex *lexia*
Herculano de Carvalho (1979)	*Sintagma livre* (Free phrase)	*Sintagma fixo* (Fixed phrase)
Lyons (1995)	Lexically composite phrasal expressions	Lexically simple phrasal expressions
Sinclair (1995)	Open choice principle	Idiom principle
Mel'čuk	Free phrases	Set phrases or phrasemes
Bosque & Demonte (2000)	Non-phrasal lexemes	Phrasal lexemes or syntagmatic compounds

Even Saussure (1992: 210-211) recognizes that it is extremely difficult to say what a free combination of words is, to distinguish between a language fact, which results from a collective use, and a speech act, which is dependent on the free will of a speaker. Lyons (1995: 52) corroborates this idea, by stating that the frontier between *composite phrasal expressions*, created by a process of grammaticalization, and *simple phrasal expressions* or lexicalized expressions, is not as straightforward as it appears.

For Lyons (1985: 145), not all lexemes correspond to a single word, because we can find numerous *phrasal lexemes* (or 'syntagmatic compounds', according to Bosque & Demonte 2000) in any language, "lexemes whose forms are phrases, in the traditional sense of the word" (Lyons 1985: 145). These phrasal lexemes tend to be grammatically or semantically idiomatic, because their distribution in sentences or their meaning is not predictable from the syntactic or semantic characteristics of their parts (e.g. red herring). When a phrasal lexeme has a non-phrasal lexeme as its equivalent, we say that the former has an idiomatic, metaphorical or figurative meaning, whereas the latter bears a

literal meaning. However, Lyons (1985: 146) believes that there is no generally accepted criterion to distinguish these phrasal lexemes from clichés or fixed collocations.

Another author, Carvalho (1979: 495-496) refers to a gradation between phrase and word which are distinguished in three levels: a *free phrase*, a *fixed phrase* and a *compound word*. A fixed phrase is defined as "uma associação de palavras em sequência fixa, que constitui uma unidade sintáctica perfeita (...) e também muitas vezes semântica (...) e morfológica". Therefore, the difference between a fixed phrase and a free phrase is that the former means a simple concept and its sense is non-compositional, whereas the latter signifies a complex one and is usually compositional. Apart from this, fixed phrases are also characterized by: the rigid order of their elements; the impossibility to change them; the possibility of belonging to different word classes; the reduced inflectional variation; the likelihood of having abnormal syntactic constructions; the syntactic functioning as a word only; and the integration within a paradigmatic lexical structure. Nevertheless, distinguishing between a fixed phrase and a compound is even harder, since a compound can have only one stressed syllable and may be made up of bound forms.

On the other hand, Pottier (1978 269-270) refers to three types of lexias: the simple lexia, the compound lexia (equivalent to a compound) and the complex lexia (e.g. cold war); apart from these, he also refers to textual lexias, such as proverbs. This author establishes criteria for identifying lexias, though considering them to be insufficient, which are: the lack of autonomy of their elements, the impossibility of partially recovering their reference, of replacement and of separation of an element, and the absence of determination.

Similarly, Fonseca (1981: 90-97) also follows this perspective, speaking of simple, compound and complex lexias: the first ones match those inseparable morphemes, while the other two will correspond to lexical and syntagmatic compounds, respectively, mentioned further on. Hence, complex lexias are units learned by heart as part of group competences, susceptible of being actualized in discourse whenever necessary and that are the basis for the syntactic and syntagmatic construction (e.g. rainbow). At last, complex lexias achieve such a level of combinatory cohesion that their meaning goes beyond the mere literal reading, such as the cases of *círculo vicioso*, *estar à vontade* and *perder o juízo*. Fonseca (1981: 97) suggests also the use of a test of

embedment to check if a certain compound constitutes or not a compound: 'inegável círculo vicioso' → '*círculo inegável vicioso'.

Finally, Bergenholtz & Tarp (1995: 117-118) refer to three types of word combinations: maxims, idioms and collocations. The first function as a whole, semantically and in terms of usage, in which the parts cannot be replaced by other possibilities; the second are "relatively fixed phrase[s] with only limited possibilities of syntactic variation" with a non-compositional meaning; and the last are made up of two or more words whose meanings may lead to the meaning of the whole unit.

3.3. Collocations

> "estereotipos" cuya regularidad, vinculación a un fenómeno o acontecimiento de la realidad, así como su frecuencia de uso le han otorgado cierta estabilidad y han hecho que la lengua las privilegie como unidades aislables y reproducibles. (Irsula 1992: 160)

From Benveniste's (1989: 175-176) standpoint, lexical combinations that are not totally free, i.e. restricted, are *collocations* or *sinapsias* that impose certain restrictions in terms of the words with which they are combined and their meaning is semantically compositional. Thus, these aspects will provide us with information about their level of frozenness, but not necessarily about their level of idiomaticity, since one is not to be misunderstood with the other. According to Iriarte Sanromán (2001: 139, 160), collocations can be more or less idiomatic, more or less transparent or opaque, bearing figurative (or improper) meanings which are only actualized when in combination with other words.

In Wierzbicka's (1985: 57) point of view, collocations are characterized according to their economy, elasticity and ability to adapt to new situations and new conceptualizations of the language. But, contrary to what is advocated by authors like Kjellmer (1994), the frequency with which certain words combine with others can not be seen as a feature of collocations, because this does not have any influence over the lexical solidarity established among the words. This relationship depends, instead, on the characteristics of the things they refer to. A restricted lexical combination is not

equivalent to a frequent combination of two or more lexemes (Iriarte Sanromán 2001: 162-163).

As we have showen above, there are tests that can be carried out to some collocations in order to analyze their higher or lower degree of freedom of combination: the use of the passive voice, of the participle adjective, of the relative clause or of the pronoun; the change of adjective, of noun or of adverb; determination; quantification; the use of the indefinite or of the plural; and the presence or absence of the article. However, it is impossible to establish rules for the syntactic behavior of all collocations, even with the use of these tests.

Collocations consist of "a phenomenon of lexical combinatorics: they involve lexical, semantic and syntactic properties of lexical items and their syntagmatic co-occurrence" (Heid in Wright & Budin 2001: 788). It is obvious that collocations pose great problems to translating even if they are fairly transparent, especially because they are "a matter of convention rather than being explicitly rule-governed" (Heid in Wright & Budin 2001: 788).

Additionally, Heid (in Wright & Budin 2001: 788-89) presents a number of principles that are generally accepted as uncontroversial: collocations involve two lexemes, not including determiners, quantifiers or prepositions that might also occur, which is a first major distinction towards multiword lexemes, such as idioms; the constituents of these collocations may be collocations themselves; collocations can be categorized according to the word classes of their elements into [noun + verb], [noun + adjective], [noun + noun], [verb + adverb] and [adjective + adverb]; collocations are polar, since one of the elements is determined (base), whereas the other determines (collocate).

We must be aware that not all chunks are accepted as being types of collocations, for instance subject/verb collocations cannot be included in this realm. In the choice of the components for noun phrases and noun compounds, there are usually combinatory preferences of conventional nature, meaning that "a given collocate with a given base being an arbitrary phenomenon (…) must be memorized" (Heid in Wright & Budin 2001: 790-93). That is why it is highly difficult to draw a line between collocations and multi-word lexemes. Again the conclusion is that there are no rules that govern collocations; they are rather idiosyncratic and must be learned by heart when learning a language, be it our mother-tongue or a foreign language.

To sum up, Heid (in Wright & Budin 2001: 791) regards collocations as being dependent on lexical factors (idiomatic or conventional preferences) and conceptual factors subtypes of concepts or typical properties, usages and actions, and partial compositionality is also an important criterion to bear in mind, i.e. collocations can be partially predictable.

3.4. Idiomatic Expressions

The word 'idiom' comes from the Greek 'idios', which means own, private or peculiar, with time assuming the meaning of any expression of peculiarity to be filled with new concepts. According to the *Oxford English Dictionary* (1998 vol. VII: 624-626), an idiom consists of:

> (1) the form of speech peculiar or proper to a people or country; own language or tongues; (2) the specific character, property or genius of any language, the manner of expression which is natural or peculiar to it; (3) a form of expression, grammatical construction, phrase, etc., peculiar to a language; a peculiarity of phraseology approved by the usage of a language and often having a signification other than its grammatical or logical one. (OED 1998 vol. VII: 624-626)

Consequently, an idiom is an expression that obeys to certain criteria, namely its fixed nature, its non-literal and sometimes metaphorical usage and the fact that is recognized as such by the speakers of a language. It is a complete expression, though there is an extensive number of single words that can also function in an idiomatic way: 'I *was over the moon* when I heard she got married'; 'It *broke my mother's heart* to see her home burn to the ground'; 'The river *flooded* several villages' versus 'The crowd *flooded* on to the pitch' (Wright 1999: 7-10). From the reading of these sentences, we can conclude that some are perfectly transparent, allowing for the creation of images that facilitate their meaning to the user (for example, 'break someone's heart' or 'flood onto'), and the others are considerably opaque, preventing the user to apprehend their sense (such as, 'to be over the moon).

To exemplify further, in French, one speaks of *expressions figées* (frozen expressions) and the process by which they are created *figement* (frozenness), defined in the *Grand Larousse Universel* (1995 vol. 6: 4256) as follows:

Le figement se caractérise par la perte du sens propre des éléments constituants le syntagme, qui apparaît alors comme une nouvelle unité lexicale, autonome et à sens complet, indépendant de ses composantes (…) Le figement joue un rôle important dans l'évolution linguistique. (*Grand Larousse Universel* 1995 vol. 6: 4256)

On the other hand, in Spanish, according to the *Diccionario Enciclopédico Espasa* (1996 vol. 6: 6487a), this type of expressions is identified by the name of *idiotismo*, that "expresión o sintagma privativo de una lengua, de forma fija y por las reglas de la gramática" and an *idiomatism* as a "rasgo lingüístico peculiar y característico de cada idioma". Similarly, the *Diccionario Enciclopédico Universal* (1986 vol. 4: 1484) defines *idiotismo* as a *modismo*, a way of speaking that goes against the grammar rules, but is particular of a language.

Finally, in Portuguese, as can be read in *Enciclopédia Verbo Luso-Brasileira de Cultura* 2000, vol. 15: 387), an idiom (which can also be designated as *idiotismo* or *idiomatismo*) consists of a phrase, a way of saying or a lexical or grammatical construction that characterizes a language when compared with others, bringing about difficulties in their translation into analogous structures between languages. Idiomatic expressions are frozen or crystallized sequences that are sometimes difficult to differentiate from some type of compounds, these being lexemes that result from the word formation process called compounding (Mateus *et al* 1994: 392), contrary to what other scholars, such as Bosque & Demonte (2000), refer to as composites.

It is worth focussing briefly on the process of compounding, according to which we can have *lexical compounds* and syntagmatic compounds. In Bosque & Demonte's (2000: 4761) view, the first ones maintain their lexical integrity and consist of a phonological blending of their constituents, having only one accent, and morphological units, also called 'syntactic islands', because they cannot establish syntactic relations with the rest of the elements of the phrase or sentence they belong to.

Table 2. Examples of lexical compounds (Bosque & Demonte 2000: 4763).

carta bomba:		mujer objeto:	
a.	cartas bomba	a.	* mujer objeto digno de estudio
b.	* cartas bombas	b.	mujer objeto digna de estudio
c.	?? carta con-bomba		

The other ones – *syntagmatic compounds* – are defined by their syntactic frozenness, which can be easily understood by the examples *fin de semana* ('weekend' in English), which is different from *fin de essa semana* ('the end of the week' in English). In these constructions, the frozenness that defines them does not change their stress nor their syntactic structure. Syntactic frozenness implies that the inflection of the phrase be dependent on its nucleus, that it does not allow for the addition of modifiers to the prepositional phrase or of adverbs to the adjective and that it be impossible to partially commute its constituents without altering the nature of the phrase. Thus, the frozenness of the syntactic properties is a sine qua non condition for the existence of this type of compounds. (Bosque & Demonte 2000: 4763-4764)

As other authors have done it (for instance, Zgusta 1971 or Fonseca 1981), Bosque & Demonte (2000: 4764) put forth a number of tests that enable to say whether a certain compound has reached such a level of frozenness as to be considered a syntagmatic compound or to be regarded only as a lexical compound.

Table 3. Examples of syntagmatic compounds (Bosque & Demonte 2000: 4764).

fin de semana:		el orden del día:	
a.	fines de semana	a.	los órdenes del día
b.	*fines de semanas	b.	*los órdenes de los días
c.	*fin de semana de vacaciones	c.	*el orden de un día
d.	*fin de esa semana	d.	*el orden del día del fiesta
e.	*término de semana	e.	*el orden diario

After analyzing these examples, it is possible to conclude that syntagmatic compounds comprehend a range of different constructions, which can do without a verb, are fixed syntactic structures, encompass whole concepts, break the compositional principle and the greater resistance to cohesion they offer, the higher their motivation or their semantic transparence turns out to be. Each constituent of these compounds does not possess syntactic autonomy, because they are part of a unitary whole.

When dealing with phrases, it is also of the highly significant to determine their nucleus, since it is in the nucleus that the possibilities of distribution for these constructions are based, influencing their morphological and syntactic category. As a result, we can have exocentric or endocentric compounds.

On the one hand, *exocentric constructions* have an idiomatic sense that presents a high degree of compositional opacity – they are demotivated constructions and are

similar to metaphors, the lower their compositionality, the higher the cohesion among their elements. On the other, *endocentric constructions* possess a nucleus and present a morphological structure that reflects the semantic relations established among their parts. (Bosque & Demonte 2000: 4783-4788)

Idioms are viewed as polylexemic expressions, opposed to polymorphemic expressions, and several scholars (such as the ones previously mentioned) approach them as a linguistic issue both interesting and problematic, justifying the number of studies that can be found on this phenomenon, which for Chomsky (cit. Strässler 1982: 21) represents "a basic stumbling block".

At this moment, we shall present a number of different approaches to idiomatic expressions, so as to show the myriad of designations and theories that populate this area of study and to attempt to reach some type of conclusion.

For Strässler (1982: 11, 15-16), idioms are "a special category of lexical items which are not only determined through their structure, but which also show a specific type of behaviour in language use", thus being a functional element of language. There are different levels of idiomaticity that are not considered as idiomatic by all groups of scholars, allowing people to be aware of the problems brought about by this complex issue, such as sayings, proverbs, phrasal verbs, prepositional verbs, tournure idioms, binomials, frozen similes, ungrammatical, but generally accepted expressions, logical connective prepositional phrases, phrasal compounds, incorporating verb idioms, and formula expressions.

By approaching idioms from a pragmatic standpoint, Strässler (cit. Fernando 1996: 13) refers to the social implications of their use, which depends on a considerable number of social variables, such as social status, age, education and profession of their users. "When using an idiom, the speaker conveys more information than its semantic content (…) [establishing] a social hierarchy or [testing] the hearer's opinion in this matter" (Strässler cit. Fernando 1996: 14), because he makes use of idioms in a deictic manner.

The deictic use of idioms comprises first person idioms, second person idioms and third person idioms. The *first and second persons idioms* are usually marked; they are the social deixis because they allow for the establishment of social relationships – the first person idioms are used among people of the same social status, but avoided by dominant speakers, and the second person idioms are only accepted among peers. At

last, the *third person idioms* are described as being non-marked and neutral and bear no restrictions to speakers of different social status. (Strässler cit. Fernando 1996: 14-15)

Therefore, idioms work as status markers and their use in a conversation consists of a way of showing membership – this is another difference of idioms towards their literal counterparts, which don't convey this pragmatic role (Strässler cit. Fernando 1996: 14-15).

Weinrich (1969 cit. Fernando 1996: 6-7) not only considers idioms as both universal and specific to each language, but also believes phraseology to be a branch of lexicology. His approach is developed under the influence of Soviet phraseologists and the generative-transformational grammar, thus being extremely formal. According to this author, only some phraseological expressions are to be considered idioms and he distinguishes between the idiomaticity of phrases and the stability of the collocations. Both constructions reflect the phenomenon of co-occurrence (see 2.1) in literal and non-literal contexts; however the co-occurrence of idioms is brought about by a particular semantic relation that is not present in collocations.

As a consequence, Weinrich defines idiom in the following way: "A phraseological unit involving at least two polysemous constituents, and in which there is a reciprocal contextual selection of subsenses will be called an idiom" (Weinrich cit. Fernando 1996: 7-8).

The Soviet phraseologists, such as Vinogradov (1977), understand phraseological units and lexemes as units of language alike, differing only in their structure. This author categorizes the former into: completely unmotivated phraseological collocations; completely motivated but in metaphorical re-interpretation unmotivated phraseological units; completely motivated phraseological combinations. For Amasova (1963), phraseology is the theory of collocations which establishes a fixed context and an idiom is a phraseological unit where there is no longer any distinction between a key word and the semantic features that it actualizes.

Reichstein (1973 and 1974 cit. Strässler 1982: 23-24) dropped the connection of phraseology to lexicology and identified two important aspects in phrasal collocations: regularity in usage (or usual) and irregularity in their structural semantic organization. This last one is defined by *idiomaticity* or semantic irregularity and *frozenness* or syntactic irregularity, but idiomaticity is a pre-requisite for frozenness. Later on, he acknowledged different degrees of frozenness – absolute, selective and preferential frozenness – and the difference between perfect phrases and non-perfect phrases, which

led to a classification of 18 categories. On the other hand, Mel'čuk's (cit. Strässler 1982: 26) idiomaticity classifies polylexemic expressions into frozen and idiomatic; frozen and not idiomatic; not frozen and idiomatic; and not frozen and not idiomatic.

As for Mackkai (cit. Fernando 1996: 3-4), idioms can be classified into *idioms of coding* and *idioms of decoding*, in that an idiom is an expression constituted by a minimum number of two lexical independent elements. Those expressions, which consist of only one element and cannot occur in other contexts, are not to be truly seen as idioms. Hence, idioms of decoding can be further organized into *lexemic and sememic idioms*: the former including phrasal verbs (e.g. bring up, get away with), tournures (e.g. fly off the handle, rain cats and dogs), irreversible binomials (e.g. salt and pepper, bag and baggage), phrasal compounds (e.g. blackmail, high-handed), incorporating verbs (e.g. eavesdrop, iceskate) and pseudo-idioms (e.g. kith and kin, spick and span); and the latter encompassing those pragmatic meanings that are already culturally institutionalized, such as proverbs, familiar quotations (e.g. not a mouse stirring), idioms of politeness (e.g. May I...?) and idioms of understatement and hyperbole (e.g. I wasn't too crazy about him) (Makkai cit. Fernando 1996: 5-6).

The difference between these two types of idioms is their function, since sememic idioms clearly play a role in interpersonal relations, which is also culturally determined, by expressing advice, requests, and assessments, among others.

Finally, Chitra Fernando (1996: 30) defines idioms as conventionalized phrases that are characterized by non-compositionality, institutionalization and semantic opacity. These three features are present in a great number of word combinations, such as slang, sayings, metaphors, social formulae, collocations, among others, but idioms are faced as an unbreakable unit, whose elements cannot be changed, unless in very restricted cases, thus showing a particular degree of frozenness.

According to Fernando's approach (1996: 33-37), it is of the utmost importance to have a *scale of idiomaticy* that would allow for the classification of idioms and overcome the criticism posed on some linguists who study idioms, which is the fact of being obsessed with their interpretation and not with their classification. Hence, idioms can be *pure idioms*, non-literal conventionalized phrases that must be understood as a whole; *semi-idioms* that possess one or more literal constituents and at least one with a non-literal meaning, specific of that relation of co-occurrence (some of these are overlapped with restricted collocations that allow some lexical variation); and *literal*

idioms, which accept invariability or a restricted variation and are lexically less complex.

Table 4. Scale of idiomaticity according to Fernando (1996: 32).

IDIOMS	HABITUAL COLLOCATIONS
I. <u>Pure idioms</u> a) Invariant and non-literal (red herring, smell a rat) b) Restricted variance and non-literal (seize/ grasp the needle)	
II. <u>Semi-literal idioms</u> a) Invariant (drop names) b) Restricted variance (good morning/ day)	I. <u>Restricted variance/ Semi-literal</u> (explode a myth/ theory/ notion/ idea/ belief; thin/ flimsy excuse)
III. <u>Literal idioms</u> a) Invariant (on foot, on the contrary) b) Restricted variance (for example/ instance)	II. <u>Restricted variance/ Literal</u> (for certain/ sure; in-the-not-too distant past/ future)
	III. <u>Unrestricted variance/ Semi-literal</u> (catch a bus/ ferry/ plane)
	IV. <u>Unrestricted variance/ Literal</u> (beautiful/ lovely/ sweet woman)
IV. <u>Literal idioms</u> a) Restricted variance with optional elements (develop from/ into)	V. <u>Restricted variance/ Literal/ Optional elements</u> (shrug one's shoulders; clap one's hands)

The several subclasses of idioms reveal that they not only resist their internal lexical substitution, because they are whole lexical units, but also that they are non-literal, a feature that results from a progressive loss of meaning of their constituents and imposes an external meaning to the whole of the expression.

Additionally, for Fernando (1996: 37-38), form is also an aspect to distinguish between idioms and non-idioms, due to the fact that conventional phrases bear lexical and grammatical characteristics that are not specific of single lexemes or ad-hoc constructions. That's why idioms are phrases that have limits to their extension, whereas non-idioms may comprehend from small phrases to recursive phrases or multi-phrasal structures.

Therefore, the higher limit to idioms would be non-idioms or complex phrases and the lower limit compounds (see sub-chapter 3.1), which share some of the features of idioms, namely the fact that they represent habitual co-occurrences between two or more words, being literal, semi-literal or non-literal.

In conclusion, after the presentation of these several approaches to word combinations, that emphasize the complexity of this issue, it is our intention to go further into this at the level of the thesis, in order to be able to summarise a number of possible tests to be applied to restricted word combinations. We believe this will allow us to identify in a trustworthy way the examples that are collocations and the ones that are idioms within our findings and also come up with explanations for this, following the principle of Toury's descriptive approach.

3.5. Metaphorical construction

Metaphors, as well as metonymies and other figurative language resources, have frequently been identified as the basis for the construction of set phrases, because of the natural connection between metaphors and imagery. Idioms, for instance, are viewed by Kövecses (2002: 193, 201) as a blend that comprehends metaphors, metonymies, word pairs, similes, sayings and proverbs, phrasal verbs, grammatical idioms, among other multilexemic expressions; they are also products of their conceptual system.

As Butcher (cit Johnson & Malgady *in* Hoffman & Honeck, 1980: 260-261) puts it metaphors demand "an eye for resemblances" – it enables to see similarities between things. They relate two concepts that, even slightly similar, are far from each other in their own conceptual fields, one being concrete, the other abstract. Their main function is for Kövecses (2002: 147) to allow the understanding of one thing in terms of the other by mapping the elements of a conceptual field into another.

According to Harris, Lahey & Marsalek (*in* Hoffman & Honeck, 1980: 164), the imagery representation functions as an effective mnemonic to understand and recall metaphors, because these are closer to memory representations than merely linguistic ones. Moreover, Cormac (1985: 143-144) claims that metaphors are linguistic expressions constitutive of cognitive processes, being one of the chief cognitive ways of understanding fuzzy concepts.

As far as Lakoff & Johnson (1980) are concerned, metaphors not only provide meaning to the human experience by highlighting some aspects and omitting others, but also organize the conceptual system. For them, there are several types of metaphors: conceptual metaphors, in which one conceptual domain is understood in accordance to another; structural metaphors, when one concept is metaphorically organized according to another concept; the conduit metaphor that represents the placement of ideas and

37

objects into words, playing the role of a container; and orientational metaphors that allow the semantic organization of concepts according to their spatial relationships within their own conceptual system, giving them a spatial orientation.

4. Audiovisual Translation

As we mentioned before, AVT as a form of translation has been given numerous names, due to the fact that the representations of translation itself vary considerably. This is due to the fact that AVT not only includes a variety of texts, so much as it encompasses a number of different tasks, such as localization, language transfer, adaptation, editing, revision, documentation management, co-authoring, technical writing, versioning, language mediation, copywriting – these are only some examples to illustrate the terminological misunderstanding that reigns in translation. (Gambier & Gottlieb in Gambier & Gottlieb 2001: ix-x)

This terminological inconsistency is also focused by, for example, Chaume (2003: 15) who summarizes a number of designations for AVT: film dubbing (Fodor 1975), constrained translation (Titford 1982), film translation (Snell-Hornby 1988), screen translation (Mason 1989), film and TV translation (Delabastita 1989), media translation (Eguíluz *et al.* 1994), *comunicación cinematográfica* (Lecuona 1994), *traducción cinematográfica* (Hurtado 1994-1995), *traducción fílmica* (Díaz-Cintas 1997), multimedia translation (Mateo 1997; Gambier & Gottlieb 2001) or *traducción para la pantalla* (Mayoral Asensio 2001).

The term AVT came to replace the initial terms of 'film translation' and 'language transfer', which strongly emphasized the linguistic elements (Gambier in Gambier 2003: 171), but this one is considered by Díaz-Cintas (2001: 24) as an entirely appropriate designation for retaining the semiotic dimension of this type of translation. However, the calque from French 'audiovisual translation' also concerned a teaching method fashionable around the 1960s, which explained why other designations were pursued, like versioning, screen translation or multimedia translation (Gambier in Gambier 2003: 171). Nonetheless, the more recent and less problematic designation is audiovisual and multimedia translation, due to the fact that it covers "totes les transferències lingüístiques i culturals d'aquells textos que es manifesten a través de diversos canals de comunicació, però també a través de diferents codis" (Chaume 2003:

16) and it comprehends "tant la traducció cinematogràfica com la traducció I adaptació de produces informatics (…) localització, però també vol incloure la traducció de l´ópera o la traducció de comics" (idem).

Additionally, these new and alternative terms also tend to reflect "the great diversity of expectations and representations related to the concept of translation" (Gambier & Gottlieb in Gambier & Gottlieb 2001: ix), along with the fact that a new conception of 'text' as "polysemiotic multi-signs" is also emerging.

The conventional way of regarding translation as a mere transfer of words from one language to (an)other(s) neglects the fact that this is a highly complex process, emphasized further by the conception of 'text' as a plain string of sentences (Gambier & Gottlieb in Gambier & Gottlieb 2001: x). These views would be enough to rule out what happens in audiovisual means of communication as a form of translation. However, even Jakobson in 1959 (in Venuti 2000: 113-118) went as far as to consider the intersemiotic translation or transmutation as one of his three types of translation (the others being intralingual translation or rewording and interlingual translation or translation proper), in which the translation of audiovisual material can easily be fit, because it involves "an interpretation of verbal signs by means of nonverbal sign systems" (Jakobson in Venuti 2000: 118). Hence, there can be no doubt whatsoever that AVT has to be viewed as a form of translation, because, as it happens in other types of translation, it may include reductions, omissions, summaries, and the like.

Another question is: should we talk about audiovisual or multimedial texts? The term 'multimedial texts' was initially introduced by Reiss & Vermeer (1984) to refer to a category of texts in which "the verbal content is supplemented by elements in other media" (Shuttleworth & Cowie 1997: 105-106). For Gambier & Gottlieb (in Gambier & Gottlieb 2001: x-xi), both texts share a number of characteristics: team work, working with intermediate texts (in between source and target texts, for example scripts), and criteria such as comprehensibility, accessibility and usability, allowing us to understand them as equally acceptable. Of course, this difference does not lie only in translation strategies, but also on the nature of the text, owing to the intersemiotic relationship established between sound and image at both the levels of decodification and transcodification. But even so, all these criteria are not exclusive of AVT, they are also common to many other types of translation.

Hence, we chose to refer to audiovisual texts as the material with which AVT works, without ignoring that the concept of 'text' in this type of translation involves

various channels of communication of the message and demands particular attention to all aspects, such as music, sound, image, among others.

In 2003, Gambier (2003: 172-177) divided AVT into dominant and challenging types, the former including interlingual subtitling, dubbing, consecutive and simultaneous interpreting, voice over, free commentary, sight translation and multilingual production (like doubled versions and remakes), whereas the latter consisting of script translation, surtitling, intralingual translation, real-time subtitling and audiodescription, many of which have now turned into mainstream ones, particularly the last three.

Therefore, AVT should not be mistaken for subtitling (though the mistake is often made), but it should be seen rather as a superordinate term that comprehends several types of subtitling, along with other forms of translation, such as dubbing, interpreting, voice-over and audiodescription. For instance, Shuttleworth & Cowie (1999: 161) regard subtitling as "one of the two main methods of language transfer used in translating types of mass audio-visual communication", completely disregarding other types of "language transfer" involved in "mass audio-visual communication", and also neglecting other designations.

4.1. Subtitling

As is common knowledge, the art of subtitling was born from the "intertitles" used in silent movies, by means of a Swedish and Hungarian invention, which was then taken to France. In this line of thought:

> Subtiltles are condensed written translation of the original dialog which appear as lines of text, usually positioned towards the foot to the screen. Subtitles appear and disappear to coincide in time with the corresponding portion of the original dialog and are almost always added to the screen at a later date as a post-production activity. (Luyken *et al.* 1991: 31)

This type of AVT has always presented a number of advantages that explain why a reasonable number of European countries (Portugal, Belgium, the Netherlands, Greece) choose it over dubbing (Spain, France, Germany) or voice-over (Russia, Poland), for instance the fact that Portuguese or Greek TV subtitlers have organized in separate departaments (RTP and ERT, respectively) that have enable the "speedy and

cost-effective production of easy to read subtitles, even at short notice and for complex subject areas" (Luyken *et al.* 1991: 36).

According to Díaz-Cintas (2001: 49-50), these benefits can be summed them up in the following way: it is a cheaper and fairly quick job; it respects the integrity of the original dialog; it develops the learning of foreign languages; it helps the development of viewers' reading ability in their mother tongue; it maintains the original voices; it is better for the deaf and hard-of-hearing and for immigrants. However, it also holds a number of disadvantages which are the fact that it contaminates the image on screen, leading to the spreading of attention across several aspects, like the image, the written text or the soundtrack; it demands more reduction of the original text because of time and space limitations; it does not allow for the overlapping of dialogs; it is hard to manipulate; if viewers get distracted or lost, they are unable to read the subtitles; it may lead to some disorientation due to the presence of (at least) two linguistic codes; and it may permit the entrance of linguistic calques.

On the other hand, dubbing also holds a set of pros and cons. Among its drawbacks, it should be mentioned that it turns out to be more expensive; it leads to the loss of the original; it is usually more laborious and slow; it intends to be a domesticating product; the voices of the actors can be repeated; and it must abide by lip synchronization. As for its advantages, it enables a less problematic manipulation of dialogs and their overlapping; it is considered more beneficial for children and illiterate; it respects the image on screen, not contaminating it, thus viewers can concentrate solely on the image and sound; it does not need to reduce that much text as in subtitling; it makes use of only one linguistic code and of oral language features; and it prevents the entrance of linguistic calques. (Díaz-Cintas 2001: 49-50)

Hence, subtitling must be regarded as a linguistic practice that wishes to offer a written text, normally at the bottom of the screen, accounting for the dialogs going on among actors or for monologues (Cintas 2001: 23), or a "kind of simultaneous written interpretation" (Gambier cit. de Linde & Kay 1999: 2). Chaume (2003: 18) describes it further as consisting of the "incorporar text escrit en la llengua meta a la pantalla on s'exhibeix una pel·lícula en versió original, de manera que aquest text en forma de subtítols coincidesca aproximadament amb les intervencions dels actors de la pantalla".

Consequently, subtitles, often referred to as captions as well, are "transcriptions of film or TV dialog, presented simultaneously on the screen", along with the image, sound, paralinguistic elements and others, and "usually consist of one or two lines of an

average maximum length of 35 characters (…) [being] placed at the bottom of the picture and [that] are either centred or left-aligned" (Gottlieb in Baker 1998: 245).

It is obvious that these definitions can cover numerous types of subtitling, each with different features and imposing different constraints to translators/subtitlers: Gambier (in Gambier 2003: 172-177) mentioned interlingual subtitling, intralingual translation, real-time subtitling and surtitling, but that sight translation and multilingual production could also involve some form of subtitling.

In a multilingual production, as the name clearly echoes, the output involves a number of different languages and, because of that, could involve sign language interpreting or subtitling for the deaf and hard-of-hearing, not to mention other types of AVT, like dubbing, interlingual subtitling or audiodescription. As for sight translation, this "appears as a hybrid and rather unexplored phenomenon, used in various contexts and with different definitions" (Agrifoglio 2004: 43), and most of the times included as a step in the training for interpreting. Nonetheless, a sight translator "reads a written text" and could also reproduce this text not orally, but in a written form, such as in live or real-time subtitling.

Furthermore, Díaz-Cintas (2001: 24-26), though not enumerating all these types of subtitling, establishes a typology of subtitling according to three criteria: formal presentation, linguistic elements and technical aspects. As far as the first one is concerned, we can have *traditional subtitling*, either maintaining complete sentences (the so-called verbatim), or being condensed or bilingual (in which each line is devoted to a different language, such as in Belgium), or *simultaneous subtitling*, typical of situations like a live interview. Linguistically speaking, there is *intralingual subtitling*, designed to satisfy different needs, those of the deaf and hard-of-hearing, needs related to the learning of languages and what Díaz-Cintas calls the 'karaoke effect' (connected with the preservation of the original soundtrack, for instance in musicals), and *interlingual subtitling*, resulting in the translation of an audiovisual 'text' from one language to another. Finally, from the technical point of view, he mentions *open subtitling* and *closed subtitling*, according to which one can have either an end product which is inseparable of the translated subtitles (open subtitling), or the audiovisual text is left untouched and is accompanied with a respective translation(s).

In open subtitling, we would be watching a subtitled programme on TV, cinema or video, i.e. with subtitles available to everyone, "forming part of the original film or broadcast" (Shuttleworth & Cowie 1997: 161), preventing us to take them off the

screen. In closed subtitling, "broadcast [is done] separately and [is] accessible (…) by means of teletext" (idem), for example the case of subtitling for the deaf and hard-of-hearing, the case of DVDs or real-media on the Internet.

Bearing in mind the several types of subtitling within AVT and their distinction according to specific criteria, it is worth mentioning Gottlieb's (in Baker 1998: 245-247) three distinctive features of subtitling as a form of translation, that will lead to the understanding of some of the constraints involved in the practice of subtitling: the semiotic composition, the time dimension and the pragmatic dimension.

According to the *semiotic composition*, translated texts can be either monosemiotic or polysemiotic, whether they use only one channel of communication, which translators control, or they use other channels of communication, such as the visual and the auditory, which will influence translators' job. In addition, polysemiotic texts can be isosemiotic if the translation uses the original channel, or diasemiotic if the translation results from a combination of different channels, which occurs in the case of subtitling. Consequently, in subtitling one has to work with four simultaneous channels: the verbal auditory channel (dialog, background voices, lyrics); the non-verbal auditory channel (music, natural sounds, sound effects); the verbal visual channel (titles, written signs on the screen); and the non-verbal visual channel (picture composition and flow). This means that every decision made by the translators/subtitlers will affect the end product in any of these four channels, which is especially relevant in intralingual subtitling.

Concerning the *time dimension*, it must be remembered that subtitling is dependent on the "time for production of the original", the "time for presentation of the original" and the "time for presentation of the translation", making it a type of synchronous translation, because it is in synchrony with the original, as well as of contemporal translation, since it is connected with the original in terms of time and space.

Finally, regarding the *pragmatic dimension*, since "intentions and effects are more important than isolated lexical elements", which make up an audiovisual 'text', translators/subtitlers will have to ensure that a considerable dialog restriction and concision is achieved, involving intersemiotic and intrasemiotic conciseness, so as to avoid redundancy of information that is given by facial expressions, tone of voice, the rhythm of music and sound effects.

To conclude, more than the idea of transferring, restricting, reducing or adapting, one should retain Gambier's concept of transadaption (in Gambier 2003: 178-199), which involves the already-mentioned temporal constraints, the density of information and the relationship established between the spoken and the written codes, and "allows us to go beyond the usual dichotomy [between] literal/free translation or translation/adaptation", as explained above.

4.2. Interlingual subtitling

Interlingual subtitling, also known as traditional subtitling or open captioning, refers to the type of transadaptation of a so-called 'source text', part of which may match a post-production script (when there is one), being associated to what goes on screen, into a set of (usually) two-line 'target text', with (normally) 34 to 40 characters each, that is to be presented to viewers (most often) at the bottom of the screen every four or every six seconds, while an audiovisual product is being broadcasted.

According to Díaz-Cintas (2001: 112-115), the presentation of subtitles on the screen must obey a series of formal, technical and linguistic conventions, and not norms.

Technically speaking, the first line of a two-line subtitle should try to be shorter than the second in order to avoid contamination of the picture, as long as it does not break units of meaning. All subtitles should be well cued as far as possible, thus reflecting the rhythm of the film, and their pace should be as stable as possible throughout the film, as well as adequate to the reading ability and reading speed of the intended audience. It is totally acceptable for the subtitle to enter up to half a second before the actor speaks and to exit half to one and a half seconds after the actor stops speaking, as long as there are no shot changes, and to separate one subtitle from another by no more than a second.

The subtitles must be legible to the readers, that is why most subtitling companies choose to use either Arial, Times New Roman or, more recently, Sans Seriff Lettering as their preferred fonts at a size of 12, normally in white (but sometimes in yellow, thic colour no longer in use in Portugal) and very rarely using background boxing colors (usually black and white, fairly frequent in Portugal, especially to cover the initial or final credits that overlap with dialogs).

In the *linguistic* point of view, subtitles must be as adequate as possible, respecting all idiomatic matrices and cultural references; each subtitle should bear a complete semantic and syntactic idea, avoiding the same idea to go on through several subtitles, unless absolutely necessary; the reduction of the dialogs must respect their coherence and cohesion; messages that appear in the picture should be conveyed in the subtitles; and lyrics should also be subtitled.

Finally, as far as *orthographical and typographical conventions* are concerned, subtitles should reproduce the rules of their target languages and reach equilibrium in the use of punctuation on the screen. For instance, according to Portuguese conventions, a hyphen is used to indicate dialog between two people; suspension points have a double function, showing that a subtitle is to continue in the next one or that there is a pause, omission or interruption in the dialog; capital letters should be scarcely used because they are difficult to read and they take up too much space on the screen; italics represents off-screen voices, voices coming from the radio or telephone, thoughts or dreams; inverted commas are used for quotations and must be repeated at the beginning of each subtitle until the quotation finishes; abbreviations (like Mr.) or numbers (333) bring about problems for readers because they are less readable than expected.

It is then clear that the usual target audiences of interlingual subtitling are viewers that (probably) are not physically challenged (neither auditorily nor visually) and that are presented with subtitles they cannot remove from the picture, because of their country's cultural habits, though possessing uneven reading abilities.

On the other hand, Ivarsson (1992: 53-72) presents an approach which is based on his professional practice developed in Sweden. In terms of legibility, he mentions the use of a simple typeface, such as Helvetica or Universe, and the unquestionable use of lower case characters, instead of capitals, and the kerning, i.e. the spacing not only between characters, but also between words. Concerning the layout of subtitles, Ivarsson discusses the placing of text at a centred position or fixed left margin, depending on the country's tradition and if it has or not adopted the principle of cinema. There is also this need "to keep the important part of the picture unobstructed, either by limiting the text to one-liners at the bottom of the screen during close-ups or by moving the text to one side of the picture" (Ivarsson 1992: 65). Although most countries choose to regard two lines as common sense and practice, being that each line "cannot usually exceed about 40 letters and spaces" (Ivarsson 1992: 66), there are exceptions such as in subtitled news presentations, that can go up to three lines, to mention only one example.

Added to the question of legibility, accuracy is also dealt with by Ivarsson (1992: 77-81) because "translations [in subtitling] simply must be correct, and omissions as few as possible within the constraints of the inexorable "time limits"". For this reason, it is important to be suspicious of everything, of one's work, of the original and its possible errors, and to proofread the entire work (if possible by some else), as well as to doublecheck the subtitles with the picture and the sound.

Another point Ivarsson (1992: 90-95) makes is to highlight the importance of editing, i.e. the need to select and thus condense the text to be subtitled, the use of omissions, paraphrases or ellipsis (if redundant towards the image), the elimination of muddled speech and the merging of short dialogs, the use of simple vocabulary, not to mention the careful and parsimonious use of punctuation signs and conventions related to letters, numbers, time, units of measurement, currency, abbreviations, titles and institutions, forms of address, songs and poetry, as briefly mentioned above.

In conclusion, going back to the issue of conventions versus norms, national and/or international norms and standards are gradually being put forward as a way to standardize subtitling practices in different professional spaces, from TV to video games. Nevertheless, in Portugal, there are no standards officially published for interlingual subtitling, except those that are used internally in Portuguese TV channels, namely SIC, and subtitling companies that also provide training. It is worth mentioning though the recent release of a guide to subtitling for the deaf and hard-of-hearing in Portugal – "Vozes que se vêem" (Neves 2007) – which is of great value not only because it is the first in this field, but also because it may serve as a comparison for future work in the standardization of interlingual subtitling in the country.

4.3. Dubbing and voice-over

According to Gambier (in Gambier 2003: 172), *dubbing* is a dominant type of AVT and "involves adapting a text for on-camera characters", requiring also lip, visual, gesture and facial synchronization, though this is a question of cultural tradition – some people may be more tolerant to dischrony than others. This is the typical AVT method to handle animation films or other children's programs.

For Chaume (2003: 17), dubbing consists of:

la traducció I ajust del guió d'un text audiovisual i la posterior interpretació d'aquesta traducció per part dels actors, sota la direcció del director de doblatge i els consells de l'assessor linguistic (...). Tècnicament, (...) es reemplaça la banda dels diàlegs originals per una altra banda en la qual aquests diàlegs s'enregistren traduïts en llengua meta i en sincronia amb la imatge. (Chaume 2003: 17)

Gambier (2003: 174) also mentions 'post-synchronization', as a type of multilingual production, meaning that actors when dubbing use their own mother tongue which will be later post-synchronized in only one language.

On the other hand, *voice-over* is designated as 'half-dubbing' or 'partial-dubbing' and it means that the original sound is reduced to a lower level so that "the target voice is superimposed on top of the source voice" (Gambier in Gambier 2003: 173-174). It occurs in documentaries (for example, in Portugal, where these programs are voiced-over and simultaneously subtitled: the voice of the narrator is voiced-over, whereas others' are subtitled) and live interviews (like the live broadcasting of the Oscar Awards from Hollywood).

However, it is interesting to notice that Shuttleworth & Cowie (1997: 44-45) consider dubbing as a type of AVT including both "any technique of covering the original voice in an audio-visual production by another voice" and "other types of revoicing, such as voice-over, narration or free commentary". Dubbing is then regarded as a lip-sync process (or "the imperfect art" in the words of Luyken *et al.* 1991: 71) that involves a considerable number of stages apart from language transfer, as well as other factors, like technical issues (checking the material to be dubbed and its script, visualizing the material, translating and adapting it to lip-sync constraints, and delivering it to the recording studio), up-to-date equipment, actors to be chosen, the competence of the dubbing editor and the sound equipment. It is definitely "an exercise of visual phonetics" (Fodor cit. Shuttleworth & Cowie 1997: 45) that requires visual and acoustic synchronization, being the latter more important than the former most of the times.

In the same way, Baker & Hochel (in Baker 1998: 74-75) also understand dubbing and revoicing as the two types of oral language transfer in the audiovisual context, though they also allude to the fact that *revoicing* may work as a generic term for "all methods of oral language transfer, including lip-sync dubbing". Nonetheless, the

several methods of revoicing may be pre-recorded or broadcast live, while dubbing is always pre-recorded.

Therefore, and taking Agost's (in Duro 2001: 242-244) standpoint, dubbing is an audiovisual choice dependent on several factors:

- technical factors (for example, the immediate nature of the broadcasting which will determine choosing one or another AVT mode);
- economic factors, because TV channels buy the products they believe will guarantee a potential group of viewers;
- political factors, since, in some countries, it is the government that choose the general audiovisual policy to be followed in the several TV channels (consider the cases of the dictators in Portugal and Spain), with view of normalizing the language use in the means of mass communication;
- the function of the product – depending on the purpose of the program, it may be dubbed or subtitled; examples of this could be weekly programs for informing viewers of the broadcasting agenda of the TV channel or programs that have pedagogical purposes, intending to develop the knowledge of a certain foreign language and culture;
- the target audience, who will determine the choice for dubbing or subtitling, which is also a question of cultural tradition of the countries;
- intertextuality, present when there are continuous references to the daily life of other societies, to the private and public life of VIPs or to programs in other TV channels or a commentary of the latest social, political and cultural events; the more intertextuality a program shows, the less likely it is of being dubbed.

In conclusion, when discussing the issue of the advantages and disadvantages of dubbing versus subtitling, Shuttleworth & Cowie (1997: 46) state that dubbing may be said to be less authentic or less flexible than subtitling; it is more expensive and demands more time to be completed (generally, a translator/subtitler may be asked to do the translation and subtitling of a two-episode series of about one and half to two hours in two days); it asks for less cognitive effort from the viewers; it requires less reduction of the message; it might have strong cultural, ideological and linguistic implications, in the sense of domesticating or naturalizing a foreign audiovisual product, thus defending the national language and culture (think of Spain, for instance), but also developing

national stereotypes. Moreover, dubbing also prevents the viewers from listening to the original foreign language, thus determining their fluency at that particular language (Baker & Hochel in Baker 1998: 75).

4.4. Domesticating and foreignizing strategies

It is worth mentioning the domesticating and foreignizing strategies that can underlie the audiovisual method to be chosen. Strategies of translation are determined by cultural, economic and political factors, which lead a country to prefer a more conservative approach, "appropriating [the foreign text] to support domestic canons, publishing trends, political alignments" (Venuti in Baker 1998: 240), or one that aims to "revise the dominant by drawing on the marginal, restoring texts excluded by domestic canons, recovering residual values (…), and cultivating emergent ones (for example, new cultural forms)" (Venuti in Baker 1998: 240).

Nietzsche (cit. Baker 1998: 241) regarded translation as a form of conquest, exemplifying with the case of the cultural and literary appropriation that Romans did with Greek culture: they attempted to delete Greek cultural markers and replaced them by Roman specific ones. Thus, domestication engages into retaining home-made canons in order to serve domestic imperialist, evangelical or professional purposes, being dependent on cultural and political developments (Venuti in Baker 1998: 241) and turning out to be a "narcissist experience" (Rodrígues Espinosa in Duro 2001: 104). It consists of translating according to a clear, fluent and acceptable style for the target audience, eliminating all possible difficulties brought about by foreign references or even replacing them (Zaro Vera in Duro 2001: 55).

In the case of AVT, *domestication* is personified by dubbing, which "is an assertion of the supremacy of the national language and its unchallenged political, economic and cultural power within the nation's boundaries" (Danan cit. Zaro Vera in Duro 2001: 55). For instance, scholars, such as Ballester (1995), defend that the dubbing policy in Spain is due to the censorship policy led by Franco, whereas others, like Agost (1996) and Ávila (1997), conclude that this was intended to prevent the development of minority languages in Spain. Both readings reflect this domesticating approach to AVT.

Conversely, Salazar managed to exercice his censorhip by using subtitling, instead of dubbing, though his concerns were similar to those of Franco. According to

Pieper (forthcoming: 3), this choice of Salazar's was due to the high rate of illiterates (almost 40% in 1950), which led to the fact that "large parts of the population were excluded from the pleasure, or understanding, of foreign films. In the full interest of the state's ideology, they were 'protected' from influences from abroad."

On the other hand, *foreignizing* translation brings about the awareness of the foreign and answers a specific domestic situation by opposing it and "challenging literary canons, professional standards and ethical norms in the target language" (Venuti in Baker 1998: 242), and by becoming a means for cultural innovation. Foreignization reproduces the cultural idiosyncrasies of the source text, by becoming an "ethnodeviant pressure of those values to register the linguistic and cultural difference of the foreign text, sending the reader abroad" (Venuti 1995: 20).

This approach is typically exemplified by subtitling, in which viewers still have access to the source text and, at the same time, to the translation of this text into their own language(s). It is more foreignizing than domesticating, because the translator's work is more visible by being simultaneous to the source text in the foreign language (Zaro Vera in Duro 2001: 57).

Consequently, the choice of one or another method of AVT within a country is determined by a myriad of factors, among which "cost, availability of relevant technology, standard of literacy, interest in foreign languages, degree of cultural openness, and the strength of the local film industry", together with "the audience habits" (Baker & Hochel in Baker 1998: 75).

5. Audiovisual genres

According to Rabadan (cit. Chaume 2003: 172), Contrastive Textology can be defined as a "marco teórico de análisis que combina los procedimientos del análisis contrastivo y el análisis del discurso y que pretende descubrir las correspondencias que existen entre las concepciones textuales y las estructuras lingüísticas de ambos polisistemas". The problem arises when the translator is confronted with the different textual conventions of the source and target cultures and needs to select the most adequate to the target audience, needless to say that it occurs in all types of translation, including in AVT.

To begin with, it is necessary to understand the difference between genre and text type, which frequently leads to terminological misunderstandings. Chaume (2003: 174) presents the definitions retrieved from Nord (1991), who based her approach on Reiss and Vermeer (1984), as a way to dispel any doubts about these two terms: *text type* refers to the function of the text, which is based on Bühler's tripartite language functions, thus leading to the expressive, informative and operative text types, whereas *genre* is also designated as text form, meaning a "forma convencialitzada de textos recognoscible per una comunitat de parlants com a tal I que acompleix una funció comunicativa" (Chaume 2003: 175).

Before the birth of the new means of mass communication – the cinema, the radio, the television or the computer –, the only audiovisual texts available were the theater and the opera. Those new ones managed to break the barriers of the present moment, of space and time, creating new spaces for the expression of reality and imagination. Among the new audiovisual spheres, Agost (in Duro 2001: 231-232) considers that television stands out as a space for antonomasia and fragmentation, due to the huge variety of programs that comprehends coming from the most various sources (e.g. the theater, music, the opera, the news or the cinema).

Consequently, *audiovisual texts* share a number of characteristics: pragmatically, the type of participants in the communicative situation, the communication contexts and the communicative intention; in terms of communication, the language types and varieties and the viewer; in the semiotic point of view, the genre.

As far as the users/viewers are concerned, and according to Agost (in Duro 2001: 233-234), audiovisual texts can be described as having an almost endless scope, including a heterogeneous number of people from both the emission (broadcasting companies, TV channels) and reception sides (end users). The communicative situation is dependent on economic criteria, since the cinema and TV have turned into money industries, which explains the obsession for TV audiences. Finally, the communicative intention is related to the main role of these texts: to entertain, to inform, to persuade and, in some cases, to try and change the public's opinion about or attitude towards a certain topic – they are then marked by a wide-ranging intention in which the exposition, narration and education are the leading ones.

From the usage perspective, audiovisual texts comprehend all language types (from language registers to dialects, from documentary to colloquial languages) and possible themes. Finally, within the semiotic dimension, one has only to consider the

importance of the discourse in these texts, that is to say the significance of ideology in politics, in the news, in the language used, in the translator and the work they do.

Discourse is the means of expression for audiovisual texts, because when they reach the viewers, they hold a conventional, pre-established form, which receives the designation of *genre*. Some genres are typical of audiovisual means of communication; others are common to different means, such as the news, literary or cinematographic, televised theater or printed advertisement. One of the problems associated with audiovisual texts is the emergence of new formats and new genres in order to catch the public's attention to new audiovisual products. (Agost in Duro 2001: 235)

From Agost's (in Duro 2001: 235) viewpoint, the need for the concept of *genre* appeared when the need to classify different spoken texts from Ancient Greece was also felt. The literary culture then established three major genres: lyric-poetic, epic-narrative and dramatic-theatrical, which were later combined with the expressive, communicative and referential text types. Nowadays, genre is applied to "las manifestaciones que han generado las nuevas tecnologías de la comunicación y en especial al cine, la radio, y la televisión" (Casalmiglia & Tusón cit. Agost in Duro 2001: 235).

For Biber (1989: 9), genres "are text categorizations made on the basis of external criteria relating author/speaker purpose and the text categories readily distinguished by mature speakers of a language"; they are based essentially on external format criteria.

It is obvious that genres are also dependent on and influenced by the society in which they appear – they accomplish a function and possess a pragmatic validity. Their function is to organize the diversity of the discourse spaces and their validity comes forth as a way of text organization into different discourses according to pragmatic, enunciative, syntactic and semantic criteria. Agost (in Duro 2001: 237-238) claims that the audiovisual text is the concept used for understanding the different discourse practices in the audiovisual context and when we talk about television, the basic unit of broadcasting and the concrete manifestation of a specific genre is the program.

This same author (in Duro 2001: 239-241) presents a number of criteria for classifying *audiovisual genres*, of which one is the mode (these texts make use of visual and linguistic codes), the other the scope of use (television, cinema, video or computer) and the last the function of the specific program (the predominant communicative intention, which for some authors, such as Nord (1991) or Bathia (1993), determine both the genre and their internal structure). The majority of programs achieve one of

these functions: telling fictional stories; informing about facts and real events; acting upon viewers so as to change their habits, attitudes or behaviors; entertaining and maintaining contact with the spectators.

Although any text, audiovisual or not, can have more than one function, it is possible to divide audiovisual texts into four chief macro-genres: dramatic, informative, advertising or entertaining genres. The *dramatic genre* comprehends narrative (films, TV series, soap operas, TV movies or cartoons), descriptive (documentary or philosophical films) or expressive texts (televised theater or opera, musicals), while the *informative* one includes documentaries, reality-shows, programs about the social life, interviews, debates, weather forecast or DIY programs. On the other hand, the *advertising genre* is characterized by its mainly educational purpose, combined with the conversational and the expositive, such as institutional campaigns, election campaigns, among others. Finally, the *entertaining genre* encompasses a more heterogeneous group, namely the narrative (social chronicles, sports emissions), the conversational (contests), the expressive (humorous or musical programs), the predictive (horoscope) or the educational (gymnastics program) types.

Apart from the function, it is of the utmost importance to consider also the mode, which allows for the distinction between those genres that are more oral and spontaneous and those that are more written and deliberate, and the tone that measures the attitude of the enunciators and their enunciative distance, distinguishing the documentary sub-genre from the reality-show, for instance. Nevertheless, this classification cannot be understood as a closed one, because of the speed at which new audiovisual formats are showing up.

5.1. Documentary films

Delabastita (1989: 196-197) when discussing mass-communication defines *film* as a "multi-channel and multicode type of communication", which takes place in two different channels – the visual and the acoustic. Although his paper focuses on fiction films, some of the aspects he refers to are also true for documentary films, especially if we think of recent ones that involve the use of short historical simulations, as if it was a theatrical performance, or of animation. These are new narrative strategies that enable viewers to better grasp the reality being presented in the documentary and visualize it as if it was happening in front of their eyes. Therefore, the sign systems involved in the

production of both types of films are the same: the verbal code (linguistic and paralinguistic features); literary and theatrical codes (conventions for the construction of the plot, for the dialogs or for the narration); proxemic, kinesic, vestimentary, make-up codes; and cinematic codes (rules, conventions, techniques). All these codes combined together present the viewers with a highly complex sign, whose translation will have to conform to a number of constraints imposed by the type of AVT to be chosen.

The same author (Delabastita 1989: 199-200) proposes to explain the choice for one or another type of AVT by applying the traditional categories of rhetoric to these sign systems:

> *repetitio* (the sign is formally reproduced in an identical manner), *adiectio* (the sign is reproduced with a certain addition), *detractio* (the reproduction is incomplete, it implies a reduction), *transmutatio* (the components of the sign are repeated in a somewhat different internal order, there being an alteration of the sign's textual relations) and *substitutio* (the sign is replaced with an altogether different sign). (Delabastita 1989: 199-200)

Afterwards, Delabastita (1989: 199-200) matches these rhetorical categories with the possibilities of AVT for films: if the acoustic verbal signs are to be combined with *substitutio*, we will have a dubbed version, in which the original source text will be completely replaced by the dubbed target text); if the visual verbal signs are to be merged with *adiectio*, then we shall get a subtitled version, that is to say that the target film is an precise reproduction of the source film, with the addition of verbal signs on the screen, the subtitles.

6. Research Methodology

The methodology chosen for the development of this descriptive-oriented piece of research in AVT which focus on phraseological issues attempts to take contributions from various sources in order to carry out the pilot study to be described further on.

Pöchhacker (2004: 63) proposes a three-fold methodological approach, which he calls research strategies: field work, survey research and experimentation. If a researcher is to make use of the first one, it leaves the phenomenon to be considered as it is, studies its context and collects data about it from its natural settings. As for the

second one, the investigator will collect data in a considerably standardized way, developing a horizontal approach in order to have a broader view of the situation and gathering as many observations as possible, since it aims at quantity. Finally, concerning the last one, the researcher controls independent variables, manipulates and changes one of these variables so as to try to assess its output with the help of dependent variables. Added to these, this author also mentions data collection methods, including documentary analysis, which intends to record; to use questionnaires, interviews and focus groups that aim to ask; and participant observation, which means to watch.

On the other hand, Williams & Chesterman (2002: 65-66) give other examples of empirical research methods: case studies (the same as Pöchhacker's field work), based on the study of limited situations in their natural environment, of which only a single feature or even a set of features will be studied; corpus studies, focusing on an extensive collection of texts, which are to be analyzed according to pre-defined objectives; survey study, exploring and describing the distribution of a phenomenon over a specific population; and historical and archival research, which investigates, analyzes and tries to understand documents and other type of data on a specific topic.

Furthermore, van Doorslaer (cit. Hermans 1999: 70) by focussing the concept of representativeness of samples in TS says that it is necessary to look at it from the quantitative and qualitative aspects: the former attempts to reach a balance between economy and credibility, i.e. it "should be large enough to be credible in light of the purpose of the exercise, but small enough to permit appropriate depth" (idem), the latter is "a matter of interpretation and judgement" (idem), in which extra-textual information might be of influence.

The need to select passages from a corpus and thus choosing samples, the siuational elements according to Holmes's terminology (cit Hermans 1999:70), should lie in the functional perspective. In Nord's words: "Who transmits to whom, what for, by which medium, where and when, why, a text with what function? On what subject-matter does he say what (what not), in which order, using which non-verbal elements, in which words, in what kind of sentences, in which tone, to what effect?" (Nord cit. Hermans 1999: 70)

For the purposes of our pilot study, we will aim at working on a case study/field work, that is to say to make the most of our documentary film, "The Real Da Vinci Code", in order to deal with some of the instances of fixed language found in the subtitled version of this program. We will single out this feature in our representative

sample within the subtitled version of this documentary film, so that later on, during the development of our doctorate thesis, we will have a starting point to compare it with the results that are to be taken from the voiced-over version of the same film.

Our working methodology consisted of the choice of a documentary film (more in the work that approaches); contacts with the broadcasting TV channel (with no success as will be explained below); the transcription of the original soundtrack; the writing down of the subtitled version; and the analysis of possible cases of set phrases.

Later on, we will collect more documentary films subtitled or revoiced from English into European Portuguese, by contacting both the Portuguese public TV channel – RTP – and several satellite channels – such as Discovery Channel, in an attempt to create a considerable corpus of the original films and their parallel texts in the audiovisual versions, and move on to the second part of our research – conducting a corpus study/survey research.

Therefore, this will allow us to have a broader perspective of our object of study, by combining a significant number of texts (quantitative approach) with the analysis of a feature of these texts (though not exclusive of them) according to the theoretical background previously presented (qualitative approach).

It is then impossible to say that any research will be done using only an empirical or a theoretical framework, and disregarding each other's methods, strategies and techniques. The desirable approach would be what Creswell (1994: 177-178) mentions as being a mixed-methodology design, according to which we would reach "the highest degree of mixing paradigms", combining aspects of theoretical and empirical approaches, of qualitative and quantitative methods, of inductive and deductive models, and their respective strategies and techniques. The same author (Creswell 1994: 173-175) speaks of a combined qualitative and quantitative research design, referring to, for instance the concept of triangulation, by which bias inherent in certain sources or investigators "would be neutralized when used in conjunction with other data sources, investigators and methods". This triangulation would range from combining observations and mixing different types of research, like experimentation and ethnographic research, and survey research and qualitative procedures.

By making use of this mixed methodological approach, we intend to confirm and/or refute our hypotheses in the various polysystems this research moves: descriptive translations studies; AVT, specifically subtitling; the issue of genres and text types in terms of audiovisual genres; and the Portuguese context.

7. A pilot study: "The Real Da Vinci Code"

7.1. Brief presentation

The pilot study for the present research project was based on a documentary film originally in English called "The Real Da Vinci Code", based on the blockbuster book by Dan Brown – "The Da Vindi Code" – which was published in 2003, had sold 60.5 copies by 2006 and has been translated in 44 languages. This documentary was subtitled and voiced-over in European Portuguese and broadcast first by RTP1 and then RTP2 on two different occasions. The original English version was produced by Wildlire Television, Channel Four Television MMV; the subtitled version was translated by Filomena Pedro and subtitled by Teresa Sustelo; and the voiced-over version was also translated by Filomena Pedro and the revoicing process was in charge of Carlos Ribeiro, Isabel Angelino, José Henrique Neto, Maria Margarida and Paulo Lázaro. The documentary film is unusually long: it's 102 minutes, when normally documentaries vary between 45 and 60 minutes.

The documentary film in its two versions was recorded on 17th May 2006 and 23rd May 2006, with a time span of one week: the subtitled version was first broadcast on RTP2 and the voice-over one afterwards on RTP1. It is worth mentioning that, in recent years, the difference between one public channel and the other has been emphasized especially by the type of programs televised, leading the general public to consider RTP1 more popular and RTP2 more elitist.

Although some efforts have been made to gather further information on the original documentary and its two versions and on the context in which the audiovisual work was conducted next to the head of the Translation Services of RTP, it was impossible to find out more details about these. It was then impossible to have access to the script and the subtitling of the documentary film, requiring the former to be transcribed and the latter to be written manually in Microsoft® Word.

Finally, we have no idea whatsoever of the reasons underlying the broadcasting of the same documentary film in both public TV channels in two different AVT types, with only a week apart. This, along with other aspects, will be the focus of further research in our thesis.

7.2. Aim

The purpose of our pilot study is to find instances of restricted co-occurrence of words or set phrases, namely idiomatic expressions (complete phrasemes) and collocations (semi-phrasemes), as opposed to free co-occurrences or free phrases, in the subtitled and voiced-over versions of the documentary film "The Real Da Vinci Code". However, within the context of this minor dissertation, the focus shall be given only to the subtitled version of this documentary film, which will be later on complemented with the same instances of the already-mentioned set phrases in the voiced-over version.

At the same time, it will be necessary to assess in what way, if any, these set phrases originally in English were identified by the translator and transferred into a Portuguese equivalent or if they were omitted in the subtitling, either because the translator missed to recognize them or due to the constraints imposed by the process of subtitling. The underlying reasons for not having included the translations of the set phrases in the Portuguese subtitling will be another aspect to be exploited at the level of the thesis.

7.3. Research questions

The research questions that guided our research were as follows:
   a. Are set phrases common in audiovisual texts, specifically in documentary films?
   b. Are they identified and taken into account by translators when translating an audiovisual text to be subtitled or revoiced?
   c. Do translators include them or omit them in the subtitled and voiced-over versions of audiovisual texts?
   d. Is this due to the constraints posed by the subtitling process or by other factors?
   e. Are set phrases strengthen or weaken by what goes on on the screen?

7.4. Hypotheses

Bearing in mind these research questions, the following hypothese were formulated:
   a. Set phrases are as common in general language as in specialized language.
   b. Audiovisual texts, especially documentary films, make use of many set phrases.

c. Translators when translating for subtitling or voice-over make different choices due to the different constraints they have.

d. There is a higher number of set phrases omitted in subtitling.

e. Set phrases are strengthened by what occurs on screen.

7.5. Materials

The materials that constituted the basis for our research are, as previously explained, one documentary film originally in English, 102 minutes long, in its audiovisual versions – subtitled and voiced-over in European Portuguese. They were recorded at the time they were broadcast (May 2006) from the two Portuguese public channels – RTP1 and RTP2.

The original sound track was obtained through the transcription, since it was impossible to get its script, and the same method was used to have the subtitled and the voiced-over versions. For this, we used URUSoft© Subtitled Workshop, version 2.51, and Microsoft® Word.

7.6. Methodology

The methodology we followed to conduct this pilot study was, once having the complete versions of the subtitling and the original sound track of the documentary film, to go through them in search of set phrases in the original and then compare them with their translation in the subtitled version.

In addition, we also attempted to identify and classify them according to the theoretical framework presented in the previous chapters, especially in terms of collocations and idiomatic expressions. In a further stage of this research, some tests will have to be undertaken in order to confirm that the set phrases are really an example of restricted word combinations and provide evidence that they are collocations or idiomatic expressions.

7.7. Results and discussion

Although it would be possible to present an extensive list of possible set phrases used in this documentary film, we chose to produce only the examples that are in table 5, because we believe they are enough to illustrate the aim of this pilot study.

Table 5. Some results taken from the pilot study "The Real Da Vinci Code".

	ORIGINAL VERSION (EN)	SUBTITLING (PT)	CLASSIFICATION OF PHRASEMES AND METAPHORS
1.	"on the heart of (any quest)"	"no cerne de qualquer busca"	Complete phraseme Metaphorical: conceptual and orientational
2.	"think-tank"	"espécie de grupo de reflexão"	Complete phraseme Metaphorical: structural and conduit
3.	"make one's legwork"	"o meu próprio caminho"	Semi-phraseme Metaphorical: structural
4.	"a spit in the ocean"	"uma gota no oceano"	Complete phraseme
5.	"one hell of a journey"	"A coisa promete…"	Complete phraseme Metaphorical: structural (and conduit)
6.	"follow the footsteps"	"seguir os passos"	Semi-phraseme

Therefore, from a first analysis of these examples, we can conclude that only two are semi-phrasemes (collocations), one of them clearly metaphorical (though some scholars, such as Firth, Searle or Goalty, would call them dead metaphors), and the other half complete phrasemes (idiomatic expressions). It is worth mentioning again that tests will have to be done before assuming that these conclusions are accurate.

Secondly, our contrastive analysis of the English examples and its respective Portuguese translations has enabled us to withdraw some conclusions.

As far as the first example is concerned, *on the heart of* means that something is at center of something else, having thus a special importance. It is a complete phraseme that has a metaphorical basis (particularly orientational), because it uses a part of a human's body – the heart, which is in the middle of the body – to represent the centre of a certain activity, in this case of the quest. In the translation, the metaphorical complete phraseme is simplified and the figurative aspect is eliminated.

The print screen in Fig. 1 shows that the action on screen is helping the strengthening of the chosen translation, since the narrator is positioned at the center of

the live representation of the "Last Supper", preventing the viewers to actually see the person playing the role of Jesus Christ, who is replaced by the narrator.

Figure 1. Prints creen of "on the heart of (any quest)".

In the second example, *think-tank* represents a group of people established by a government or organization in order to advise them on particular subjects and to suggest ideas. Here, it seems that the translator considered a paraphrase ("grupo de reflexão) more appropriate to the context, which can be arguable but perfectly possible.

Fig. 2 seems to be an example in which the image does not bring any advantage to the understanding of the set phrase in question, since the only characteristic that could be related to *think-tank* would be the setting, which is a library where a scholar is reading.

Figure 2. Prints creen of "think-tank".
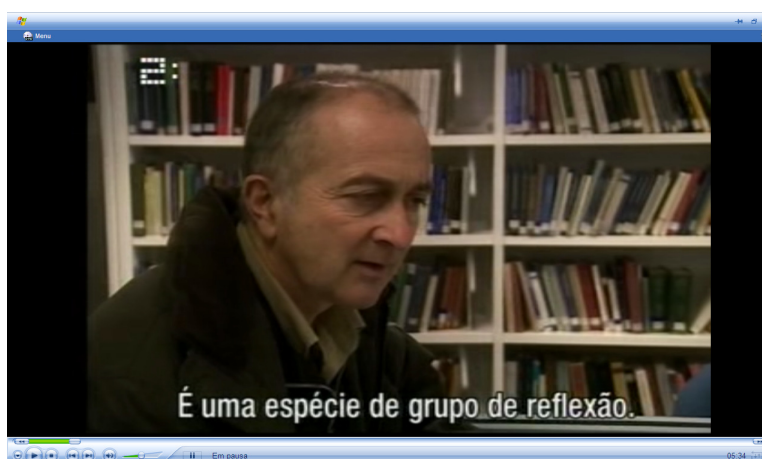
Concerning the third example, *legwork* means practical or boring work that needs to be done. This meaning was clearly lost in the translation and replaced by a somewhat different sense, which is equivalent in English to *follow a certain path*. At first sight, this could be considered an incorrect translation, but it is actually a set phrase in Portuguese – *fazer o meu próprio caminho* – meaning to do something myself. An alternative translation could be *partir pedra*, but that would raise the issue of language registers, something to be explored in the thesis.

The action preceeding the print screen in Fig. 3 is the narrator gathering books from several shelves so as to say that he has to make his own legwork, represented by the pile of books he carries on his arms.

Figure 3. Print screen of "(make one's) legwork".
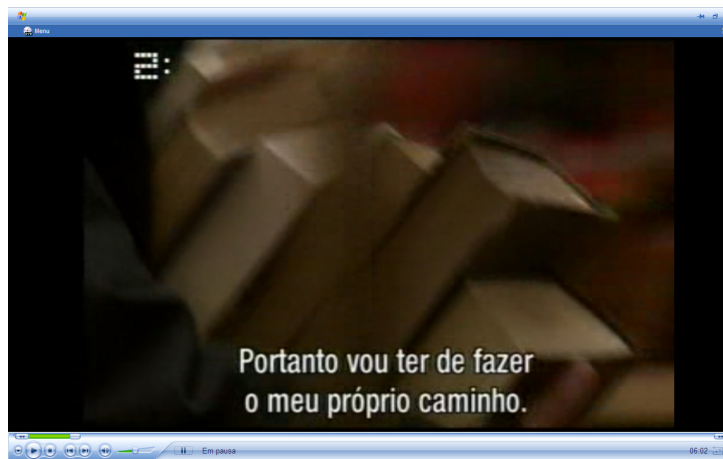
As far as *a spit in the ocean* is concerned, this was definitely the most interesting example, because the most usual synonym is actually *a drop in the ocean*, something in a very small amount compared to the amount necessary, and this was also the translation chosen by the translator, also a set phrase in Portuguese.

In Fig. 4, we see the narrator putting books into a plastic book and concluding with this remark, which emphasizes the fact that this little box carries only a small part of all that has been written on the Holy Grail.

Figure 4. Print screen of "a spit in the ocean".

The fifth example, *one hell of* functions as a qualifying expression meaning *a great*, *awesome* or similar idea. The Portuguese translation followed another path and changed the expression into a paraphrase that attempted to replicate the same significance – *A coisa promete...* However, it would be possible to use a set phrase, though it could be considered slangish (again the issue of language registers), such as *e peras*, which can mean both interesting and difficult or tricky.

Fig. 5 illustrates the narrator preparing himself to start a journey, by looking to the back mirror and the side mirrors, and by admiting that it would a both interesting and difficult journey.

Figure 5. Printscreen of "one hell of (a journey)".

Finally, *follow the footsteps* means to do the same job or the same things in your life as someone else, which is exactly what is expressed by the Portuguese equivalent that seems to lose a bit of the idiomaticity of the original. Fig. 6 exemplifies this

replication of one's footsteps because the narrator is redoing the main character's steps in "The Da Vinci Code".

Figure 6. Print screen of "follow the footsteps".

To conclude, through the analysis of these set phrases and their translations into the subtitled version, we could claim that in all cases, with the exception perhaps of *think-tank*, everything that happens on screen, before and after what is represented in the print screens, ensures the strengthening of their meaning. Thus, setting is important – the narrator in the centre of the representation of "The Last Supper"; action is also essential – the car coming out of parking lot; the pile of books; the books inside the plastic box (container); details could also have a saying – the library surrounded by bookshelves and the scholar.

7.8. Conclusions

It is our strongest belief that the work done up until now is undeniably *a drop in the ocean*, because it has only shown us *the tip of the iceberg*. We will have *one hell of a journey* in front of us and a lot of *legwork* still to do.

Firstly, these examples and all others that we might find in this subtitled documentary film will have to go through a number of tests to verify their level of frozenness or setness and their potential level of idiomaticity. Secondly, all the classifications and interpretations of the set phrases will have to be re-evaluated, not to mention the translations obtained from the subtitling, which will require further research in the area of norms and standards for AVT and contrastive linguistics.

Finally, the completion of our pilot study will be achieved when we are able to compare the translation results of the subtitling with those of the voiced-over version, so as to test one of our main hypotheses, which will shed light on the possible *footsteps* that our research *will follow*.

8. Conclusion

This piece of research examined the language used in AVT, specifically set phrases, within the subtitled and voiced-over versions of the documentary film "The Real Da Vinci Code". The basic premise for the choice of this topic was the social and cultural role that AVT plays in a country such as Portugal with such a large tradition in subtitling, though also making occasional use of dubbing and voice-over.

The theoretical framework enabled us to insert this research within the descriptive TS paradigm and the norm-oriented perspective, within the AVT polysystem, namely interlingual subtitling and the Portuguese context. At the same time, we focused on the field of AVT, especially the difference between subtitling and revoicing (dubbing and voice-over), their advantages and disadvantages and the translation strategies they typically embody (either domestication or foreignization), as well as the importance of understanding the issue of audiovisual genres and their text types.

Moreover, we dealt with the question of word combinations, distinguishing between free and restricted lexical co-occurrence and exploring further the concepts and numerous definitions of both collocations and idiomatic expressions according to the perspective of a number of authors in the field of Lexicogrpahy and Phraseology.

The research methodology introduced our study into the field of case study to be complemented, further on, with corpus studies. Our pilot study was presented within this context, explaining the background in which our material was obtained and its main characteristics, and clarifying the research questions and the hypotheses it allowed us to design. Afterwards, some examples in English and in Portuguese were taken from the documentary film and analyzed under the perspective of classifying them as collocations or idiomatic expressions and of contrasting the English original with the Portuguese translations, weighing also the significance that events on screen might have on the ultimate effect on viewers.

Our discussion enabled us to conclude that set phrases are extremely common in audiovisual texts, as well as in general language, represented by our prototypical case, a documentary film done on the best-selling book by Dan Brown. Apart from this, it was also possible to confirm that most set phrases were maintained with almost "perfect" or close equivalents in Portuguese, whereas others were either mistranslated or ignored and omitted, due to the constraints imposed by the language itself or by subtitling.

Moreover, it was also possible to conclude that all of the given examples were strengthened by what was going on in the image, in terms of the action, the setting or details, thus taking advantage of the richness of an audiovisual text to support the effect of the set phrase in question.

Nonetheless, this minor dissertation has allowed us to be aware of the amount of work still to be done, namely in further developing the issue of frozenness and idiomaticity, since we still feel the need to clarify these concepts and design idiomaticity scales and reliable tests to verify if they are really instances of setness. It will also be necessary to continue researching in the area of AVT, especially in terms of audiovisual genres and in the definition of documentary films, so as to permit a more thorough description of "The Real Da Vinci Code" and of other documentaries to be collected.

This further collection of documentaries will be the starting point for the corpus study to be developed within the Corpógrafo, an integrated suite of online tools produced by Linguateca (www.linguateca.pt/corpografo) and *Pólo CLUP* (Faculty of Letters of the University of Porto), eventually with view to helping design a method for automatic extraction of set phrases.

References

Agrifoglio, Marjorie. 2004. "Sight translation and interpreting: a comparative analysis of constraints and failures". In *Interpreting* 6:1. Amsterdam and Philadelphia: John Benjamins Publishing. 43-47.

Alonso Ramos, Margarita. 1989. "Aproximación a um nuevo modelo lexicográfico: *El Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques de I. Mel'uk*". In *Verba*. Anuario Galego de Filoloxía, vol. 16. 421-450.

Araújo, Vera Lúcia Santiago. 2004. "To be or not to be natural: clichés of emotion in Screen translation". In *Méta – Traduction Audiovisuelle*. Vol. 49. No. 1. April 2004. Montreal: Les Presses de l'Úniversité de Montreal.

Baker, Mona ed. 1998. *Routledge Encyclopedia of Translation Studies*. London and New York: Routledge.

Benveniste, Émile. 1989. *Problemas de Lingüística Geral*. vol. 2. Campinas and São Paulo: Pontes.

Bergenholtz, Henning and Sven Tarp. 1995. *Manual of Specialised Lexicography*. Amsterdam and Philadelphia: John Benjamins Publishing.

Bosque, Ignacio and Violeta Demonte. org. 2000. *Gramática Descriptiva de la Lengua Española – Entre la oración y el discurso. Morfologia*. vol. 3. Madrid: Espasa.

Carvalho, José Herculano de. 1979. *Teoria da Linguagem. Natureza do Fenómeno Linguístico e a Análise das Línguas*. Vols. 2. Coimbra Editora: Atlântida.

Chaume, Frederic. 2003. *Doblatge i subtitulació per a la TV*. Eumo Editorial: Vic.

Cormac, Earl Mac. 1985. *A Cognitive Theory of Metaphor*. Cambridge and Massachusetts: The MIT Press.

Corpas Pastor, Gloria. 1995. *Un estudio paralelo de los sistemas fraseológicos del inglés y del español*. (Doctorate thesis at the Universidad Complutense de Madrid). Málaga: Secretariado de Publicaciones de la Universidad de Málaga.

Coseriu, Eugenio. 1977. *Principios de semántica estructural*. Madrid: Gredos.

Creswell, John. 1994. *Research Design: Qualitative and Quantitative Approaches*. Thousand Oaks and London and New Delhi: Sage Publications.

Curse, Alan. 2000. *Meaning in Language: an Introduction to Semantics and Pragmatics*. New York: Oxford University Press.

de Linde, Zoé and Neil Kay. 1999. *The Semiotics of Subtitling*. Manchester: St Jerome Publishing.

Díaz-Cintas, Jorge. 2001. *La Traducción Audiovisual – El Subtitulado*. Salamanca: Ediciones Almar.

Diccionario Enciclopédico Espasa. 1996. Madrid: Espasa.

Diccionario Enciclopedico Universal. 1986. Leon: Editorial Everest S.A.

Duro, Miguel. ed. 2001. *La Traducción para el Doblage y la Subtitulación*. Madrid: Ediciones Cátadra.

Enciclopédia Verbo Luso-Brasileira de Cultura. 2000. Edição séc. XXI. Lisboa and São Paulo: Verbo.

Fernando, Chitra. 1996. *Idioms and Idiomaticity*. Oxford: Oxford University Press.

Fonseca, Joaquim. 1981. *Coesão em Português: Semântica – Pragmática – Sintaxe* (Doctorate Thesis), Porto: FLUP.

Fromkin, Victoria and Robert Rodman. 1993. *Introdução à Linguagem*. Coimbra: Livraria Almedina.

Fuentes, Adrián. 2001. *La recepción del humor audiovisual traducido: estudio comparativo de fragmentos de las versions doblada y subtitulada al español de la película* 'Duck Soup'*, de los Hermanos Marx*. PhD thesis. Granada: University of Granada.

Gambier, Yves guest ed. 2003. *Screen Translation*. Special Issue of The Translator – Studies in Intercultural Communication (9), no. 2. Manchester: St Jerome Publishing.

Gambier, Yves and Henrik Gottlieb. eds 2001. *(Multi)Media Translation – Concepts, Practices, and Research*. Amsterdam and Philadelphia: John Benjamins Publishing.

Gómez Capuz, Juan. 1994. "Calcos sintácticos, fraseológicos y pragmáticos en los doblajes del inglés al español". In *Actas del XII Congreso Nacional de la Asociación Española de Lingüística Aplicada*. Barcelona: Universidad Autónoma de Barcelona. 171-178.

Gottlieb, Henrik. 1997. *Subtitles, Translation and Idioms*. Copenhagen: University of Copenhagen. PhD Thesis.

Grand Larousse Universel. 1995. Paris: Larousse.

Grande Enciclopédia Portuguesa e Brasileira. No date. Lisboa and Rio de Janeiro: Editorial Enciclopédia Limitada.

Hartmann, Reinhard Rudolf Karl. 1983. *Lexicography: Principles and Practice*. London: Academic Press.

Haegeman, Liliane and Jacqueline Guéron. 1999. *English Grammar: A Generative Approach*. Oxford and Boston: Blackwell Publishers.

Heid, Ulrich. 2001. "Collocations in Sublanguage Texts: Extraction from Corpora" In *Handbook of Terminology Management*. vol. 2. Wright, Sue Ellen and Gerhard Budin ed. Amsterdam and Philadelphia: John Benjamins Publishing. 788-808.

Hermans, Theo. 1999. *Translation in Systems – Descriptive and System-oriented Approaches Explained*. Manchester: St Jerome Publishing.

Honeck, Richard & Robert Hoffman ed. 1980. *Cognition and Figurative Language*. USA: Lawrence Erlbaum Associates.

Hudson, Grover. 1999. *Essential Introductory Linguistics*. Oxford: Blackwell Publishers.

Iriarte Sanromán, Álvaro. 2001. *A Unidade Lexicográfica: Palavras, Colocações, Frasemas, Pragmatemas*. Braga: Universidade do Minho and Centro de Estudos Humanísticos.

Irsula, Jesus. 1992. "Colocaciones sustantivo-verbo" In *Estudios de lexicología y metalexicografía del español actual*. Wotjak ed. Lexicographica. Series Maior, 47. Tübingen: Max Niemeyer Verlag. 159-167.

Ivarsson, Jan. 1992. *Subtitling for the Media – A Handbook of an Art*. Stockholm: Transedit.

Kövecses, Zoltán. 1986. *Metaphors of Anger, Pride and Love*. Amsterdam and Philadelphia: John Benjamins Publishing.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago and London: University of Chicago Press.

Lyons, John. 1985. *Language and Linguistics: an introduction*. Cambridge: Cambridge University Press.

Lyons, John 1995. *Linguistic Semantics. An Introduction*. Cambridge: Cambridge University Press.

Luyken, Georg-Michael *et al*. 1991. *Overcoming language barriers in television – Dubbing and subtitling for the European audience*. Manchester: European Institute for the Media.

Mateus, Maria Helena Mira *et al*. 1994. *Gramática da Língua Portuguesa*. Lisboa: Caminho.

Mel'čuk, Igor. 1995. "Phrasemes in Language and Phraseology in Linguistics" In *Idioms: Structural and Psychological Perspectives*. Everaert, M. *et al.* ed. Hillsdale, New Jersey and UK: Lawrence Erlbaum Associates. 167-232.

Mel'čuk, Igor. 1998. "Collocations and Lexical Functions" In *Phraseology. Theory, Analysis and Applications*. Cowie ed. Oxford: Clarendon Press. 23-53.

Munday, Jeremy. *Introducing Translation Studies – Theories and Applications*. London and New York: Routledge

Neves, Josélia. 2007. *Guia de Legendagem para Surdos. Vozes que se vêem*. Leiria: Instituto Politécnico de Leiria & Universidade de Aveiro.

Orero, Pilar ed. 2004. *Topics in Audiovisual Translation*. Amsterdam and Philadelphia: John Benjamins Publishing.

Pieper, Katrin. *Censorship in subtitling during the Estado Novo in Portugal*. (forthcoming).

Pöchhacker, Franz. 2004. *Introducing Interpreting Studies*. London and New York: Routledge.

Pottier, Bernard. 1978. *Linguística Geral. Teoria e Descrição*. Rio de Janeiro: Presença.

Rey-Debove, Josette. 1971. *Étude Linguistique et sémiotique dês dictionnaires français contemporains*. The Hague and Paris: Mouton.

Schröter, Thorsten. 2005. *Shun the pun, rescue the rhyme? The dubbing and subtitling of language-play in film*. Doctoral thesis. Karlstad: Karlstad University.

Seco, Manuel. 1987. *Estudios de lexografía española*. Madrid: Paraninfo.

Shuttleworth, Mark and Moira Cowie. eds 1999. *Dictionary of Translation Studies*. Manchester: St Jerome Publishing.

Strässler, Jürg. 1982. *Idioms in English – a Pragmatic Analysis*. Tübingen: Gunter Narr.

The Oxford English Dictionary. 1998. Oxford: Clarendon Press.

Toury, Gideon. 1995. *Descriptive Translation Studies and beyond*. Amsterdam and Philadelphia: John Benjamins Publishing.

Venuti, Lawrence. 1995. *The Translator's Invisibility*. London and New York: Routledge.

Venuti, Lawrence ed. 2001. *The Translation Studies Reader*. London and New York: Routledge.

Vilela, Mário. 1979. Estruturas Léxicas do Português. Coimbra: Almedina.

Vilela, Mário. 1994. *Estudos de Lexicologia do Português*. Coimbra: Almedina.

Wierzbicka, Anna. 1985. *Lexicography and Conceptual Analysis*. Ann Arbor and MI: Karoma.

Williams, Jenny & Andrew Chesterman. 2002. *The Map – A Beginner's Guide to Doing Research in Translation Studies*. Manchester: St. Jerome Publishing.

Wright, Jon. 1999. *Idioms organiser – organised by metaphor, topic and key Word*. London: LTP.

Zgusta, Ladislav *et al*. 1971. *Manual of Lexicography*. The Hague and Paris: Mouton de Gruyter.