

## **A Multimedia Database for the Training of Audiovisual Translators**

**Cristina Valentini**  
**University of Bologna at Forlì, Italy**

### **ABSTRACT**

This paper sets out to describe a multimedia database, the Forlì Corpus of Screen Translation, developed at the University of Bologna's Department of Interdisciplinary Studies in Translation, Languages and Culture (SITLeC), emphasizing the advantages it offers for audiovisual translator training. The database is actually a corpus of 20 original and dubbed films, fully indexed on the basis of a set of predefined linguistic, cultural and pragmatic categories. The tagging of the corpus makes it possible to extract concordances in the form of transcripts of film dialogues along with the original audiovisual scenes. Thanks to these features, the database is a valid tool to be used as a component in the training of audiovisual translators, in that it can help develop traditional linguistic, but also communicative and cultural skills. In addition, from a methodological point of view, it offers an ideal basis on which to ground empirical and quantitative research, thus helping establish a scientific approach in the new discipline of audiovisual translation studies.

### **KEYWORDS**

Multimedia database, AVT teaching, professional training, corpus linguistics, empirical research

### **BIOGRAPHY**

Cristina Valentini got her degree in Conference Interpreting from the SSLMIT (University of Bologna), in 2001, with a dissertation based on a survey conducted among simultaneous interpreters to assess the use of PCs in the booth. In 2003, after having completed a traineeship at the SDT of the European Commission, she became research assistant at the SITLEC (University of Bologna) at Forlì, specializing in the development of terminology and multimedia databases. Her research interests currently focus on multimedia corpora, dealing particularly with polysemiotic aspects of film translation and film discourse, and on technical and legal terminology pertaining to the specialised domain of health and safety at work.

E-mail: [CValentini@sslimit.unibo.it](mailto:CValentini@sslimit.unibo.it)

## **1. Introduction**

The last decade has witnessed a massive growth in the creation and use of corpora, which have arguably become the necessary hallmark of all scientific linguistic analysis. Electronic corpora nowadays provide the basis for empirical research in translation-based studies, with positive repercussions which have long been discussed in the literature from a theoretical, practical and pedagogic point of view (Baker, 1995; Aston, 2001; Zorzi, 2001; Zanettin, 2001; Ulrych, 2001).

While the potential usefulness of corpora for various fields of translation – e.g. technical, scientific, literary, legal – has been extensively investigated in recent years, multimedia translation scholars have not yet adequately focussed their attention on corpus linguistics tools and methods. The only experience in this direction, to our knowledge, comes from systemic-

functional linguistics, and in particular, multimodal analysis theory (Thibault, 2000; Baldry & Thibault, 2001; Taylor, 2003; Taylor, 2004). The investigation of multimodality in 'film texts' has provided the basis for developing a web-based system,<sup>1</sup> whose main aim is "the study of the synchronisation between the meaning-making resources deployed in these texts" (Baldry, 2004). Although this application offers an empirical basis for comprehensive description and theoretical analysis of the different semiotic modalities involved in multimedia texts, in its current form it does not allow contrastive study of multimedia products in general, and films in particular. As such, audiovisual translation has until now mainly relied upon the contribution of individual scholars, most of whom have adopted a case-study approach (Gambier & Gottlieb, 2001).

In 2003, in response to the need for more empirical studies, a research group on film translation at the University of Bologna's Department of Interdisciplinary Studies in Translation, Languages and Cultures (SITLeC), made up of Christine Heiss, Marcello Soffritti and Cristina Valentini set out to design a textual and audiovisual database for the collection and study of film translation data. This corpus, called Forlixt 1 (Forlì Corpus of Screen Translation), is part of a more general ongoing project that involves the study of other AVT topics and modalities, such as interlingual and intralingual subtitling and quality of dubbed products (Rundle, 2000; Chiaro, 2004; Antonini & Chiaro, 2005 and forthcoming).

For the purposes of the present paper, Forlixt 1 will be the starting point for reviewing current needs and challenges of audiovisual translator training, a field within which this resource can help develop a number of specific skills traditionally situated at the crossroads of written and oral competence. The definition of the hybrid nature of audiovisual translation (AVT) will also help make the case for pointing out specific training needs and skills required of translators in this particular field.<sup>2</sup> A description of corpus construction methods and details of the conceptual architecture, concordancing system and browsing utilities will then be given, with an indication of potential end-users. In conclusion, limits and benefits of using a multimedia database of this kind as a component of AVT training will be discussed.

## **2. Forlixt 1: the Forlì Corpus of Screen Translation**

Forlixt is presently composed of 11 Italian and 9 German films, including the complete transcription of the dialogues of each film.<sup>3</sup> The corpus currently amounts to approximately 32 hours of fully transcribed audiovisual material (equivalent to about 200,000 words) stored in a SQL dedicated server.<sup>4</sup> The database is essentially a bilingual, parallel and comparable corpus.

By 'parallel corpus' we mean a collection of texts in language A and their translations into language B, C, D (Teubert, 1996). As a matter of fact,

although Forlixt presently contains only products from Italian and German-speaking countries, the software tool has been designed in such a way as to be able to deal with an indefinite number of translations (dubbed versions) from source language A. Moreover, following Teubert's definition, Forlixt comprises scenes of films and transcripts aligned on a 'scene-by-scene' basis, so that sequences can be retrieved along with their translation, not only from language A into language B but also from language B into language A. Forlixt is thus a multidirectional database. Finally, and again in accordance with Teubert's classification, it can also be defined as 'comparable' in that it is made up of original films in both languages, A and B, which can be ascribed to the same film genre, currently comedy, and are therefore similar in terms of their content and narrative structure.

Forlixt has been designed first and foremost as an archive for storing films and their (tran)scripts, making them accessible both in their original (linear) form and through an index of pre-established categories, in order to satisfy academic research needs on screen translation. The corpus is representative of the linguistic variety characterising traditional "written to be spoken as if not written" film discourse (Gregory & Carroll, 1978) and translated discourse, which can take the form of either a peculiar spoken (dubbing, audio-description, media interpreting) or written (subtitling) variety.<sup>5</sup>

### **3. Query utilities**

#### **3.1 Free-text search**

Two search systems have been designed: a free-text and a guided search tool. The free-text modality enables the user to look for words or strings of words, i.e. a linear sequence of words as written. Queries can be extracted by exact match, selecting the option 'full word', or by approximate match, selecting the option 'any part of the word'. Queries can also be restricted by sub-domain, namely selecting one of the given options: dialogues or subtitles or dialogues and subtitles.

The software provides results in the form of a list of occurrences of the word or string of words queried, accompanied by its textual context, namely the complete line in which the query word appears (Figure 1). In addition, for each occurrence it also provides some general information concerning the film: whether it is an original or dubbed version, the film title, the name of the character uttering the line, and the language of the film. Finally, this page contains the hypertextual link to the scene associated with each line and a list of categories attributed to the scene that can be browsed via the guided search to observe specific phenomena.

Figure 1 Free-text search page

Ricerca libera su dialoghi e sottotitoli

Cane

Cerca in:  Dialoghi e Sottotitoli  Dialoghi  Sottotitoli | Tipo di ricerca:  Parte di testo  Parola intera  
(La ricerca non fa distinzione tra maiuscole e minuscole)

Battute trovate: 27

Risultati visualizzati 1-10 (di 27)  
Prima « Precedente - Successiva » Ultima | Pagine: 1 2 3

Film: "Caro Diario", ita , versione originale  
Battuta: dialogo , ita , personaggio "Moretti Nanni"  
Testo: *D'estate a Roma i cinema sono tutti chiusi, oppure ci sono film come "Sesso, amore e pastorizia", "Desideri bestiali", "Biancaneve e i sette negri", oppure qualche film dell'orrore come "Henry", oppure qualche film italiano.*  
Visualizza la sequenza:  
01:00.10-02:25.59 (01:25.49) [Humour verbale, Libri, film, opera, riviste e programmi televisivi]

Film: "Caruso Pascoski (di padre polacco)", ita , versione originale  
Battuta: dialogo , ita , personaggio "Caruso"  
Testo: *La mortadella è buonissima, un c'è niente da fare, è proprio buona. La mortadella è comunista, il salame socialista, il prosciutto è democristiano, la coppa, liberale, le salcicce repubblicane, il prosciutto cotto è fascista.*  
Visualizza la sequenza:  
28:36.00-29:34.47 (00:58.47) [Insulto, Scherno, Figure di parola, Humour verbale, Nomi di istituzioni, associazioni, enti, partiti, Nomi di cibi e bevande]

Film: "Caruso Pascoski (di padre polacco)", ita , versione originale  
Battuta: dialogo , ita , personaggio "Caruso"  
Testo: *C'ho il cane grosso. Lei come sta?*  
Visualizza la sequenza:  
29:34.48-30:09.12 (00:34.64) [Saluto, Scherno, Allocutivi, Titoli onorifici e professionali]

Film: "Funny Games", ita (film originale "" , )  
Battuta: dialogo , ita , personaggio "Anna"  
Testo: *Lascia chiuse quelle che danno sul lago altrimenti entra il caldo! Dai, smettila, ora ti do qualcosa ma prima devo mettere a posto. Potresti chiamare il cane? Non mi fa combinare niente!*  
Visualizza la sequenza:  
05:30.16-07:02.95 (01:32.79) [Ordine, Rimprovero]

Visualizza la sequenza:  
28:36.00-29:34.47 (00:58.47) [Insulto, Scherno, Figure di parola, Humour verbale, Nomi di istituzioni, associazioni, enti, partiti, Nomi di cibi e bevande]

Film: "Caruso Pascoski (di padre polacco)", ita , versione originale  
Battuta: dialogo , ita , personaggio "Caruso"  
Testo: *C'ho il cane grosso. Lei come sta?*  
Visualizza la sequenza:  
29:34.48-30:09.12 (00:34.64) [Saluto, Scherno, Allocutivi, Titoli onorifici e professionali]

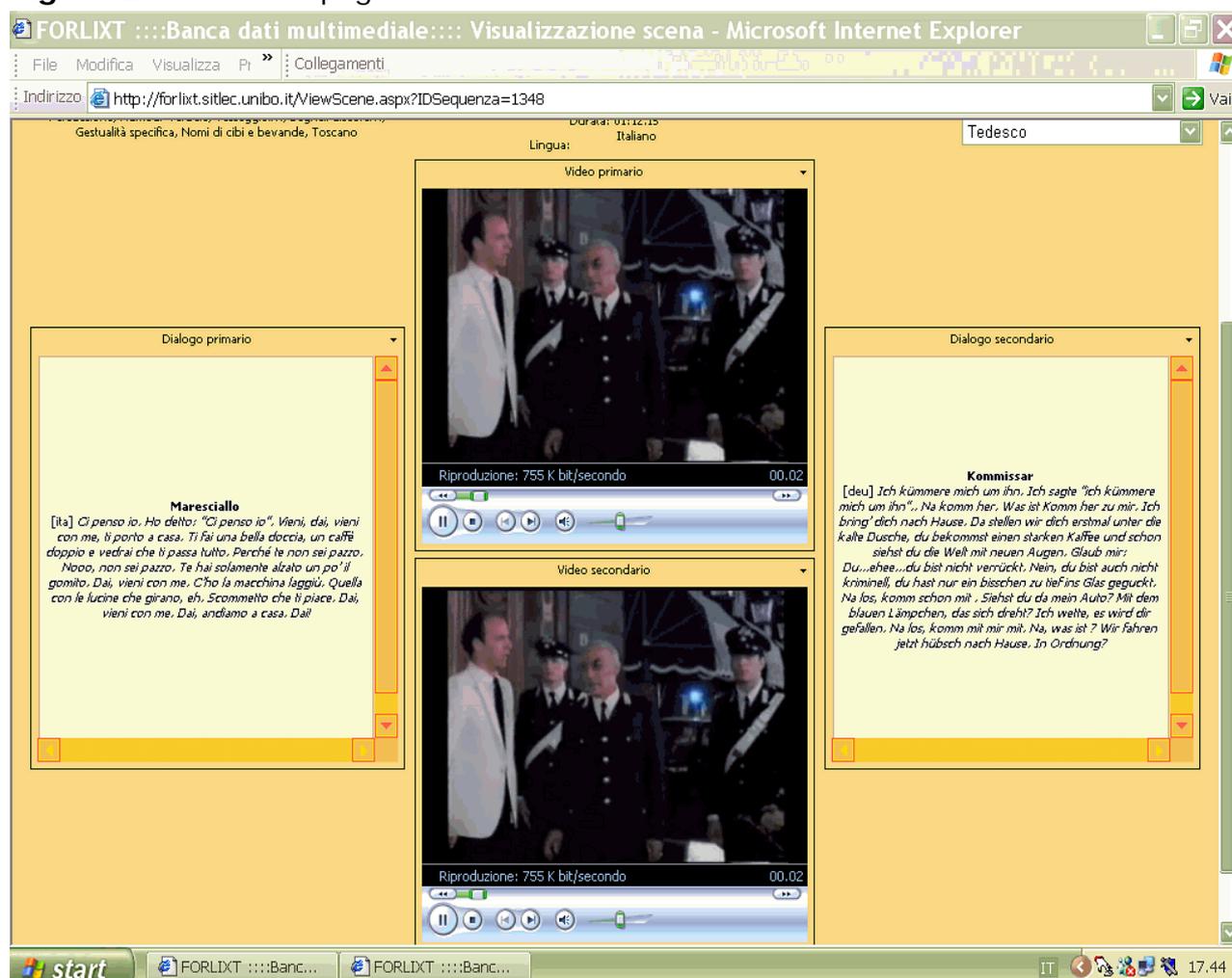
Film: "Funny Games", ita (film originale "" , )  
Battuta: dialogo , ita , personaggio "Anna"  
Testo: *Lascia chiuse quelle che danno sul lago altrimenti entra il caldo! Dai, smettila, ora ti do qualcosa ma prima devo mettere a posto. Potresti chiamare il cane? Non mi fa combinare niente!*  
Visualizza la sequenza:  
05:30.16-07:02.95 (01:32.79) [Ordine, Rimprovero]

Film: "Funny Games", ita (film originale "" , )  
Battuta: dialogo , ita , personaggio "Georgi"  
Testo: *Non è detto. L'anno scorso, quando non in D-85, Ermete un fondo con a un mi Ermete il kama. Ti ricordi?*

start FORLIXT :::Banc... 16.23

A new page (Figure 2) allows the user to play the video scene including the hit line queried and to compare it with its transcription. A bilingual comparative analysis can be made by selecting a target language from the language menu bar. The software retrieves the parallel text in the language selected, together with the associated scene. It also makes it possible to browse other scenes from the film, by clicking on the next or previous arrows on top of the first video.

Figure 2 View scene page



### 3.2. Guided search

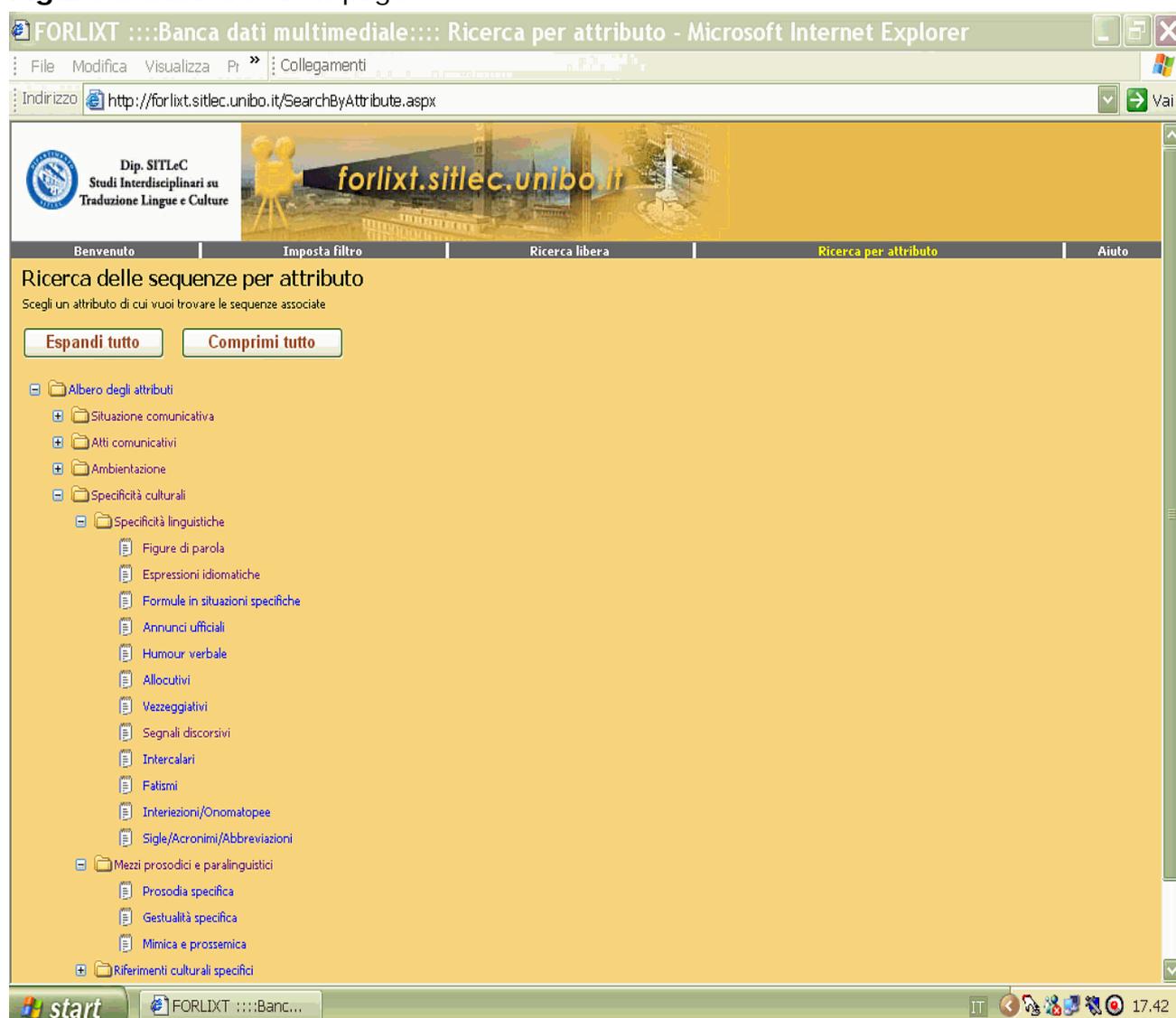
The second type of query tool is the guided search utility that will guide the user through the corpus according to a hybrid set of pre-established linguistic-pragmatic criteria inherited from previous research conducted on AVT products in the field of descriptive linguistics, translation theory and pragmatics (Heiss, 2000; Heiss, 2004; Pavesi 1994; Baccolini *et al.*, 1994; Bazzanella, 1994). These categories correspond to the list of attributes used for tagging the corpus. Textual and audiovisual material has been segmented into a number of scenes and subsequently annotated using this list, making it possible to focus queries according to the user's specific needs. In particular, categories have been clustered into groups:

- ❑ Pragmatic categories (communicative situations, communicative acts).
- ❑ Encyclopaedic categories (cultural, temporal and geographical setting).

- Linguistic-cultural categories (linguistic specificities, prosodic and paralinguistic means, specific cultural references, names of specific entities).
- Linguistic varieties (jargon, dialect, LSP, register, etc.).<sup>6</sup>

Each of these groups actually represents a macro-category or, as it is called by software engineers, a node to which an unlimited number of leaves or attributes can be assigned. For instance, the category 'names of specific entities' comprises labels such as names of famous characters and people, titles and names of cultural or trade products. The macro-category 'linguistic specificities' includes labels such as figures of speech, idioms and verbally expressed humour. There is also one category for specific film-related contents, so that a filter can be applied to restrict the corpus by selecting language, director, film genre and actor (Figure 3).

**Figure 3** Guided search page



#### 4. Specific aspects of audiovisual translation

Many scholars have tried to define the concept behind the term audiovisual translation and, in this effort, most of them have eventually ended up with the enumeration of a list of features characterising this particular type of translation, and with a list of professional activities - e.g. interlingual/intralingual subtitling, audio description, surtitling, dubbing, voice-over, etc. (Gambier, 2003) - whose common denominator is to be found in various forms of 'audiovisual' support that they have to consider.

The term 'audiovisual' thus implies at least two channels of communication, namely the acoustic medium (speech and soundtrack) and the visual medium (images, gestures of characters, facial expressions, etc.). This means taking into account the various ways in which a number of distinct semiotic resource systems are co-deployed in the creation of a polysemiotic text. Consequently, the linguistic code cannot be regarded as more important than other semiotic resources involved in the creation of meaning, which is the result of the multiplication, rather than the addition, of the variety of ways in which different classes of phenomena – words, actions, objects, visual images, sounds and so on – are related to each other (Taylor, 1999).

Audiovisual translators are therefore confronted with a twofold challenge of deconstruction and reconstruction of a highly complex semiotic system comprising a number of different codes, ranging from linguistic, paralinguistic and musical items to special effects, iconography, photography, planning, graphics and syntax (Chaume, 2004). Accordingly, what should mainly interest the translator is to convey these signs, both linguistic and non-linguistic. To do so, it is important that AVT teachers provide trainees with authentic material for contrastive analysis of both source (original) and target (translated) texts. This makes it possible to preserve the 'original entity' of the film material, adopting the so-called *in vivo*, as opposed to the classical transcription-based *in vitro*, approach (Baldry, 2004: 24).

Though audiovisual translation comprises a vast array of professional activities, often involving technical specialisation, Gambier (2004:1) highlights three main aspects worthy of attention in audiovisual translation:

- The relationship between images, sound and speech.
- The relationship between one or more source foreign language(s) and the target language.
- The relationship between the oral and written code.

The first and third points require further comment, in relation to the characteristics of film dialogues which are the object of this peculiar form

of linguistic transfer. Film dialogue is in most cases a particular form of “written to be spoken as if not written” language variety, according to the already classic definition of Gregory and Carroll (1978). As a result, it is essentially a hybrid variety, written by scriptwriters to reproduce spontaneous speech and natural face-to-face conversation. The success or failure of this attempt depends on the power of observation of filmmakers, but also on their idiosyncrasies and intended audience impact.<sup>7</sup> Be this as it may, the final product should by and large reflect the norms and rules characterising oral speech. The resulting scripts are then performed by actors. In this process, they are often changed and adapted to the constraints of ‘real’ interactions and further enriched with an array of paralinguistic features (mimics, body language, prosody), which help construct the full meaning of the film. Audiovisual translators should thus try to study the relationship between the verbal code (e.g. oral discourse markers) and the paralinguistic code conveyed through images and sound (e.g. mimics, gestures, prosody), which in many cases help contextualise and disambiguate certain words or expressions as well as deictic references.

## **5. Audiovisual Translators’ Skills**

Drawing on these theoretical assumptions, a list of general skills is presented, which, in our view, can be regarded as the primary objectives of any academic course in audiovisual translation, whatever the specific modalities taken into consideration (e.g. dubbing, subtitling, voice-over, audio description, free commentary, etc.):<sup>8</sup>

- (1) Linguistic competence.
- (2) Pragmatic, communicative and interactional competence.
- (3) Paralinguistic competence.
- (4) Cultural (encyclopaedic) competence.
- (5) Technical competence.

The importance of a linguistic pre-analysis of all film material has been recently emphasised in Italy, both by organisers of professional courses and by film adapters (Di Fortunato & Paolinelli, 1996; Paolinelli, 2004). Correct emphasis needs to be placed on linguistic skills, with particular attention to analysing syntactic features of oral discourse in various languages, e.g. use of cleft sentences, topicalisation and discourse markers. In addition, students should be made aware of different social and register varieties, and of their specific use in definite context-driven communicative situations. This means fostering knowledge of ritual and conventional formulae, as well as of sociolectal and regional varieties, which, along with pronunciation and prosody, are often exploited in films to produce expressive and comic effects (Heiss, 2000a).

Partly overlapping this linguistic competence, the development of pragmatic, communicative and interactional skills is also a basic

requirement for future audiovisual translators. This competence is fruitfully enhanced through the study of specific linguistic elements (discourse markers, fillers, etc.), culture-specific communication rules and elements pertaining to the paralinguistic dimension of face-to-face interaction. The latter is regarded by some scholars (Herbst, 1987 and 1994; Pavesi, 1994) as one of the most difficult and vital aspects for the successful transposition of a film's semiotic content.

A further added value for the audiovisual translator comes from the development of paralinguistic competence, focusing on mimics, prosody, gestures and behavioural patterns as well as on their interaction. As indicators of cultural diversity, their interpretation and examination is useful to the training of the translator as a *Kulturmittler*, i.e. an intermediary between two cultures (Heiss, 2000b: 189). In particular, dubbing is especially concerned with the development of this specific competence, in that lip-synchrony is very closely related to mimics and gestures (Herbst, 1994).

Given that films are essentially culture-specific products, particular attention should also be given to fostering cultural and encyclopaedic knowledge in the trainee translator. The analysis of cultural references in films can be carried out at different levels: they can be overtly expressed in words (names of places, food and drinks, famous characters and people, cultural and trade products, etc.), embedded in discourse (quotations, aphorisms, metaphors, cultural stereotypes, etc.) or conveyed through gestures and behaviours of characters, photography, icons and writings displayed through images (Antonini & Chiaro, forthcoming; Baccolini *et al.*, 1994).

Last but not least, if we think of screen translation and its many modalities, the development of a certain degree of basic technical competence should not be overlooked. In particular, trainees must be provided with a set of technical skills involving knowledge of film editing, isochrony, i.e. the attempt to attain an equivalent duration of source text utterances and the utterances of the target text (Chaume 2004), as well as lip-synchronisation techniques. As far as dubbing is concerned, this aspect generally goes beyond the scope of academic teaching settings and is usually learnt in a practical way, through experience in dubbing studios. Nonetheless, the possibility of accessing different linguistic versions of films in their aural and visual entirety can help focus trainees' attention on kinetic movements of actors on stage and thus help provide a theoretical background for their subsequent practical experience.

## 6. Examples

Some benefits pertaining to the use of multimedia corpora in general and Forlixt 1 in particular, as a useful aid in academic translation teaching settings will now be discussed, providing examples taken from films in the

database. These examples will specifically aim to show how audiovisual support can help contextualise and successfully interpret the meaning of discourse, demonstrating that video often provides vital clues to the identification and understanding of cultural and pragmatic references.

### Example 1

Film: <i>Mimi metallurgico ferito nell'onore</i> <sup>9</sup>	
Ispettore: Giovanotto! Giovanotto! Ci dica com'è andata! Ci racconti bene i fatti. Eh, eh... l'assassino lei l'ha visto in faccia. Per forza! No!?	Inspector: Hey, man! Man! Tell us what happened! Tell us exactly what happened. You saw the killer directly in the face. No doubt! Didn't you?
Mimì: Eh no! Purtroppo non so niente. Nun ve pozzu dire niente. Nun vidi niente! Sono entrato e sono svenuto.	Mimì: No! I'm afraid not. Sorry to say. I don't know. I can't tell you. I didn't see anything. I walked in and I failed.
Ispettore: Ma l'assassino potrà almeno descrivercelo!	Inspector: But you should at least be able to tell us what the killer looked like!
Mimì: Nnt! Me dispiace! Nun ve pozzo aiutare. Nun sacciu niente. Nnt! Me sento poco bene.	Mimì: No I can't! I'm sorry! I can't help you. I don't know. I don't feel very well.

In this shot, after a shoot-out involving Mafia members, the inspector urges Mimì, the film's main character, to report what he has just seen. Mimì, in a typically Sicilian way, denies having seen anybody. Here the video helps the learner contextualise the communicative act of 'refusal', which is expressed in non-standard Italian. Mimì has a typical Sicilian accent and uses a regional variety of Italian. The way in which he says *no* is typical of Sicilian speech, and of Mafia slang in general. The words 'I didn't see, hear anything, I don't know anything, I can't tell you' are stereotyped references to the mafia culture of the 'code of silence', which Mimì must respect despite himself. This is an example of how translators might have to deal with a scene that is the repository of many cultural references, combining visual elements (the setting, the guns), prosody (the inspector and Mimì's marked Sicilian pronunciation), mimics and gestures (the act of refusal is emphasised by the movement of Mimì's eyes and head), together with implicit cultural references in discourse (Mimì's respect for the code of silence). If the film is to be dubbed, the translator will also be required to put a special effort into lip-synchronisation. Hence the importance of basing class discussion on the whole of the scene's audiovisual content, allowing students to focus on the paralinguistic elements which contribute to its overall meaning.

Similarly, in the following example humour can be successfully conveyed if speech is appropriately contextualised from a multimedial point of view:

### Example 2

Film: *La Stazione*

Domenico: Certo che lei è proprio alta, eh sì!	Domenico: Yes, you are very tall, indeed!
Flavia: Alta? Boh, io senza i tacchi mi sembro una specie di papera.	Flavia: Tall? I don't know, without high heels I see myself rather like a sort of duck.
Domenico: Papera? Che papera signorina? Deve vedere le papere che teniamo qua al paese, le paperissime proprio, tutte con quel...Faccio un caffè, così si sveglia un poco?	Domenico: What...a duck? A duck, signorina? You should have a look at the ducks we have here, very, very duck, all of them with that... Shall I make a coffee to wake you up?

In this scene, the close-up of the young lady's legs indicates that Domenico, the stationmaster, has been admiring them. Not knowing how to justify himself, he resorts to a euphemism and comments 'You are very tall indeed'. The humour is conveyed not only by words, but most significantly by words associated with images. The whole shot is very amusing because of the explicit body language used by the stationmaster to answer the young lady, arguing that where he lives there are real ducks, meaning 'duck-like girls', while the movement of his hands suggests that these 'duck-like girls' have large thighs. Finally, there is also a cultural reference which adds to the humour of the scene: *paperissima*, a recently coined term meaning 'very duck-like', is the name of a well-known Italian TV programme where people are shown in very awkward situations acting like ducks, i.e. very stupidly.

From a linguistic point of view, Domenico's idiolect features an accumulation of discourse markers, fillers and conversational indicators, complemented by his gestures and mimic expressions. All of these should be taken into account, since many of the dubbing difficulties connected with the specificity of local spoken varieties are often glossed over. As Nadiani (1996:1) states with reference to the German version, the result is "highly simplified spoken language produced for a mass audience which often sounds as if it is 'cleared' of the specific traits characterizing a film's cultural context".

Finally, let us take a look at the last of our examples:

### Example 3

Film: <i>La vita è bella</i>	
Giosuè: Quanti ce n'è?	Giosuè: How many kids are there?
Guido: Un vespaio. E' pieno così! Son son tutti nascosti.	Guido: It's a wasps' net, it's full of them. All hiding.
Giosuè: Babbo guarda!	Giosuè: Dad, look!
Guido : Oh! C'è un covo! Visto visto visto visto. Eliminato. Andiamo, andiamo.	Guido: Hey, it's a hideout! Picked, Picked! Picked! Picked! Out! Let's go, let's go.

Again, spatial deictic references are fully understandable only if speech is appropriately associated with the relevant scene. The images here complement the metaphor used by Guido to convince his son that the prison camp is full of children hiding everywhere. He says that there is a 'wasps' nest' of children [a crowd of children], and the following shot actually shows children coming out of nowhere, literally buzzing like wasps. The perlocutionary effect of Guido's words in the Austinian sense is not fully achieved until Giosuè sees all these children.

These examples help illustrate that an audiovisual translator's curriculum should take into account a number of aspects, all of which can give considerable help to future dubbers and subtitlers in coming to grips with the real challenges of the profession. AVT is in fact a highly structured and complex process that often obliges translators to make choices and to sacrifice certain elements in favour of others after careful examination of pros and cons.

## 8. Conclusion

To sum up, Forlixt 1 can serve a variety of uses and purposes. As an empirical tool, it offers translation students an opportunity to start to come to grips with the challenges of film translation by accessing a reservoir of examples. In addition, the database can be used by professional practitioners to look for ready-made solutions, to study the language specific to a certain period, genre, jargon or style, as is the case with traditional lemma-based corpora compiled for specialised translation purposes. Finally, scholars can exploit the database to analyse translation strategies and processes for multimedia texts, affording a privileged empirical observation post on which to base theoretical assumptions.

On the other hand, there are obviously some limits to the potential use of Forlixt 1 in specialised AVT courses. First, the database does not yet include categories to take account of specific 'adaptation' aspects, which are fundamental in order to achieve satisfactory translated products. Second, as a reservoir of ready-made solutions, it might tend to encourage crystallisation of customary translations. This is a problem which has been tackled by many scholars in the Italian context, where the disparaging term *doppiaggese* has been coined to describe the Italian used in dubbed dialogues (Pavesi, 1996). Third, the quality of a corpus depends on qualitative criteria applied in selecting materials, an issue which in turn raises the problem of translation quality in dubbed and subtitled films, and perception of this quality (Chiaro, 2004, Bucaria & Chiaro, forthcoming). Finally, the still unaddressed issue of copyright must be considered.<sup>10</sup>

Notwithstanding the latter considerations, these limits may however be considered not just limits as such, but as challenges for potential future developments of the database, especially if we consider the fact that with

this instrument we have tried to provide for the first time an answer to the appeal to academics for more detailed theoretical and empirical analysis of AVT products:

Ainsi ne sont pas traités les problèmes de formation des traducteurs de l'audiovisuel, les compétences et comportements qu'ils doivent acquérir [...] Ont été également négligés, malgré quelques allusions ici et là, les impacts de la TAV sur la traductologie, avec par exemple les notions de texte, d'original, de sens, d'interprétation, de normes descriptives, d'équivalence, de skopos, d'écrit (notamment dans ses rapports à l'oral) (Gambier, 2004: 11).

[As a result, audiovisual translators' problems and particularly the issue of the competences and skills they need to acquire have not been expressly tackled [...] Similarly, despite some sporadic studies, inadequate attention has been put on the impact of audiovisual translation on translation studies, involving the notion of text, original, meaning, interpretation, descriptive norms, equivalence, skopos, written code (with particular emphasis to it in relation with the oral code)] (author's translation).

The Forlixt database, far from being a panacea for all ills, is an experimental tool which we hope will help substantiate researchers' theoretical assumptions with practical evidence, in order to achieve a certain level of theorisation of specific strategies at work in audiovisual translation and thus potentially benefit the profession as a whole.

## Acknowledgments

Special thanks go to Piero Conficoni, who is responsible for the Forlixt 1 software engineering, to Sabrina Linardi for the hard work she put into revising the transcriptions and entry data and analysing the German films, and to all the SSLiMIT students who suggested film titles and enabled us to include material they had collected in the database.

## Notes

<sup>1</sup> I refer in particular to MCA, the Multimodal Concordancing System developed by a joint research team of the Universities of Pavia and Trieste: <http://mca.unipv.it>.

<sup>2</sup> Given that the database presently includes only dubbed products, the term 'audiovisual translator' is hereby intended to refer mainly to professionals who take part in the dubbing process and whose status and role, at least in Italy, is still blurred and requires professional certification (Paolinelli 2004). However, many of the skills that will be discussed in this paper have also a bearing on the training of other translators working with audiovisuals, among which subtitlers certainly represent the most significant group. In this respect, we think that the acknowledgment of the role of institutions like universities where such training is provided is all the more important, along with the quality of the teaching and the development of tools which can successfully contribute to enhancing it.

<sup>3</sup> We opted for a plain transcription because the concordancer will always display the queried item together with its multimedial co-text, thus entirely preserving the film text in its original form.

<sup>4</sup> Since we are presently tackling the problem of copyright licenses for the use and reproduction of films, the main entry page of the system is presently located in a dedicated Internet domain of the SITLeC Department, (<http://forlixt.sitlec.unibo.it/login.aspx>) and is protected by password and its consultation is therefore forbidden to outsiders.

<sup>5</sup> For a more comprehensive overview of corpus construction methods and steps see Valentini (forthcoming).

<sup>6</sup> For a more detailed description of the categories applied, see Heiss (2004) and Heiss & Soffritti (2005).

<sup>7</sup> Taylor (1999) offers a list of provisos which are thought to constrain the representation of authentic dialogue in written scripts.

<sup>8</sup> For a more exhaustive reference about screen translators' skills and competences as well as AVT aids, see Heulwen *et al.* (1996), Gambier & Gottlieb (2001), Díaz Cintas & Orero (2003).

<sup>9</sup> The English translation for all the examples cited in this paper is not the official one, but the author's. It is meant to simply support the understanding of the Italian example.

<sup>10</sup> In this respect, it should be pointed out, however, that streaming technology used to retrieve videos in our corpus allows playback of sound or video without downloading the entire resource file in advance, thus making it possible to partly comply with copyright legislation.

## References

- Antonini, R. and D. Chiaro. 2005. "The Quality of Dubbed Television Programmes in Italy: the experimental design of an empirical study." Marina Bondi e Nick Maxwell, (eds.) *Cross-Cultural Encounters: Linguistic Perspectives*. Roma: Officina Edizioni, 35-44.
- Antonini, R. and D. Chiaro. Forthcoming. "The General Perception of Likelihood of Occurrence: the translational norms of language-specific features on the Italian screen", in Anderman, G. and J. Diaz-Cintas (eds.) *Audiovisual Translation – Language Transfer on the Screen*. Basingstoke: Palgrave Macmillan.
- Aston, G. (ed.) 2001. *Learning with Corpora*. Bologna: CLUEB.
- Baccolini, R., R.M. Bosinelli and L. Gavioli (eds.) 1994. *Il doppiaggio: Trasposizioni linguistiche e culturali*. Bologna: CLUEB.
- Baker, M. 1995. "Corpora in translation studies: an overview and some suggestions for future research". *Target* 7: 223-243.
- Baldry, A. 2004. "Concordancing film corpora: the role of MCA in multimodal discourse analysis", in Taylor, C. *et al.* (eds.) *Proceedings of the Seminar, LinguatEL: Progetto telematico di lingua e mediazione interlinguistica*. Padova: Unipress, 19-32.
- Baldry, A. and M. Beltrami. 2004. "Report: The MCA system: its role in the development of multimodal corpus linguistics". *DIDACTAS: Didattica dell'analisi testuale, dei corpora, della traduzione e della sottotitolazione*. [www.univ.trieste.it/~didactas/pagine/relazione.php?id=20041009084724](http://www.univ.trieste.it/~didactas/pagine/relazione.php?id=20041009084724).
- Baldry, A. and J.P. Thibault. 2001. "Towards multimodal corpora", in Aston, G. and L. Bernard (eds.) *Corpora in the description and teaching of English*. Bologna: CLUEB, 87-102.
- Bazzanella, C. 1994. *Le facce del parlare*. Florence: La nova Italia.

- Bucaria, C. and D. Chiaro. Forthcoming. "The quality of screen translation on Italian television". *Target*.
- Chaume, F. 2004. "Film studies and Translation Studies: two disciplines at stake in audiovisual translation". *Meta* 49(1): 12-24.
- Chiaro, D. 2000. "Servizio completo. On the (un)translatability of puns on screen", in Bosinelli, R.M. *et al.* (eds.) *La traduzione multimediale. Quale traduzione per quale testo?* Bologna: CLUEB, 27-42.
- Chiaro, D. 2004. "Investigating the perception of translated Verbally Expressed Humour on Italian TV". *ESP Across Cultures* 1: 35-52.
- Di Fortunato E. and M. Paolinelli. 1996. *Barriere Linguistiche e la Circolazione delle opere audiovisive. La questione doppiaggio*. Roma: AIDAC, Associazione italiana dialoghisti adattatori cinetelevisivi.
- Díaz Cintas, J. and P. Orero. 2003. "Postgraduate Course in Audiovisual Translation". *The Translator* (9)2: 371-388.
- Gambier Y. and H. Gottlieb (eds.) 2001. *(Multi)media translation. Concepts, Practices and Research*. Amsterdam and Philadelphia: John Benjamins.
- Gambier, Y. 2003. "Screen Transadaptation: Perception and Reception". *The Translator* 9(2): 171-189.
- Gambier, Y. 2004. "La traduction audiovisuelle: un genre en expansion". *Meta* 49(1): 1-11.
- Gregory M. and S. Carroll. 1978. *Language and Situation: language varieties and their social context*. London: Routledge & Kegan Paul.
- Heiss C. 2000a. "Quanto è tedesco Mimì metallurgico. Qualità e strategie di doppiaggio in alcuni esempi di commedia all'italiana", in Taylor, C. (ed.) *Tradurre il cinema*. Trieste: University of Trieste, 59-73.  
<http://citam01.lingue.unibo.it/intralineia>
- Heiss, C. 2000b. "La traduzione filmica come pratica didattica", in Bosinelli, R.M. *et al.* (eds.) *La traduzione multimediale. Quale traduzione per quale testo?* Bologna: CLEUB, 183-196.
- Heiss, C. 2004. "Translatorische Aspekte der Filmsynchronisation Deutsch/Italienisch/Deutsch". *Moderne Sprachen* 48(2): 39-59.
- Heiss, C. In press. "Filmdialoge als zu übersetzender Text? Ein kontrastiver Ansatz Deutsch - Italienisch, Italienisch-Deutsch", in *Kontinuität, Konkurrenz und Kompetenz – Wohin will die Deutsche Romanistik? XXVIII Deutscher Romanistentag* (Kiel, 2003).
- Heiss, C. and M. Soffritti. 2005. "Parallelkorpora 'gesprochener Sprache' aus Filmdialogen? Ein multimedialer Ansatz für das Sprachenpaar Deutsch-Italienisch". *Korpus Linguistik Deutsch – synchron, diachron, kontrastiv*. Tübingen: Niemeyer, 207-217.
- Herbst, T. 1987. "A pragmatic translation approach to dubbing". *EBU Reviews* 6: 21-23.
- Herbst, T. 1994. *Linguistische Aspekte der Synchronisation von Fernsehserien. Phonetik, Textlinguistik, Übersetzungstheorie*. Tübingen: Niemeyer.
- Heulwen, J. Roffe, I. Thorne, D. 1996. "Assessment and Skills in Screen Translation", in Dollerup, C. and V. Appel (eds.) *Teaching Translation and Interpreting 3: New Horizons. Papers from the Third Language*

- International Conference, Elsinore, 9-11 June 1995*. Amsterdam and Philadelphia: John Benjamins, 177-186.
- Nadiani, G. 1996. "Di alcuni segnali discorsivi nell'analisi contrastiva dei dialoghi in italiano e in tedesco del film *La stazione*", in Heiss, C. and R.M. Bollettieri Bosinelli (eds.) *Traduzione multimediale per il cinema, la televisione e la scena*. Bologna: CLUEB, 185-194.
- Orero, P. (ed). 2004. *Topics in Audiovisual Translation*. Amsterdam and Philadelphia: John Benjamins.
- Paolinelli, M. 2004. "Nodes and boundaries of Global communications: Notes on the translation and dubbing of audiovisual". *Meta* 49(1): 172-181.
- Pavesi, M. 1994. "Osservazioni sulla (socio)linguistica del doppiaggio", in Baccolini, R. et al. (eds.) *Il doppiaggio: Trasposizioni linguistiche e culturali*. Bologna: CLUEB, 51-60.
- Pavesi, M. 1996. "L'allocuzione nel doppiaggio dall'inglese in italiano", in Bosinelli, R.M. et al. (eds.) *La traduzione multimediale. Quale traduzione per quale testo?* Bologna: CLEUB, 117-130.
- Pavesi, M. 2004. "Dubbing English into Italian: a Closer Look at the Translation of Spoken Language". Paper presented at the International Conference *In So Many Words. Language Transfer on the Screen*. London. Manuscript.
- Rossi, F. 1999. *Le parole dello schermo. Analisi linguistica del parlato di sei film dal 1948 al 1957*. Roma: Bulzoni.
- Rundle, C. 2000. "Using subtitles to teach translation", in Bosinelli, R.M. et al. (eds) *La traduzione multimediale. Quale traduzione per quale testo?* Bologna: CLEUB, 167-181.
- Taylor, C. 1999. "The Translation of Film Dialogue", in Bassnett, S. et al. (eds.) *Textus*, 12/2, Genova: Tilgher.
- Taylor, C. 2003. "Multimodal transcription in the analysis, translation and subtitling of Italian films". *The Translator* 9(2): 191-205.
- Taylor, C. 2004. "Subtitling, filmese and MCA", in Taylor, C. et al. (eds.) *Proceedings of the Seminar, Linguatel: Progetto telematico di lingua e mediazione interlinguistica*. Padova: Unipress, 3-18.
- Teubert, W. 1996. "Comparable or parallel corpora?". *International Journal of Lexicography* 9(3): 238-264.
- Thibault, J.P. 2000. "The multimodal transcription of a television advertisement: theory and practice", in Baldry A. (ed.) *Multimodality and Multimediality in the Distance Learning Age*. Campobasso: Palladino editore, 311-85.
- Ulrych, M. 2001. "What corpora for what purpose: a translation-based view", in Taylor Torsello, C. et al. (eds.) *Corpora testuali per ricerca, traduzione e apprendimento linguistico. Studi Linguistici Applicati*. Padova: Unipress, 361-371.
- Valentini, C. Forthcoming. "Using multimedia corpora for advanced language learning". MuTra Conference *Challenges of Multidimensional Translation*, Saarbrücken, 2005.
- Zanettin, F. 2001. "Swimming in words", in Aston, G. (ed.) *Learning with Corpora*. Bologna: CLEUB, 177-198.

Zorzi, D. 2001. "The pedagogic use of spoken corpora: learners discourse markers in Italian", in Aston, G. (ed.) *Learning with Corpora*. Bologna: CLUEB, 85-108.

### **Filmography**

*Mimi metallurgico ferito nell'onore* [*The Seduction of Mimi*]. 1971. L. Wertmüller. Italy.

*La Stazione* [*The Station*]. 1990. S. Rubini. Italy.

*La vita è bella* [*Life is beautiful*]. 1998. R. Benigni. Italy.