

Continuity and Variation across Translations: Phraseology in the *Pavia* *Corpus of Film Dialogue*

MARIA FREDDI
(University of Pavia)

1. INTRODUCTION

The present paper describes the phraseology of original American and British filmic speech by relating it to issues of translation – dubbing – into Italian. In so doing, it brings together stances from audiovisual translation research, theory of phraseology, as developed by corpus linguists both with reference to monolingual analysis and in relation to the translation process, and corpus-based translation studies.

Recently, stress has been placed on sociocultural contextual factors affecting the translation process and products (see in particular Baker 2004; Laviosa 2004) and especially research from within corpus-based translation studies (hereafter CTS) has advocated that empirical studies of translators' variability should be carried out. This is in order to incorporate contextual parameters such as the actual agents of the translation process, professional and commitments constraints, etc. into the analysis of translated texts, something corpus methodology allows for. Strong criticism of a logocentric view of text corpora stems from translation scholars such as Mona Baker, who want to value translation as a variety in its own right, showing specificities and tendencies which are not present in other varieties of original, non-translated texts (cf. Baker 1995; Olohan 2004). Such specificities, often classified as 'translation universals', are related to contextual, non-linguistic factors. So, for example, the ob-

served tendency of translations towards explicitation practices is interpreted as a result of the translator's need to spell out things more explicitly for the target community.

The same invite to study the policy that regulates the whole translation process and a similar call for the incorporation of context into text corpora also comes from audiovisual translation (hereafter AVT), see in particular Diaz Cintas (2004: 25-29), and seems to share with CTS scholars many concerns, including the very notion of 'translation norms', although the two groups rarely speak to each other. The present paper uses the annotated corpus as a tool to study similarities and patterns of translation behaviour in film translation from this context-sensitive perspective.

2. AIMS AND METHODOLOGY

Within AVT, work on translated film language has shown how routinized translations are so common in film dubbing that it is possible to advance the hypothesis of a 'third norm', which is neither the source nor the target language norm (Pavesi 1994; Herbst 1996; Malinverno 1999). This has also been termed 'dubese', stressing both its autonomy and internal consistency (Freddi 1999). From a target-oriented and translational perspective, dubbese shows significant differences from non-translated language (Alfieri et al. 2003; Bucaria 2008; Bruti and Pavesi 2008), while from a source-oriented perspective, it is described as a series of semantic and structural calques which are found to occur repeatedly, hence the term 'translational routines' (Pavesi 2005: 48), such as, for example, the pairs *l'hai detto-you said it*, *la sai una cosa?-you know something?*, *scordatelo-forget about it*, etc. (for the full list of examples, see Pavesi 2005: 49).

A more recent angle on translated film language, particularly from English into Italian, but also from English into Spanish, has shown how distribution of certain expressions varies when dubbed film and television dialogues are compared with corpora of natural spoken discourse chosen as reference corpora (Romero Fresco 2006, 2009; Pavesi 2008; Bruti and Pavesi 2008), or with data found in reference grammars and dictionaries (Alfieri et al. 2003). This occurs at a time when translation scholars in general, not just re audiovisual translation, tend to focus on translated language as a variety in its own right to be compared with non-translated language, though not necessarily thought of as ancillary to it or related to the source language (Kenny 2005). This perspective is paralleled by the comparisons of original film and television dialogue with spontaneous conversation, again highlighting different distributional profiles (Quaglio 2008; Romero Fresco 2009) with divergence explained by Quaglio 2008 in terms of functional (re, for example, the role of vagueness and emotional language) and situational differences (limited range of conversation topics and settings).

This AVT perspective on the formulaic nature of film translation matches well with theory of phraseology, particularly with regard to Sinclair's 'idiom principle', or the 'single choices' the language user makes when producing and processing a text (Sinclair 1991: 110).

Within monolingual corpus linguistics, various scholars have shown patterns that come out as a result of repeated usage. In particular, lexical bundles, as defined by Biber and Conrad (1999: 183), are recurrent lexical sequences 'of three or more words that show a statistical tendency to co-occur', e.g., in conversation, *I don't know what, I don't know how, well I don't know, I don't want to, I don't think so*, etc. As is clear from the examples, lexical bundles do not require structural completion and become a reliable indicator of genre variation (Biber and Conrad 1999; Biber et al. 1999) as different registers or genres are characterised by different distributions of clusters. The extreme example of this tendency words have to cluster together is idiom in the traditional sense where frequency is accompanied by lexico-grammatical fixedness and semantic non-compositionality (Moon 1998: 7-8). A broad concept of idiomaticity as a continuum together with a pragmatic classification of idioms allow us to move from frequent co-occurrence to pure idioms and to study phraseology in film dialogues as strategies for fostering interaction as well as markers of textual organization (Moon 1998: 17).

Also of help is the notion of 'functional equivalence' coming from bilingual contrastive corpus analyses (Tognini Bonelli 2001, 2002), whereby two forms chosen as translational equivalents are such if their respective environments are considered and found to be equivalent as well, ensuing the recognition that the phraseological unit reunites lexical and grammatical patterns with a specific semantic dimension and a specific pragmatic function.

Therefore, this kind of analysis can be carried out by means of the *Pavia Corpus of Film Dialogue* (see Freddi and Pavesi in press), an annotated corpus that not only aligns the texts sampled, but also combines them with the contextual information that has been encoded in it and consequently establishes multi-level relations between data which are apparently distant, in an attempt to provide the 'situated explanation' Baker (2004: 183) calls for.

3. THE PHRASEOLOGY OF FILMIC SPEECH

The perspective of the present study is both monolingual – is there a phraseology of filmic speech? – and contrastive, that specific to AVT – is the repetitiveness and fixedness of certain translation solutions the direct consequence of the formulaicity of original film dialogue? The approach therefore exploits the parallel corpus in both directions, shunting between original texts and their translations examined in parallel and translated texts as such (the translational component of the parallel film corpus) in the framework of CTS.

In order to identify phrases that might be typical of film dialogue, both original and translated, the interrogation of the raw corpus has been combined with the annotated corpus search. More in detail, to approach the first perspective, the overall corpus frequency wordlist is looked at to hypothesise frequent words which might belong in frequent phrases. Then, concordances are generated through the database (the annotated corpus) query and analysed, allowing for further distinctions to be made between the American and British sub-corpora and for searches of single films as well as across films.

Once a phraseology of original film dialogue is identified, the corresponding translations are examined to see whether an equivalent set of recurring patterns are observable, what their specific pragmatic functions are and how the same functions are expressed in original and translated texts. Finally, these findings are compared with the instances of dubbese found in the literature (cf. Pavesi 2005; Taylor 2000, 2006, his ‘framing moves’ and predictability of filmic speech, i.e. ‘predictability of textual occurrences and frequencies as associated with particular scenes and scene types’), which are the result of a translational perspective. Analysis of data at this stage is aimed at checking whether the same kinds of patterns are present in the corpus. In the following paragraphs the findings are discussed in relation to the methodology followed.

3.1 FROM FREQUENT (LEXICAL) WORDS TO LEXICAL BUNDLES TO PRAGMATIC IDIOMS

With more than a thousand occurrences, *what* is a high frequency word in the overall corpus. Coming after the two top frequencies corresponding to the second and first person singular pronouns *you* and *I*, it ranks twelfth right after other usually very frequent grammatical words. It is evenly distributed across the British and the American film sub-corpora and found to occur in each film in the corpus.

Starting from frequency counts of individual words such as this, information on frequent collocates can also be retrieved through a software for automatic text processing (*Wordsmith Tools*). An interesting kind of output is the list of 3-word clusters (bundles) calculated within a span of 5 words to the left to 5 to the right from node word *what*, with minimum frequency set at 15, of which the top three 3-word clusters are as follows:

| Cluster | Freq. |
|---------------|-------|
| WHAT DO YOU | 93 |
| WHAT ARE YOU | 71 |
| YOU KNOW WHAT | 48 |

Table 1 – *Wordsmith Tools*’ list of 3-word clusters (both left and right horizons)

The first two both present subject verb inversion typical of the structure of questions, with *what* functioning as interrogative pronoun.

The next most frequent one, *you know what*, is interesting as it includes the fixed phrase *You know what?* functioning as turn launcher. This is also listed by Romero Fresco 2006 among the set phrases typifying the original dialogue of the television series *Friends* and matches Pavesi’s *You know something?* whose translation she records among the calques typical of dubbed Italian (Pavesi 2005: 49). In the whole corpus the string *You know what?* occurs 15 times, almost all of which distributed across 2 films, both American ones, namely *Erin Brockovich* and *Crash*. There is one single occurrence in one British English film,

Mike Leigh's *Secrets and Lies*, where the sequence combines with *Listen* in initial position, used as attention-getting device, and is translated as *Te lo dico io, sai cosa facciamo?*

The other occurrences in the corpus are variously translated as *Smettila* (perhaps a more apt equivalent for *stop it than you know what?*) / *Sai che ti dico* (3 times) / *Facciamo così / Sa(i) una cosa?* and *La sai una cosa?* (3 times). Among the renditions there are also 2 zero correspondences and hapaxes such as *Hey senti okay / Senta / Ecco, visto / Beh* and *lo sai?* pointing backwards, thus annulling the projecting cataphoric function of *you know what?*, by reason of which it has been labelled 'utterance launcher' (Biber et al. 1999).

It is interesting to notice that the variety of translation solutions includes both that which according to Pavesi 2005 sounds like natural Italian, namely *Sai che ti dico?*, and its more literal counterpart *La sai una cosa?*

A similar query was generated with min. frequency set at 5 and horizon 0 to the left to 5 to the right from node word *what*, thus considering only the clustering taking place in the right co-text:

| Cluster | Freq. |
|------------------------|-----------|
| WHAT DO YOU | 91 |
| WHAT ARE YOU | 70 |
| ARE YOU DOING | 30 |
| WHAT IS IT | 25 |
| DO YOU THINK | 24 |
| WHAT THE FUCK | 21 |
| WHAT DID YOU | 21 |
| DO YOU MEAN | 21 |
| WHAT THE HELL | 16 |
| DO YOU DO | 12 |
| WHAT KIND OF | 11 |
| DID YOU SAY | 10 |
| DO YOU WANT | 10 |
| WHAT WAS THAT | 9 |
| WHAT IT IS | 8 |
| WHAT WOULD YOU | 7 |
| WHAT I WAS | 7 |
| WHAT HAPPENED TO | 7 |
| WHAT AM I | 7 |
| ARE YOU TRYING | 6 |
| THE FUCK ARE | 6 |
| WHAT IT WAS | 6 |
| ARE YOU TALKING | 6 |
| WHAT YOU WANT | 6 |
| WHAT YOU THINK | 6 |

| | |
|------------------------|----------|
| ARE YOU GONNA | 6 |
| WHAT CAN I | 5 |
| WHAT I WOULD | 5 |
| WHAT THE HELL'S | 5 |
| WHAT DID HE | 5 |
| WHAT I SAID | 5 |
| WHAT YOU MEAN | 5 |
| WHAT IS THAT | 5 |
| WHAT DO I | 5 |
| WHAT IS THE | 5 |
| WHAT IF I | 5 |
| WHAT ARE THEY | 5 |
| WHAT ABOUT YOU | 5 |

Table 2 – Wordsmith Tools' list of 3-word clusters (right horizon only)

All the clusters above are characterised by structural incompleteness, that is, they do not fulfil a syntactic function, rather they are simply the result of statistical association and even distribution across the texts in the corpus.

However, a first look at the list shows that the overall number of occurrences of *what do you* (91) partially overlaps with occurrences of *do you think*, *do you mean*, *do you do* and *do you want* (highlighted in bold). Moreover, as a consequence of the transcription process, orthographic variants such as *d'you* / *do you* should also be taken into account when searching the corpus, so numbers might become even bigger.

In the same way, *what are you* (70) is to be seen in combination with *are you doing* (30), *are you trying* and *are you talking*. We thus get a series of *What*-questions *What are you doing?*, *What are you trying to say? / to do?* and *What are you talking about?* as frequent patterns of film dialogue.

On the basis of these results, some database queries were set up in order to be able to relate the occurrence of these phrases to the parameters encoded in the corpus (e.g. film, line number, or cue, character, language, translator-adaptor). Results for the phrase *what do you mean* are shown below in Table 3:

Table 3 – what do you mean

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|------------------|---------|----------|------------------|---|-------------------|
| Ae Fond Kiss | 661 | English | ROISIN | What do you mean? Why not? Did you talk to your folks? | |
| Ae Fond Kiss | 661 | Italian | ROISIN | <u>Cosa? Perché no ? Hai parlato con i tuoi?</u> | Depaolis Federica |
| Billy Elliot | 441 | English | BILLY | What do you mean? | |
| Billy Elliot | 441 | Italian | BILLY | <u>E perché ?</u> | Cosolo Carlo |
| Crash | 493 | English | CAMERON | What do you mean? | |
| Crash | 493 | Italian | CAMERON | <u>Che vuoi dire?</u> | Ottoni Filippo |
| Dead Man Walking | 58 | English | SISTER HELEN | What do you mean? | |
| Dead Man Walking | 58 | Italian | SISTER HELEN | <u>Che vuoi dire?</u> | Bertini Lorena |
| Dead Man Walking | 626 | English | SISTER HELEN | A lot of commotion for nothing. I... What do you mean you didn't know? | |
| Dead Man Walking | 626 | Italian | SISTER HELEN | <u>Vuoi dire che non lo sapevi?</u> | Bertini Lorena |
| Dead Man Walking | 798 | English | MATTHEW PONCELET | What do you mean? | |
| Dead Man Walking | 798 | Italian | MATTHEW PONCELET | <u>Che vuoi dire?</u> | Bertini Lorena |
| Erin Brockovich | 228 | English | GEORGE | What do you mean I cheat? | |
| Erin Brockovich | 228 | Italian | GEORGE | Eh eh... <u>Come sarebbe imbroglio?</u> | Mete Marco |
| Erin Brockovich | 396 | English | GEORGE | What? What do you mean you got fired? You're working so hard. | |
| Erin Brockovich | 396 | Italian | GEORGE | <u>Cosa? Come sarebbe licenziato? Non fai altro che lavorare.</u> | Mete Marco |
| Erin Brockovic | 736 | English | ERIN | What do you mean? | |
| Erin Brockovich | 736 | Italian | ERIN | <u>Che vuoi dire?</u> | Mete Marco |

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|-------------------|---------|----------|----------------|--|-------------------|
| Finding Forrester | 95 | English | TERREL | What do you mean “nothing”? This getting in the way of your plans or something? | |
| Finding Forrester | 95 | Italian | TERREL | Come sarebbe niente? È, è una cosa che rovina i tuoi piani per caso? | Caporello Elettra |
| Finding Forrester | 138 | English | JAMAL'S MOTHER | Don't know? What do you mean you don't know? | |
| Finding Forrester | 138 | Italian | JAMAL'S MOTHER | Non lo so? <u>Cosa vorrebbe dire</u> questo non lo so? | Caporello Elettra |
| Finding Forrester | 412 | English | JAMAL | What do you mean “that's a foul”? I had the spot. | |
| Finding Forrester | 412 | Italian | JAMAL | <u>Ma quale</u> fallo? Il posto era mio. | Caporello Elettra |
| Notting Hill | 91 | English | ANNA | Alright. Well, what do you mean “just over the street”? Give it to me in yards. | |
| Notting Hill | 91 | Italian | ANNA | D'accordo, bé, <u>cosa intende per</u> “altro lato della strada”? Quanti metri dista? | Vairano Francesco |
| Ocean's Eleven | 1080 | English | MR BENEDICT | What do you mean there's no money in the bags? | |
| Ocean's Eleven | 1080 | Italian | MR BENEDICT | <u>Come sarebbe</u> non c'erano soldi nelle borse? | Mete Marco |
| One Hour Photo | 147 | English | NINA YORKIN | What do you mean you feel bad? | |
| One Hour Photo | 147 | Italian | NINA YORKIN | <u>In che senso</u> sei preoccupato? | Valli Carlo |

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|----------------|---------|----------|-----------|--|------------------------|
| Secrets & Lies | 572 | English | MONICA | What do you mean ? | |
| Secrets & Lies | 572 | Italian | MONICA | <u>In che senso?</u> | Bucciarelli Elisabetta |
| Secrets & Lies | 647 | English | ROXANNE | What do you mean? | |
| Secrets & Lies | 647 | Italian | ROXANNE | <u>Che vuoi dire?</u> | Bucciarelli Elisabetta |
| Secrets & Lies | 1140 | English | CYNTHIA | What do you mean? | |
| Secrets & Lies | 1140 | Italian | CYNTHIA | <u>Beh che dici?</u> | Bucciarelli Elisabetta |
| Sliding Doors | 172 | English | HELEN 2 | Anna! I'm over him. What do you mean I'm not? How do you know I'm not? | |
| Sliding Doors | 172 | Italian | HELEN 2 | Anna, con lui ho chiuso. <u>Perché dici</u> il contrario? Come fai a saperlo? | Vairano Francesco |
| Sliding Doors | 365 | English | HELEN 2 | What do you mean you want to see me? Jesus Christ ,Gerry! What.. How did you get in? | |
| Sliding Doors | 365 | Italian | HELEN 2 | <u>Come sarebbe a dire</u> che volevi vedermi? Santo Dio! Gerry, cosa... come sei entrato? | Vairano Francesco |
| Sliding Doors | 418 | English | LYDIA | What do you mean "what am I trying to do"? I think I've broken my toe. | |
| Sliding Doors | 418 | Italian | LYDIA | <u>Come sarebbe</u> "qual è il mio scopo"? Si sarà fratturato l'alluce. | Vairano Francesco |

The following observations can be made: the string *what do you mean* occurs in 11 out of the 12 films in the corpus, unevenly distributed across the films but always filling the same function of taking up the interlocutor's words to rebuke them. As far as the Italian version is concerned, there seems to be no preferred rendering, translations oscillating between *che vuoi dire?* and *come sarebbe?* also depending on the translator-adaptor: the same translator, Marco Mete, chooses both *come sarebbe* and *che vuoi dire* within the same film; *come sarebbe* is chosen by another translator, Francesco Vairano in another film. More peripheral instances such as *in che senso?* together with *cosa intende per...* make it difficult to see any regularity in the translation, thus we can only tentatively hypothesise a tendency towards functional equivalence between *what do you mean?* and *che vuoi dire? / come sarebbe?*

What is preferably translated as *che* than *cosa* (cf. also the fixed equivalence of the first most frequent phrase *What do you think? - Che ne pensi?*). Rather than a feature of dubbese, *cosa* being the unmarked expression in natural spoken Italian in the north, *che* is unmarked in Southern Italy (see also Pavesi 2005: 39 on colloquial and regional forms). Indeed, most translator-adaptors are based in Rome. In another query, analysis of *What are you doing?* gave *Che stai facendo?* far outnumbering other less frequent solutions like *Cosa fai? / Che fai / etc.* without the continuous aspect. Interestingly, when the expression is followed by the place adverbial as in *What are you doing here?* (6 occ.), the continuous aspect disappears and becomes *Che ci fai qui?* (3) / *Che ci fai tu qui?* / *Come mai sei qui?* / *Che cosa ci fai qui?*

Some of the occurrences contain a swearword stressing the action-quality of the situation, *What the fuck are you doing?*, or another mark of the colloquial register as in *What are you doin' man?* with the contracted form and the Vocative *man*, which in the Italian dubbing are rendered by the patterned formulations *Che credi di fare?*, the allocution having totally disappeared, or *Dove accidenti vai?* where the loss of the Vocative *man* is compensated for by the expletive *accidenti*. Sometimes, the tension accrued by the use of the Vocative in the source-text, *Tony, what are you doing?*, is made explicit in the target-text by means of the weak connective together with the verb of volition, as in *Tony, ma che vuoi fare?*

Analogously, the expression *What are you talking about?* (12 occ. of which 4 are transcribed as *What're you talking about?*) regularly occurs in confrontational scenes in each film, as is emphasised by the one occurrence, from *Erin Brockovich*, *What the fuck are you talking about?* with the swearword in between the interrogative pronoun and the finite element. Its function is to carry on the plot by forcing the interlocutor to produce the missing bit of information. The following translation solutions were found for this set phrase: *Ma che cosa stai dicendo? / (Adesso) di che cosa stai parlando? / Di che stai parlando? / Ma di che cazzo stai parlando? / Che cosa stai dicendo? / Ø*, while the 4 instances of the contracted form are translated as *Che stai dicendo? / Ma cosa stai dicendo?* Again, the continuous aspect together with the weak connective *ma* seem to be recurrent features (cf. Pavesi 2005: 49).

A similar fixed discursal function can be seen in the occurrences of *What are you trying to do?* (4) and *What are you trying to say?* (2). They are both stock expressions in that they help carry the action forward, with the interlocutor usually being prompted an explanation or summary of some events that will be seen later on in another scene, or else just talked about in the next turn.

As far as their repetitiveness in the translation goes, the literal **che stai cercando di fare?* is never used. Rather, two quite unusual renderings *mi dici qual è il tuo scopo?* and *E allora, qual è il tuo scopo, eh?* come from the same film and turn. Their fixedness being internal to a single speaker's turn, they can hardly be significant in quantitative terms and only indicative of a very specific stylistic choice (of translator Francesco Vairano). The remaining two occurrences are instances of a specific construct, whereby a bare infinitive lexicalizes the action verb *do*: *What are you trying to do, just rub it in my face?*-*Così pensi di sbattermi in faccia la realtà?*; *What are you trying to do, make him a fucking scab for the rest of his life?*-*Stai cercando di farne un crumiro del cazzo per tutta la vita?*

On the other hand, there is a perfect correspondence between the repetition of *What are you trying to say?* in the source-text and its translation as *che stai cercando di dire?* in the target-text. The Italian corpus also has an instance of *è questo che stai cercando di dirmi?* resulting from the pseudo-cleft structure *Is that what you're trying to say?* More data is needed, however, to prove the hypothesis of the expression being fixed and a calque of the original.

If we go back to the 3-word clusters list above, the fourth most frequent one is *what is it* for which another corpus query was made. The results are best displayed in Table 4 where the English version has been kept separate from the Italian to help visualize any regularities in the translation.

Table 4 – *what is it*

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|------------------|---------|----------|---------------|---|---------|
| Ae Fond Kiss | 135 | English | MR KHAN | What is it? | |
| Ae Fond Kiss | 634 | English | CASIM | What is it? | |
| Ae Fond Kiss | 676 | English | MR KHAN | What is it? | |
| Ae Fond Kiss | 808 | English | ROISIN | Jamin, thanks very much. What is it? | |
| Ae Fond Kiss | 986 | English | ROISIN | What is it? | |
| Billy Elliot | 304 | English | MRS WILKINSON | What is it? | |
| Dead Man Walking | 46 | English | FARLEY | Do you know what you're getting into? So what is it , Sister? Morbid fascination? Bleeding heart sympathy? | |
| Dead Man Walking | 600 | English | UNIDENTIFIED | What is it? What happened? | |
| Dead Man Walking | 850 | English | SISTER HELEN | What, Mat? What is it? | |

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|-----------------|---------|----------|------------|---|---------|
| Erin Brockovich | 2 | English | ERIN | No. I have kids. Learned a lot right there. Yeah, I mean, I've seen nurses give my son throat cultures, and, what is it? | |
| Erin Brockovich | 831 | English | ERIN | There's no baby in here. What is it? | |
| Notting Hill | 741 | English | ANNA | What is it about men and nudity, huh? | |
| Notting Hill | 757 | English | ANNA | What? What is it? | |
| Notting Hill | 807 | English | WILLIAM | Blimey, what is it? | |
| Ocean's Eleven | 358 | English | LINUS | What is it? | |
| One Hour Photo | 495 | English | BILL OWENS | What is it , Yoshi? You can't just leave your station unmanned down there. | |
| Secrets & Lies | 5 | English | MONICA | What is it? | |
| Secrets & Lies | 314 | English | HORTENSE | What is it? | |
| Secrets & Lies | 689 | English | CYNTHIA | Yes, what is it you want, darling? Hello? Did you want Roxanne? She's gone out. | |
| Secrets & Lies | 704 | English | CYNTHIA | Listen, darling, what is it you want? | |
| Secrets & Lies | 770 | English | CYNTHIA | What is it , sweetheart? | |
| Secrets & Lies | 788 | English | CYNTHIA | Why? What is it? | |
| Sliding Doors | 80 | English | ANNA | What is it you've done, anyway? It can't be that nice if she's walked out on you. | |

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|------------------|---------|----------|---------------|--|-------------------|
| Sliding Doors | 547 | English | ANNA | She saw you, James, at the hospital. With your wife. You know, the one with the wedding ring. What is it with you bloody men? | |
| Ae Fond Kiss | 135 | Italian | MR KHAN | <u>Che c'è?</u> | Depaolis Federica |
| Ae Fond Kiss | 634 | Italian | CASIM | <u>Ma cosa?</u> | Depaolis Federica |
| Ae Fond Kiss | 676 | Italian | MR KHAN | <u>Che c'è?</u> | Depaolis Federica |
| Ae Fond Kiss | 808 | Italian | ROISIN | Grazie, <u>ma che cos'è?</u> | Depaolis Federica |
| Ae Fond Kiss | 986 | Italian | ROISIN | <u>Di che si tratta?</u> | Depaolis Federica |
| Billy Elliot | 304 | Italian | MRS WILKINSON | <u>Che cos'è?</u> | Cosolo Carlo |
| Dead Man Walking | 46 | Italian | FARLEY | Lo sa a cosa sta andando incontro? <u>Che cosa la spinge sorella?</u> Un'attrazione morbosa? Pietà per le anime dannate? | Bertini Lorena |
| Dead Man Walking | 600 | Italian | UNIDENTIFIED | <u>Che c'è, che è successo?</u> | Bertini Lorena |
| Dead Man Walking | 850 | Italian | SISTER HELEN | Vuoi parlare. <u>Che cosa c'è?</u> | Bertini Lorena |
| Erin Brockovich | 2 | Italian | ERIN | No. Ma ho dei figli. E ho imparato parecchio. Sì, cioè, ho visto le infermiere fare i tamponi per la gola a mio figlio e, <u>che ci vuole...</u> | Mete Marco |
| Erin Brockovich | 831 | Italian | ERIN | Non c'è un bambino qui dentro. <u>Che cos'è?</u> | Mete Marco |
| Notting Hill | 741 | Italian | ANNA | Perché gli uomini sono fissati con il nudo, huh? | Vairano Francesco |
| Notting Hill | 757 | Italian | ANNA | Chi è? <u>Cosa c'è?</u> | Vairano Francesco |
| Notting Hill | 807 | Italian | WILLIAM | Perbacco! <u>Cos'è?</u> | Vairano Francesco |

| FILM | ID_CUES | LANGUAGE | CHARACTER | TEXT | ADAPTER |
|----------------|---------|----------|------------|--|------------------------|
| Ocean's Eleven | 358 | Italian | LINUS | <u>Che cos'è?</u> | Mete Marco |
| One Hour Photo | 495 | Italian | BILL OWENS | Che c'è, Yoshi? Lo sai che non puoi lasciare il laboratorio. | Valli Carlo |
| Secrets & Lies | 5 | Italian | MONICA | <u>Che c'è?</u> | Bucciarelli Elisabetta |
| Secrets & Lies | 314 | Italian | HORTENSE | <u>Cos'è?</u> | Bucciarelli Elisabetta |
| Secrets & Lies | 689 | Italian | CYNTHIA | Sì, <u>che c'è gioia?</u> .. Pronto?..Volevi Roxanne? È uscita. | Bucciarelli Elisabetta |
| Secrets & Lies | 704 | Italian | CYNTHIA | <u>Che cosa vuoi?</u> | Bucciarelli Elisabetta |
| Secrets & Lies | 770 | Italian | CYNTHIA | Sì, <u>che c'è?</u> Dica. | Bucciarelli Elisabetta |
| Secrets & Lies | 788 | Italian | CYNTHIA | Perché, <u>che c'è?</u> | Bucciarelli Elisabetta |
| Sliding Doors | 80 | Italian | ANNA | <u>Che cosa ha fatto</u> si può sapere? Niente di carino se ti ha piantato. | Vairano Francesco |
| Sliding Doors | 547 | Italian | ANNA | Ti ha visto, James, all'ospedale con tua moglie. Sai, quella con la fede al dito. <u>Ma che vi prende</u> a voi uomini. Non siamo qui per... | Vairano Francesco |

This cluster contains the question *What is it?* meaning 'is there any problem?' or simply replacing the whole-turn *What?* translated as *Che c'è?*. Sometimes it is followed by the preposition *with* as in *What is it with you?* with a function akin to *What (i)'s the matter with you?* and *What (i)'s wrong with you?*

These expressions, in their turn, are present in the corpus and translated as *Ma che cosa hai?* / *Cosa ti succede?* / *Che succede?* / *Ma che ti/vi prende?*, occasionally with an intervening swearword as in *What the hell's wrong with you?* translated as

Ma che diavolo ti prende?. Notice that *what the hell* also appears among the next most frequent 3-word clusters with 16 occurrences in the overall corpus (although not in each film, and yet in all American ones), see Table 2 above.

The Italian string *che ti prende* can also be the result of the English *What is / What's going on?* (and the cursing *What the fuck is going on? / What the hell's going on?*). These are more commonly rendered as *Che (ti) succede* than *che sta succedendo* (apart from one *Che diavolo sta succedendo?*), with the plain present tense in lieu of the progressive seen before. However, the frequent use of the weak connector *ma* in initial position as in *What's going on with you? - Ma che ti prende adesso?* – a typical feature of dubbed Italian (Malinverno 1999; Pavesi 2005, 2008) – adds to the register-specificity of a phrase characterised by the carry-the-action-forward function which has already been observed with reference to many of these set phrases.

Finally, it is interesting to notice that Table 4 also exemplifies instances of the split structure *what is it you've done?* and *what is it you want?*. The translation, however, neutralizes the emphasis of the English construct by choosing the unmarked word order *che cosa hai fatto? / che c'è? / che cosa vuoi?* in a trend contrary to what has been observed for marked word orders in dubbed texts (Pavesi 2005, 2008).

It is clear how analysis of the clusters in the list and their collocates can proceed along the same lines to get the complete mapping of the phraseology of filmic speech. What emerges is that the idiomaticity of these expressions does not entail semantic opacity, rather it is pragmatic and register-specific to the extent that it works at the communicative and narrative levels.

3.2 TOP-DOWN INVESTIGATION OF PHRASES IN FILM DIALOGUE

Another way of looking at phraseology in film dialogue is by resorting to the inventory of fixed or semi-fixed phrases that are known in the literature to be used in natural oral conversation to see whether they are present in the corpus and to what extent. This approach is closer to a top-down rather than the bottom-up approach adopted so far.

Thus, the following expressions from Pavesi (2005: 49) were searched for: *Forget it!* (4 occ.) translated as *scordatelo* (once), *non ci raccontiamo scemenze / non importa / neanche per sogno* together with *(you can) forget about... ti puoi scordare / ti scordi pure / lascia perdere*. Again, it is interesting that the double clitic construction, said to be calqued from English, only once is the result of the expression *Forget it! Scordatelo*, however, counts another 2 occurrences in the corpus which are resorted to to translate the two elliptical answers *We're not* and *I can't*.

Conversely, search for *lascia perdere* gave 9 occurrences which, if cross-checked with the instances of *forget about*, only once translates it, while it is rather the result of various formulas in the original text, particularly *leave it (alone)* (4), *doesn't matter, it's fine, I don't care what* and is once inserted anew by the translator in lieu of a turn uttered in Punjabi in the original version. In the same vein, the Italian string *lascia stare* was found to occur 8 times translating very different realizations including *It's okay, My sister's okay! - Lascia stare mia sorella!*, *Leave her alone - Lascia stare tua sorella*, *Never mind tea - Lascia stare il tè adesso* and *don't touch your nose - lascia stare il naso*.

Also from Pavesi 2005, node word *kidding* was searched for to get all patterns of *Are you kidding (me)?* (4), *You're kidding?!* (4), *You're fucking kidding me?* (1) and *You've got to be kidding me* (1). The 10 instances of this expression come from just 7 out of the 12 films in the corpus, translated as *sta(i) scherzando?* (5 times), twice as *scherzi?* and once as *vuoi scherzare?*. The remaining two instances *You're fucking kidding me?* and *Are you kidding me?* both come from the same quarrel scene and are rendered as *Mi stai prendendo per il culo?* and *Mi stai prendendo in giro?*, uttered by the same character addressing the ambiguous main protagonist of *One Hour Photo* in both cases.

Other expressions known to characterise spontaneous conversation (Taylor 2000; Quaglio and Biber 2006: 704) were found, among which the utterance launchers and attention getting devices *I'm telling you*, *Tell you what*, *Guess what*, and *Thing is*.

The expression *I'm telling you* is not very frequent (4 occ. in overall corpus) and gets even more diluted in the translation, cf. *E adesso ti dico che / Ma te lo ripeto / Dammi retta / Vedrai* (notice, however, the reiterated turn-initial use of the pragmatic connectives *e* and *ma*).

The cognate expression *Tell you what* together with its variants *I'll tell you what* and *I tell you what*, also not very frequent (5 occ.), is translated as *Sai che facciamo? / Facciamo così* (2) / *Sai che ti dico?* and *Va bene*. A similar function is shared by *Guess what?* (3 occ.) invariably translated as *indovina un po'?* Similarly, *Thing is* (1) plus *The thing is* (5) is almost invariably translated as *E' che / La cosa è ... che / Il fatto è che*.

Unquestionably contributing to the simulation of interactivity (cf. Quaglio and Biber 2006: 716), the sparse occurrence of these phrases throughout the corpus triggers the usage of fixed forms in the target-texts, particularly pragmatic connectives and split structures, which have been shown to be a translation strategy typical of dubbed texts (Pavesi 2005, 2008; Pavesi and Perego 2008: 127). Once more is the idiomaticity of these expressions linked to their pragmatic function in the original product.

4. CONCLUSIONS

The analysis thus carried out allows for some conclusions to be drawn. First, a frequency-based description of contemporary American and British film dialogue brings to the fore holistic chunks of speech with specific functions. The functions are primarily pragmatic in that the phrases thus identified help achieve a move in dialogue. This move can be glossed as speaker A challenging speaker B usually by means of a question whose tenor is confrontational. This is why routines can be said to also have a fundamental diegetic function, in that they help advance the situations and events occurring within the fictional world of films. In fact, unlike natural conversation, film dialogues are not only linked to the situational context and the moment of uttering, they also link up to the narrative dimension of the diegesis, so that formulas like *What do you mean?*, *What're you trying to say?*, etc. have a function similar to that of a narrator whispering to the audience "hey audience, watch out as something is going to happen and character X is going

to tell us about it". The phraseology identified is therefore part of the narrative space of film dialogue.

Because of this combination of pragmatic – internal to the dialogue as anchored to the communicative situation – and diegetic – internal to the fictional world narrated by film – functions, phrases affect the kinds of interactions taking place on screen and acquire conventional significance for the type of text examined, i.e. film dialogue, thus contributing to the making of register-specificity. The second question of what the implication of this clustering in the original texts is for film translation, dubbing in particular, has been answered, showing, however, that in some instances translational routines leave place for variation and creativity on the part of the translators recreating a similar discourse function and register characterisation of the original. When translating-adapting, some of these formulas retain their formulaicity and communicative function, e.g. through repeated usage of weak connectives, progressive tense and split structures, others are diluted through variation across translations. We could then ask if such fixedness is a feature of 'filmese', as in Taylor 2006, thus reflected in dubbese, or else it is mirrored in original Italian film dialogue. To this purpose, a comparable corpus of original Italian films sampled according to the same criteria should be compiled to allow for quantitative comparisons to be made.

From a translational viewpoint, the paper has tried to show how phraseology, by reuniting the lexical and the syntactic levels, is the preferred carrier of the register-specificity of the original product and therefore also the place where translators' variability can be best observed. This is why the annotated parallel corpus and the comparative approach allow to better investigate the relation between original and translated film dialogue thought of as a particular text-type, thus showing a possible outcome of the fruitful cross-fertilization of translation studies and corpus linguistics.

- Alfieri, G., S. Contarino, D. Motta 2003, "Interferenze fraseologiche nel doppiaggio televisivo: l'italiano di E.R. e di Beautiful" in A. Sullam Calimani (a cura di) *Italiano e inglese a confronto. Atti del convegno "Italiano e inglese a confronto: problemi di interferenza linguistica"*, Venezia, 12-13 aprile 2002, 127-149
- Baker, M. 1995, "Corpora in translation studies: An overview and some suggestions for future research", *Target*, 7 (2), 223-243.
- Baker, M. 2004 "A corpus-based view of similarity and difference in translation" *International Journal of Corpus Linguistics*, 9:2, 167-193.
- Biber, D. and S. Conrad 1999, "Lexical bundles in conversation and academic prose", in H. Hasselgard and S. Oksefjell (eds.), *Out of Corpora. Studies in Honour of Stig Johansson*. Amsterdam: Rodopi, 181-190.
- Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan 1999, *Longman Grammar of Spoken and Written English*. London: Longman.
- Bruti, S. and M. Pavesi 2008, "Interjections in translated Italian: Looking for traces of dubbed language", in A. Martelli and V. Pulcini (eds.), *Investigating English with Corpora. Studies in Honour of Maria Teresa Prat*. Monza: Polimetrica, 207-221.
- Bucaria, C. 2008, "Acceptance of the norm or suspension of disbelief? The case of formulaic language in dubbese", in D. Chiaro, C. Heiss and C. Bucaria (eds.), *Between Text and Image. Updating Research in Screen Translation*. Amsterdam/Philadelphia: John Benjamins, 149-163.
- Diaz Cintas, J. 2004, "In search of a theoretical framework for the study of audiovisual translation" in Orero (ed.), *Topics in Audiovisual Translation*, Amsterdam/Philadelphia: John Benjamins, 21-34.
- Freddi, M. 1999 "Dialogue analysis and multimedia translation" in B. Naumann (ed.), *Dialogue Analysis and the Mass Media. Beiträge zur Dialogforschung*, 20 *Proceedings of the I.A.D.A International Conference*. Tübingen: Niemeyer, 149-158.
- Freddi, M. and M. Pavesi (eds.) in press, *Analysing Audiovisual Dialogue. Linguistic and Translational Insights*. Bologna: Clueb.
- Herbst, T. 1996 "Why dubbing is impossible", in C. Heiss, R.M. Bollettieri Bosinelli (a cura di), *Traduzione multimediale per il cinema, la televisione, la scena*. Bologna: Clueb, 97-115.
- Kenny, D. 2005, "Parallel corpora and translation studies: old questions, new perspectives? Reporting that in GEPOLC: a case study" in G. Barnbrook, P. Danielsson, M. Mahlberg (eds.), *Meaningful Texts: The extraction of semantic information from monolingual and multilingual corpora*. London: Continuum, 155-165.
- Laviosa, S. 2004, "Corpus-based translation studies: Where does it come from? Where is it going?", *Language Matters*, 35 (1), 6-23.
- Malinverno, A. 1999, "La resa delle varietà non-standard dal film americano all'italiano del doppiaggio", in M. Vedovelli (a cura di), *Indagini sociolinguistiche nella scuola e nella società italiana in evoluzione*, Milano: Franco Angeli, 51-72.
- Moon, R. 1998, *Fixed Expressions and Idioms in English. A corpus-based approach*. Oxford: Clarendon.
- Olohan, M. 2004, *Introducing Corpora in Translation Studies*. London: Routledge.
- Pavesi, M. 1994, "Osservazioni sulla (socio)linguistica del doppiaggio", in R. Baccolini, R.M. Bollettieri Bosinelli, L. Gavioli (a cura di), *Il doppiaggio. Trasposizioni linguistiche e culturali*. Bologna: Clueb, 129-142.
- Pavesi, M. 2005, *La traduzione filmica. Aspetti del parlato doppiato dall'inglese all'italiano*. Roma: Carocci.
- Pavesi, M. 2008, "Spoken language in film dubbing: Target language norms, interference and transla-

- tional routines”, in D. Chiaro, C. Heiss and C. Bucaria (eds.), *Between Text and Image. Updating Research in Screen Translation*. Amsterdam/Philadelphia: John Benjamins, 79-99.
- Pavesi, M. and E. Perego 2008, “Il dialoghista in Italia: indagine sociologica e norme linguistiche”, in A. Baldry, M. Pavesi, C. Taylor Torsello and C. Taylor (eds.), *From Didactas to Ecolingua. An Ongoing Research Project on Translation and Corpus Linguistics*, Trieste: EUT, 111-139.
- Quaglio, P. and D. Biber 2006, “The grammar of conversation” in B. Aarts and A. McMahan (eds.), *The Handbook of English Linguistics*. Oxford: Blackwells, 692-723.
- Quaglio, P. 2008, “Television dialogue and natural conversation: Linguistic similarities and functional differences”, in A. Ädel and R. Reppen, (eds.), *Corpora and Discourse: The challenges of different settings*, Amsterdam/Philadelphia: John Benjamins, 198-210.
- Romero Fresco, P. 2006, “The Spanish dubbese: a case of (un)idiomatic Friends”, *Journal of Specialized Translation* 6.
- Romero Fresco, P. 2009, “Naturalness in the Spanish dubbing language: A case of not-so-close Friends” *Meta*.
- Sinclair, J. 1991, *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Taylor, C. 2000, “Look Who’s Talking. An analysis of film dialogue as a variety of spoken discourse”, in L. Lombardo, L. Haarenan and J. Morley (eds.), *Massed Medias: Linguistic tools for interpreting media discourse*, Milano: LED, 247-278.
- Taylor, C. 2006, “‘I knew he’d say that!’ A consideration of the predictability of language use in film”, in *MuTra 2006 – Audiovisual Translation Scenarios*, EU-High-Level Scientific Conference Series, 1-11.
- Tognini Bonelli, E. 2001, *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.
- Tognini Bonelli, E. 2002, “Functionally complete units of meaning across English and Italian: Towards a corpus-driven approach”, in B. Altenberg and S. Granger (eds.), *Lexis in Contrast. Corpus-based approaches*. Amsterdam/Philadelphia: John Benjamins, 73-95.