CONVERSATION AS A UNIT OF FILM ANALYSIS.
DATABASES OF L3 TRANSLATION AND AUDIOVISUAL SAMPLES OF MULTILINGUALISM

Patrick Zabalbeascoa
patrick.zabalbeascoa@upf.edu
Universitat Pompeu Fabra

Montse Corrius
montse.corrius@uvic.cat
Universitat de Vic-Universitat Central de Catalunya

Abstract

This chapter argues in favour of analysing audiovisual texts (cinema and TV) by using the concept and unit of “conversation”, theoretically (Bakhtin 1982; Grice 1975; Sidnell 2010) and practically, for the purpose of creating an audiovisual corpus or database. The point is illustrated and defended from the experience of research carried out within the Trafilm project (trafilm.net), whose aim is to create a database of instances of multilinguism in films and samples of how they are rendered in dubbing, subtitling (including SDH) and audio description, mostly (but not exclusively, because it is open to researchers with other interests) from English into Spanish. Conversations are seen as the smallest possible units that contain instances of L3 (any language which is not the main language) with sufficient contextual and pragmatic information, and possibly useful for analysing other audiovisual (AV) features (e.g., representation of identities), too, and not only L3.

1. This study was supported by the grant FFI2014-55952-P from the Ministerio de Economía y Competitividad for the TRAFILM project, on the translation of multilingual films in Spain.
Resumen

Este capítulo presenta argumentos teóricos (Bakhtin 1982; Grice 1975; Sidnell 2010) y prácticos para defender el concepto de “conversación” como unidad análisis de textos audiovisuales (cine y televisión), y su utilidad para un corpus audiovisual o una base de datos. Este planteamiento es ilustrado y defendido a partir de la experiencia de investigaciones realizadas dentro del proyecto trafilm.net, cuyo objetivo es crear una base de datos de casos de multilingüismo en películas y muestras de cómo se traducen en doblajes, subtítulos (incluida la accesibilidad) y audiodescriptiones, mayoritariamente del inglés al español (pero no exclusivamente, dado que está abierto a otros investigadores). La conversación se concibe como la unidad más pequeña que contiene muestra(s) de L3 (cualquier lengua aparte de la lengua principal) para la que proporciona suficiente información contextual y pragmática. Se espera que sirva también para analizar otros elementos audiovisuales (AV) además de L3 (p.ej. la representación de identidades).

Keywords: Conversation. Multilingualism. Audiovisual translation. L3. Unit of analysis.

1. The translation of multilingual audiovisual films: an overview

There seems to be a growing trend in recent years for producing audiovisual texts (cinema or TV) that are not entirely monolingual, for various reasons. According to Heiss (2004), the number of films requiring the audience to deal with communication in more than one language has increased since the 1980s and we might add that this tendency has grown even more since Heiss made this claim. A wide variety of English-language films display two or more languages to a greater or lesser degree. They constitute instances of the kind of language variation we call “third language” (Corrius 2008), or L3 (Corrius & Zabalbeascoa 2011). For example, *A Passage to India* (David Lean 1984) includes native Urdu, non-native English, and non-native Urdu. *Beauty and the Beast* (Gary Trousdale & Kirk Wise 1991) adds instances of French; *Lost in Translation* (Sofia Coppola 2003), as the title suggests, depicts scenes of translating and interpreting (intratexual translation, Corrius & Zabalbeascoa 2011) between English and Japanese, as language and communication become a theme within the film. French and German can also be heard in *Lost in Translation* but only once. Certain directors, even with different agendas, use language variation quite frequently in their films, like Quentin Tarantino (e.g., *Inglorious Basterds* 2009, *Django Unchained* 2012), who reinforces stereotypes, or Ken Loach (e.g., *Bread and Roses* 2000, *Land and Freedom* 2004), who does not. In science fiction, Peter Jackson’s *The Lord of the Rings* trilogy (2001, 2002, 2003) provides good samples of constructed (as opposed to real) third languages. Apart from the important case of films produced in English, there are productions from countries in virtually every continent, whose main language (L1) is not English. For example, Almodóvar’s L1-Spanish *Todo sobre mi madre* (1999), with a few words in L3-Catalan (L3), or de Oliveira’s L1-Portuguese *Um filme falado* (2000) with French, Italian, English and Greek characters all speaking their own L3 language.

Needless to say in recent years the topic has raised interest among scholars devoted to audiovisual translation (AVT) and a number of researchers to a
greater or lesser degree, have done research in this field, rising to the challenge. In the context of Heiss’s (2004) claims that in the case of multilingual films both subtitling and dubbing need to be taken into account, it can be said that Corrius & Zabalbeascoa’s (2011) research is aimed at all AVT forms but draws more samples from dubbing, whereas more recent studies, such as Krämer & Eppler (forthcoming 2018) and Iberg (forthcoming 2018) are proposing specific theoretical models for L3 renderings in subtitles, based on case studies of the TV series *Breaking Bad* (AMC 2008-2013) and *Game of Thrones* (HBO 2011-2018), respectively. Delabastita and Grutman (2005: 16) described the multilingual text as a text that “is worded in different languages” and in the same year Dwyer (2005: 296) described it as a text “marked by the naturalistic presence of two or more languages at the level of dialogue and narrative”. Wahl (2005: 43) described polyglot films as “anti-illusionist” because they show the diversity of human life and do not hide it behind the mask of a universal language, and proposed a classification of six subgenres for polyglot films: episode, alliance, globalisation, immigrant, colonial, and existential. Bleichenbacher (2008) studied the presence and purpose of languages other than English in Hollywood movies but not really their foreign versions, whereas O’Sullivan (2011) focused on the role languages play in the construction of film narratives. Corrius & Zabalbeascoa (2011) worked on providing a more systematic understanding of the nature of L3, as it appears in audiovisual texts, formally rather than functionally, identifying the variables that can affect and characterise L3, such as its comprehensibility or familiarity. More recently, de Higes-Andino (2014) and the TRAMA group based in Castelló studied linguistic diversity in Spanish immigration films (Martínez-Sierra et al. 2010; de Higes-Andino et al. 2013). Meylaerts & Serban (2014) remarked the intradiegetic functions of third languages, such as character portrayal, voice and point of view and its extradiegetic role (used for the sake of authenticity or exoticism). Zhang (2015) carried out a study investigating the role of subtitling in multilingual films in China. Trafilm was started in 2015 as a three-year project funded by the Spanish Government, aimed at advancing in the study of the translation of multilingual audiovisual texts, by validating and refining existing theoretical models for AVT and multilingualism. The project set out to describe and analyse a rich collection of data which involved creating a consultable database of audiovisual (AV) instances of L3 along with their dubbed, subtitled (SDH produced specifically for the deaf and hard of hearing and otherwise) or audio described versions, as a relevant research tool for AVT studies, to ultimately reach out to teachers and professional practitioners.
2. In search of an AV unit that can work in practical and theoretical terms

In order to carry out the trafilm.net project, it was necessary to find or create a suitable software tool for storing and sorting a large number of multilingual samples taken from 21st-century films so that they could be descriptively analysed. The aim is to discover how the phenomenon of multilingualism is dealt with in AVT, its features and the tendencies that arise from rendering linguistic variation for dubbing, subtitling (including SDH) and audio description. Driven by the ambition to collect samples and metadata, it seemed only natural to establish the boundaries for each sample and the location for each piece of information. The challenge was to find a unit of analysis that could work theoretically and respond adequately to the practical needs of building a database. Because the project is about L3 within AVT, the obvious candidates for the database were primarily the clip and the utterance, but both of these soon presented problems that called for rethinking.

We propose the term “conversation” as a unit of analysis capable of providing context as well as the exact location of each instance of L3 within a given AV text; audiovisual context (linguistic, diegetic and visual information) helps to draw the meaning of what is shown on the screen. In this sense it is worth highlighting that, in the analysis of multilingual AV texts, conversation is precisely what helps viewers understand the presence and the purpose of each L3-instance. Conversation, which connects one thing to so many other things, is borrowed from the field of Conversation Analysis, “an approach within the social sciences that aims to describe, analyse and understand talk as a basic and constitutive feature of human social life” (Sidnell 2010: 32) but it also happens to fit in well with certain practical recommendations of guides for scriptwriters (e.g., Bordwell 2006).

As Sidnell (2010: 22) himself has stated, Conversation Analysis, which studies conversation as an institution, is based on a set of methods that are useful for working with audio and video recordings of talk and social interaction since talk is at the heart of human social life. As Goffman puts it:

Conversation has a life of its own and makes demands on its own behalf. It is a little social system with its own boundary-maintaining tendencies; it is a little patch of commitment and loyalty with its own villains. (1957:47)

According to The Complete Film Dictionary (Konigsberg 1993: 53), a clip is a “brief selection from a film used as an illustration of the work to which it belongs or as an illustration of some point (usually removed from a movie or film footage). The Trafilm database primarily consists of a comprehensive amount of metadata concerning dozens of L3 features and factors but it can also
hold AV clip samples (AV source text, or ST, AVT sample, or both). Such clips require certain rules or common features so they can belong to the database according to well-defined criteria. Thus, clip length and contents were considered. Should a clip be as long as (and no longer than) a scene, for example? Should it contain only one L3 utterance and what are the implications of using one set of criteria or another? In the final instance, we decided that a clip should coincide with the start and end time of a Conversation, allowing for a bit more at either end if justified and provided it does not exceed a duration of 3 minutes or a size of 40 Mb.

Another unit candidate was the scene, a well-established unit of film analysis: “a unified action within a film that carries the plot forward, unfolding events in the story and providing new information” (Kuhn & Westwel 2012: 359), with the advantage of presenting clearly defined boundaries, by means of AV textual punctuation (film transition techniques, like fading in and out).

At the very least, a scene transcript provides a support for, if not an alternative to, film analysts’ use of descriptive labels such as the Feliz Navidad scene or the Bookstore scene to identify the various parts of this episode. (Baldry 2016: 3)

Although the scene is usually set in one location and in a single period of time, the problem it poses as a unit of analysis is its complexity and variability. Some scenes can include several characters doing all sorts of things connected with language variation. This kind of complication makes it extremely difficult to foresee all of the possible combinations to be included in the design of a questionnaire that researchers need to answer in order to provide the required metadata tags. For example, if the questionnaire’s items (Zabalbeascoa & Sokoli 2017: 6) provided a multiple-choice tagging system, the options would be too many to be practical and never enough to cover all cases given the amount of different types of scenes and L1/L3 combinations—especially considering that film directors attempt to be creative in developing innovative types of scenes. Such variation also affects the length a scene can have; some scenes can be very long and internally complicated. Baldry (2016) identifies the problems involved in taking the whole scene as the unit of analysis and advocates for something that he calls the mini-scene. The mini-scene can be advantageous for describing certain AV products, like TV drama with multiple plotlines, and is apparently quite closely related to our vision of Conversation as an attempt to find a smaller, more stable, unit than the scene.

Within the Trafilm project, the smallest L3 unit is the utterance (word or phrase) bound by pauses or an utterance in another language (L1 or a different L3). The complication of the utterance is that certain utterances can be repeated several times within a dialogue or conversation without really
adding anything new from the point of view of analysis. This led us to consider a more meaningful unit that could be called L3-instance. An L3-instance is one or more L3 utterances provided they all share the same properties. For example, if three characters say “ciao” (L3 Italian) in the same conversation (or one character says it three times at different points) and all of the other L3 variables remain the same, we will consider this to be a single instance of L3-Italian rather than three.

The L3-instance itself could then be taken as a unit of analysis but the problem of L3-instances is that they lack context, pragmatic sense and purpose. L3-instances are often as short as a single word. The same word may appear in different places in the same film or even in different films with very different effects and implications. So, we need something more than just the metadata that characterises an L3-instance as belonging to one language or another, and how competent the utterer is, and whether it is comprehensible to the audience, etc. We need to know its justification and purpose in the film to be able to appreciate its value more completely. This is not to say that the L3-instance is not an essential level of analysis, and Trafilm reflects this by providing as many as 15 tagging options, or variables, for L3ST and 20 for L3TT, levels 3 and 4, respectively (see figure 4).

This is in keeping with scholars who have theorised about utterances (or, in our case, instances), such as Grice and Bakhtin. Grice (1975, 1989) proposed four maxims to account for a collegial conversation in which utterances are understood. For Bakhtin (1986), there are four accepted properties for utterances, including dialogicity and generic form. Essentially, Grice and Bakhtin agree that an utterance means very little without a proper understanding of its immediate context.

From our analysis it might be inferred that a suitable unit for the purpose of full analysis should be a manageable chunk of oral or written language that is smaller than a scene and more meaningful than an utterance or instance of L3. To couple the practical needs of creating a database within the Trafilm project and the desire to be in line with theories from essential scholarly work in the field (like Bakhtin and Grice, who specifically deal with dialogue and conversation analysis), we have chosen to call our unit “conversation”. This term is chosen to show its intention to be in line with the findings of conversational analysis and discourse analysis, as well as Bakhtin's proposal for dialogism, capable of including written instances as well as oral utterances into his analysis.

A Conversation may be an oral or written exchange (e.g., question and answer) or a pragmatically significant utterance (e.g., a question-invitation like
“Coffee?”). It should also have clear boundaries marking its beginning and end; for example, a greeting, an interruption, a scene boundary, a change of topic, a shift to a new addressee, or a conversational close for oral instances, and signs, letters, notes, placards, billboards, subtitles, captions or inserts that constitute an announcement or utterance for written instances. This unit might consist of: (1) a sequence of turns in a dialogue among various speakers (example 1), or just one turn that can be understood pragmatically without the other turns (e.g., explicitly formulated invitations or threats); or (2) a monologue (or a longish single turn or utterance with monologue-like traits) with unity of purpose and or topic; for example, a complete statement by one speaker expressing an opinion; or an answer that makes sense as preceded by a question (either present or implied, as in example 2). Whether it is one turn or more, or one speaker or more, what matters is that a Conversation provides sufficient contextual clues for pragmatic, rhetorical and stylistic analysis. In example (1), we can see how the shop owner and Vianne take turns in their dialogue, whereas example (2) is an instance of a monologue where Captain is the only speaker.

Each unit of analysis is uniquely identified by its start time and duration within a film. For our concept of Conversation—for Trafilm—there is no minimum duration, and for the purpose of analysis it is easier to break down complex film dialogues into shorter conversations that can lend themselves to more straightforward labelling (e.g., to cut up a dialogue that includes a salutation, a joke and a threat into three Conversations, one for each part that can be labelled accordingly). This is partly why we recommend a maximum of 4 minutes per conversation unit, given that even within the field of scriptwriting dialogues are not recommended or found to be longer (Bordwell 2006: 57-58). In the Trafilm project, therefore, longer conversations with two or three distinct parts are duly split, but we allow for exceptional cases.

From 1930 to 1960, most films averaged 2 to 4 minutes per scene, and many scenes ran 4 minutes or more... In films made after 1961 most scenes run between 1.5 and 3 minutes. The practice reflects the contemporary screenwriter’s rule of thumb that a scene should consume no more than two or three pages (with a page counting as a minute of screen time). The average two-hour script, many manuals suggest, should contain forty to sixty scenes. In more recent years, the tempo has become even faster. All the Pretty Horses (2000) averages 76 seconds per scene, while Singles (1992) averages a mere 66 seconds. One reason for this acceleration would seem to be the new habit of getting into and out of the scenes quickly (Bordwell 2006: 57-58).
In figure 1, we can see features (elements, each represented by an X) of an AV text that might be interesting or challenging for translation or some other AV analysis. The elements could be words or phrases that make up an L3-instance (or even some other audiovisual device used as an ingredient of film composition). However, these instances just seem to be scattered along the timeline of the film; i.e., they require further contextualisation for a more meaningful analysis. By contrast, figure 2 shows the same elements within a film that is made up of scene units. Scenes can be any combination of “silence” and Conversation units. Silence, here, simply refers to total absence of words, written or oral. Thus, silences should be distinguished from pauses as silences constitute a type of conversational boundary whereas pauses are (usually meaningful) conversational components appearing somewhere within a conversation (onset, middle or end). From this point of view, a verbally “silent” moment or scene may include sound effects; and certain (written) Conversations may have no (spoken) sound, e.g., signs, notes, inserts, intertitles, captions.

- Scene type A (no conversation) = silence only
- Scene type B (no silence) = Conversation + Conversation +...
- Scene type C = silence + Conversation + silence + ...

Conversations are made up of words (combined, of course, with accompanying paralinguistic and non-verbal elements, including pauses) that are grouped into utterances. Repeated utterances within a given Conversation may be regarded as a single instance (of L3, for example, but also instances of other AV elements and features).

- Instance type A = 1 utterance
- Instance type B = repeated utterances
This allows for a richer analysis, because elements are not only counted and analysed as they are distributed within a film, but also according to the number and types of scenes and Conversations they appear in. Conversations might be classified by various different criteria, including a typology by conversational functions (regardless of L3), which must not be confused with L3-instance functions. Another potentially interesting criterion is the number of elements appearing within the same Conversation, to distinguish Conversations that have a higher or lesser density of the element under study (L3, metaphor, joke, cultural element, character trait, etc.).
be given by figures 1, 2 and 3. For reasons of space we have not added a figure showing overlapping units of any sort. Scenes, Conversations, instances, words and other (verbal and non-verbal) elements may all overlap occasionally; overlapping is a feature of film language and audiovisual composition (Metz 1974). The important point is that even when Conversations overlap, partially or totally, they can still be regarded as units that have a start time and a duration, making them good candidates to establish clip boundaries when considering criteria for building an AV database, such as the Trafilm project. Overlaps tend to occur at the level of utterances, to reflect a natural characteristic of everyday oral communication. In relation to Trafilm's proposal for including start time and duration as metadata, Baldry (2016: 19) can be cited as a researcher particularly interested in drawing attention to the importance of what he calls “timepoint referencing”. Baldry uses this term to refer to the need for (more than the existence of) a systematic way of matching points in the AV with its transcript.

Example 1. Renting the shop (Chocolat, Lasse Hallström 2000)

| Languages: L1-English / L3-French (in italics) |
| Start Time: min. 5.2 Duration: 28 sec. (Trafilm: 21-40 sec.) |
| SHOP OWNER: Who the hell are you? |
| VIANNE: We’re here about the *patisserie*. We’d like to rent it. And the apartment above. |
| SHOP OWNER: Where are you from? |
| VIANNE: Well, we lived in Andalusia for a while. And, let me see, before that, Vienna. And before that... |
| ANOUK: Before that ... Pavia. |

Example 2. At the dinner table (Um Filme Falado, Manoel de Oliveira 2000)

| Languages: L1-Portuguese / L3-English |
| Start Time: min. 64. Duration: 14 sec. |
| CAPTAIN: I’m beginning to think that the three of you should join to gather and create a more harmonious Tower of Babel, in which we all speak one language and live forever under the shadow of the tree of goodness. |

2. Trafilm metadata entries are designed for start time to be entered as digit with no fractions to show time elapsed from the beginning of the film, e.g. 5 means five minutes from the beginning of the film.

3. Duration is expressed in Trafilm in spans of 20-second intervals to show approximately how long the conversation lasts, e.g. 1-20, 21-40, 41-60 and so on.
3. Conversation as a unit of multilingual translational analysis

Various scholars (Corrius 2008; Corrius & Zabalbeascoa 2011; Martínez-Sierra et al 2010, Meylaerts & Serban 2014) have proven, through descriptive studies, that there is more than one way to render multilingual audiovisual texts (films and TV) in their translated and accessible versions.

Given the multisemiotic nature of these cultural products, the many forms multilingualism may take, the different functions it fulfils, as well as the tremendous variety of reception situations, the answer is not straightforward. Nor can there be a single answer, an algorithm or a protocol that would work in every circumstance (Meylaerts & Serban 2014: 9-10).

The complexity of this issue arises not only from the many ways L3 might be (and is) solved in translation (TT) but already in the whole range of different instances of L3 that can be found in the source texts (ST). Thus, Conversation is adopted as an important unit of analysis for the study of multilingual AV texts. One reason for doing this is to make it possible to provide reliable metadata for a database like the one in trafilm.net or any other AV databases or corpora. Hopefully, the results of descriptive analyses based on such collections of samples will help to refine the existing theoretical models of multilingualism in AVT and compare tendencies across distinct translation types. The Conversation unit also provides the advantage of enabling ST-Conversations to be straightforwardly compared with TT-Conversations. In other words, once a Conversation has been established in the ST, its TT rendering can then be found quite uncontentiously for comparative purposes. An alternative research strategy would involve establishing (identifying) Conversations by looking first at the TT and then finding their ST counterparts, and the same principle of “uncontentious straightforwardness” would still be true. However, this does not mean that the initial process of establishing conversational boundaries can be entirely objective, even though every effort is made to produce convincing arguments for establishing these boundaries. Despite this seemingly unavoidable element of subjectivity, the benefit of providing meaningful contexts for L3-instances is still clear and worthwhile.

As shown in figures 2 and 3, a film is made up of Conversations alternating with all of the silences that there may be before or after each Conversation. The Trafilm project’s methodology involves collecting only ST Conversations (and their TT renderings) that include at least one L3-instance and each one must make sense as a Conversation unit even if it is not the whole dialogue (as mentioned above, a Conversation can be an autonomous part of a larger conversation). In Trafilm, almost all conversations meet the following requirements:
— They have one or more turns, either a short dialogue or a monologue (off or on screen) with clear boundaries (i.e., start and end times can be established straightforwardly), or a piece of onscreen writing.
— They last for no longer than four minutes per Conversation; there is no minimum length.
— We assume that L3-instances originate in the ST and are somehow rendered in the translated version (TT), so the Conversation is taken from the ST, and is assumed to be the same for the dubbed, audio described and subtitled versions (including SDH).
— The ST Conversation must have the same start time and duration as the Conversation in the TT version, even for exceptional cases where there might be L3-instances in the TT (L3TT) without there being any (L3ST) in ST, either through some sort of compensation strategy or by leaving some parts of L1 untranslated in TT (L1 unchanged in TT, perceived as “not L2”, hence L3TT).

The structure and logic of the software for the Trafilm database imposes the need to start always at the ST. This entails some restrictions and some advantages, namely that ST conversations and L3ST-instances will be shared by all TT versions. This makes it possible to compare how L3ST is dealt with in different TT versions, and that is the purpose of the project. From the perspective of

4. To avoid cluttering the diagram with an excessive amount of writing, the option of subtitling for the deaf and hard of hearing (SDH) is implied and included as an option within the more general name of “subtitling”.

MonTI Special Issue 4 (2019: 57-85). ISSN 1889-4178
descriptivism within Translation Studies, first we describe (by providing meta-data) the ST film, then a selection of its Conversations (the ones containing L3-instances), then L3ST and, finally a TT version of the L3ST-instance. Then, other TT versions (dubbing, subtitling, SDH, audio descriptions, or other foreign-languages versions). For the (rare) case an L3TT-instance that has no L3ST counterpart for the same Conversation, the flexibility of the Trafilm metadata forms for entering the data can allow analysts to include it as a case of L3ST=L1 rendered in the TT as L3TT≠L2 (see the Trafilm Guide, 2017, for a full account).

We must not forget that the goal for Trafilm is to find out how instances of L3ST are rendered in foreign-language versions, not (primarily, at least) to find an account for the presence of all instances of L3TT.

Conversation is the name given to the communicative context of an instance of L3. A Conversation may contain more than one L3-instance. All L3-instances belonging to the same Conversation (and their respective TT versions) share the same start time and duration. We might distinguish between two types of Conversations depending on whether they are (monolingual) entirely in L3 (examples 2 and 4); or multilingual, with a particular L3 (Spanish, in example 5) combined with L1 (English, in examples 5 and 7) and/or some other L3 (e.g., Catalan, in example 7).

A) Monolingual L3 Conversations consist of a single L3, which might be spoken by the same character or by different characters, i.e., by one or by more than one. Thus, we might distinguish the following subtypes:

a.1) Monolingual Conversation spoken by just one character. In example (3), the L3-Spanish Conversation takes place at the end of the film when Blanca is being buried.


| Languages: L1-English / L3-Spanish |
| Start Time: min. 100. Duration: 20 sec. |

OFF-SCREEN VOICE (echoing Blanca’s words): Vamos a dejarlos en la tierra pero esta tierra ahora nos pertenece, compañeros. Y aquí tenemos que sacar la fuerza para seguir luchando porque la batalla es larga y son muchos, pero nosotros somos muchos más, siempre seremos muchos más. El mañana es nuestro, compañeros.

5. In Trafilm notation, “different”, as in “different characters” or “different languages”, means more than one. Thus, “different” is complemented by “same”, which refers to just one (see multilingual Conversation type b.2 and example 6).
a.2) Monolingual Conversation spoken by more than one character. Example (4) is from the beginning of *Um Filme Falado* and the main character, Rosa Maria, is talking, in L3-French, to a fisherman in Marseille, i.e., there are two characters, both speaking in one L3:

Example 4. First Port of Call – Marseille (*Um Filme Falado*, Manoel de Oliveira 2000)

<table>
<thead>
<tr>
<th>Languages: L1-Portuguese / L3-French</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Time: min. 9. Duration: 19 sec.</td>
</tr>
</tbody>
</table>

ROSA MARIA: Et vous, comment vous appellez?  
FISHERMAN: Moi, je m’appelle Joao.  
ROSA MARIA: C’est comme mon mari, aussi. Il s’appelle Joao. Moi et ma fille on va le retrouver.  
FISHERMAN: Très bien. Et où est-ce qu’il se trouve?  
FISHERMAN: Et vous? Vos prénez le bateau?

Although example (4) is a clear sample of a monolingual conversation, it is worth mentioning that this film, as the title indicates, is about language and, thus, five different languages (Portuguese as L1 and French, English, Italian and Greek as L3) can be heard. The film is about a Portuguese history teacher (Rosa Maria) who takes her daughter (Maria Joana) on a cruise to meet her father in Bombay. It is a kind of guided tour in which the histories of each country are discussed and the architecture and famous religious and cultural locations are presented. What is interesting about it is that every time they anchor at a port of call a different language is heard, representing each local identity. In example (4), French is spoken because they are in Marseille. The second part of the film consists mainly of a dinner conversation between the ship captain and three passengers on board from different countries: a French businesswoman, a former Italian model and a Greek singer. Each character speaks their own language and they comprehend each other easily. The linguistic element is used to characterise different identities.
B) Multilingual Conversations include various combinations, as shown in examples (5) and (6).

b.1) A mix of a main language (L1 or L2) and a third language (L3) spoken by different characters; each speaker uses a distinct language. For instance, in the Conversation in example (5), Amelia, the baby-sitter, is speaking in L3-Spanish to Debbie, one of the children, whereas the child is speaking in L1-English.

Example 5. Debbie is scared in bed (*Babel*, Alejandro González Iñárritu 2006)

<table>
<thead>
<tr>
<th>Languages: L1-English / L3-Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Time: min. 10. Duration: 47 sec.</td>
</tr>
<tr>
<td>DEBBIE: Amelia. Amelia? Could you leave the light on?</td>
</tr>
<tr>
<td>AMELIA: Pero si ya están durmiendo juntos. ¿No quedamos que se iban a dormir con la luz apagada?</td>
</tr>
<tr>
<td>DEBBIE: But I'm scared.</td>
</tr>
<tr>
<td>AMELIA: No pasa nada, hijita.</td>
</tr>
<tr>
<td>DEBBIE: But I'm scared that what happened to Sam is going to happen to me.</td>
</tr>
<tr>
<td>AMELIA: No, hijita, esto no te va a pasar a ti. Esto no te va a pasar.</td>
</tr>
<tr>
<td>DEBBIE: Sam died while he was sleeping.</td>
</tr>
<tr>
<td>AMELIA: No, baby, no, no. That only happens to some babies when they're really, really little. Your brother is already high in heaven, baby. I'll sit here with you. But you can't talk any more, ok? Go to sleep.</td>
</tr>
</tbody>
</table>

b.2) A mix of a main language and a single third language spoken by the same character, for example by code-switching or self-translation. This case is illustrated in example (6) by mixing L1-English with a few words in L3-French.

Example 6. At the Perfume Shop (*An American in Paris*, Vincente Minnelli 1951)

<table>
<thead>
<tr>
<th>Languages: L1-English / L3-French</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start Time: min. 44. Duration: 6 sec.</td>
</tr>
<tr>
<td>JERRY (addressing the shop assistant): ‘Mademoiselle’, there’s no happier man in Paris than ‘monsieur’ Mulligan.</td>
</tr>
</tbody>
</table>
b.3) Two or more different L3 languages all spoken by the same character. In (7), Man 3 speaks in L3-Catalan and L3-Spanish. Three Catalan men are at a table drinking and talking in Catalan about militiamen when a person from another table (David) hears them and goes to talk to them. He addresses them in Spanish, so one of the men (otherwise speaking in Catalan to his colleagues) answers him in Spanish.

Example 7. At the bar (Land and Freedom, Ken Loach 2004)

| Languages: L1-English / L3-Catalan / L3-Spanish (italics) /L3-non-native Spanish (bold) |
| Start Time: min. 80. Duration: 56 sec. |

MAN 1: Ahir el Martí, un amic meu, va tornar del front, mort, mort de gana. Allò és un desastre!
MAN 2: De quin front?
MAN 3: A mi m’han dit que no saben ni carregar les armes.
MAN 2: Això potser és massa, no? A mi m’han assegurat que les dones milicianes en comptes de matar als feixistes, que és el que haurien de fer, el que fan es que se’ls carden!
MAN 3: No fan res. Són una merda!
MAN 1: Són un desastre els milicians. Son quatre arreplegats que han agafat quatre fusells i galons i es pensen...
DAVID: ¡Es mentira! Yo estuve en el frente. Sí, milicianos. (with an English accent)
MAN 3: ¿Eres miliciano?
DAVID: Vienen aquí.
MAN 3: ¿Eres miliciano?
DAVID: Sí.
MAN 1: Si no sap ni parlar. Deixa’l estar.
DAVID: What are you saying? You’re a fucking kid!
MAN 1: A kid?
DAVID: Yea, a kid, niño. You don’t know how... Sit down!
b.4) Two or more different L3 languages spoken by different characters (each speaker using a distinct third language). Example (8) is a Conversation where the policeman speaks in Korean, Gaff speaks Cityspeak (a constructed language made up of words from different real languages), Deckard speaks in English, with Sushi Master acting as the interpreter, in non-native English.

Example 8. At a noodle bar (Blade Runner, Ridley Scott 1982)

| Languages: L1-English / L3-Korean (bold) / L3-Cityspeak (italics) |
| Start Time: min. 9. Duration: 33 sec. |

**POLICEMAN:** Hey, idi-wa! [Ey, come here]
**GAFF:** Monsieur, azonnal Kövessen engem bitte. [Sir, follow me immediately, please]
**SUSHI MASTER:** He say you under arrest, Mr. Deckard.
**DECKARD:** You got the wrong guy, pal.
**GAFF:** Lófaszt, nehogy már. Te vagy a Blade... Blade Runner. [Horsedick, no way! You are the Blade... Blade Runner]
**SUSHI MASTER:** He say you Blade Runner.
**DECKARD:** I'm eating.
**GAFF:** Captain Bryant toka. Me ni omae yo. [Captain Bryant wants to see your mug in front of his immediately!]
**DECKARD:** Bryant, huh?
**GAFF:** Hai! [yes]

In example (8) no subtitles are provided for the Korean or Cityspeak languages so the audience cannot understand what the Policeman and Gaff are saying. However, this reflects the scriptwriter's intention with Sushi Master playing the role of interpreter. In this instance, we have provided (in square brackets) the English translation of Cityspeak6 to help the reader.

It is now clear that a Conversation can have: a) one speaker or more, b) one L3 or more, c) a mix of L3 and the main language (L1 or L2, for ST and TT, respectively), d) and one mode (oral or written) or both, along with any of the possible resulting combinations, e.g., one speaker / one language; one speaker more than one language; two speakers one language; two speakers, more than one language; one L3 in two different modes, etc. Because we include written occurrences we use the term source to include both speakers and written sources such as inserts, subtitles or written props, like placards.

---

Both monolingual and multilingual conversations that contain L3-instances help the audience to distinguish between the in-group and the out-group. Scholars such as Corrius (2008) have linked the presence of a third language in an AV text to the presence of a foreign group (out-group), with which the audience do not feel identified; namely a group with a different identity from the main group with which the audience feels identified (in-group). For instance, in example (3) there is an off-screen voice echoing Blanca’s words. Blanca, who embodies the revolution in the Spanish Civil War is from Spain and speaks Spanish in order to represent her Spanish identity: “one of the main characters had to be Spanish and must in some way embody the revolution” (Fuller 1998: 101). Similarly, in example (5) Amelia represents the Mexican identity as opposed to the US one represented by the kid, Debbie.

4. Conversation: A relevant concept for AVT studies beyond the case of L3

Examples (9) and (10), drawn from the romantic comedy Love Actually (Richard Curtis 2003), illustrate the importance (and the complexity) of the Conversation concept for the purpose of humour, identity (personal as well as cultural) as a component of character portrayal, and relationships (romantic, in this case), all combined in these examples with associated features of L3-instances, inextricably, in this particular case. This is why examples (9) and (10) are so important from the point of view of highlighting the value of the concept of Conversation beyond the particular case of L3. Taken together, these two examples reveal Aurelia’s strong female identity both physically and psychologically, in contrast to Jamie, who is identified as a man who cannot stand the cold water like Aurelia, disorganised in his working habits, and quite insecure.

<table>
<thead>
<tr>
<th>Languages: L1-English / L3-Portuguese</th>
<th>Start Time: min. 43. Duration: 119 sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 JAMIE: Thank you.</td>
<td></td>
</tr>
<tr>
<td>2 AURELIA: Nao! Eu peco imensa desculpa.</td>
<td></td>
</tr>
<tr>
<td>4 AURELIA: Qué desastre.</td>
<td></td>
</tr>
<tr>
<td>5 JAMIE: Just leave them, please! They're not important. They're not worth it! Stop! Stop! It's all just rubbish. Just leave it. Oh, God, she's in... And now she'll think I'm a total spaz if I don't go in too.</td>
<td></td>
</tr>
<tr>
<td>8 AURELIA: [Fuck - it's cold!]</td>
<td></td>
</tr>
<tr>
<td>9 JAMIE: Fuck - it's freezing! Fuck!</td>
<td></td>
</tr>
<tr>
<td>10 AURELIA: This stuff better be good.</td>
<td></td>
</tr>
<tr>
<td>11 JAMIE: It's not worth it, this isn't bloody Shakespeare.</td>
<td></td>
</tr>
<tr>
<td>12 AURELIA: [I don't want to drown saving some shit my grandmother could have written.]</td>
<td></td>
</tr>
<tr>
<td>13 JAMIE: Just stop. Stop.</td>
<td></td>
</tr>
<tr>
<td>14 AURELIA: [What kind of idiot doesn't make copies?]</td>
<td></td>
</tr>
<tr>
<td>15 JAMIE: I really must do copies. You know there'd better not be eels in here. I can't stand eels.</td>
<td></td>
</tr>
<tr>
<td>17 AURELIA: [Try not to disturb the eels.]</td>
<td></td>
</tr>
<tr>
<td>18 JAMIE: Oh, what the hell is that?</td>
<td></td>
</tr>
</tbody>
</table>

Example 10. After the lake scene (*Love Actually*, Richard Curtis 2003)

<table>
<thead>
<tr>
<th>Languages: L1-English / L3-Portuguese</th>
<th>Start Time: min. 45. Duration: 87 sec.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 JAMIE: Thank you. Thank you so much. I know. I'll name one of the characters after you.</td>
<td></td>
</tr>
<tr>
<td>3 AURELIA: [Maybe you could name one of the characters after me. Or give me 50% of the profits.]</td>
<td></td>
</tr>
<tr>
<td>5 JAMIE: Or I could give you 5% of the profits.</td>
<td></td>
</tr>
<tr>
<td>6 AURELIA: [What kind of book is it? Romance?]</td>
<td></td>
</tr>
<tr>
<td>7 JAMIE: Yes. It's, erm...</td>
<td></td>
</tr>
<tr>
<td>8 AURELIA: [Thriller... crime... ]</td>
<td></td>
</tr>
<tr>
<td>10 AURELIA: [Frightening?]</td>
<td></td>
</tr>
<tr>
<td>11 JAMIE: Er, scary? Yes, sometimes scary. And, er, sometimes not. Mainly scary how bad the writing is. Mm.</td>
<td></td>
</tr>
<tr>
<td>13 AURELIA: [I'd better get back to work. And then later you'll drive me home?]</td>
<td></td>
</tr>
<tr>
<td>14 JAMIE: Sure. It's my favourite time of day... driving you.</td>
<td></td>
</tr>
<tr>
<td>15 AURELIA: [It's the saddest part of my day, leaving you.]</td>
<td></td>
</tr>
</tbody>
</table>

The English words in square brackets in examples (9) and (10) are the transcription of the ST subtitles, and the words that can be heard diegetically are actually L3-Portuguese utterances. The L3-Portuguese utterances are not
transcribed because the point is not to compare the Portuguese with their corresponding English ST subtitles, but to give a flavour of the way this was intended for the original (English-language) target audience; i.e., Aurelia is speaking Portuguese, which is scripted as incomprehensible to Jamie and to the English-language spectators. This leads to a distinction between utterances which are subtitled in the ST, and those which are not (e.g., ‘Qué desastre’ is not subtitled in the ST). The implication for the audience is that, when Aurelia’s L3ST utterances are not subtitled, the viewers are invited to experience empathetically the same sense of incomprehension as Jamie. By contrast, when Aurelia’s utterances are subtitled it is because their contents are important for the development of the story, for Aurelia’s character portrayal and for appreciation of the scenes and their intended humour. The scenes are meant to be funny (as well as romantically tense) because of the irony they convey; the similarities between Jamie’s and Aurelia’s turns give the impression that they could almost be considered as some sort of translation (for more on intratextual translation see Zabalbeascoa & Voellmer 2015). The irony is that the characters are completely unaware that much of what they are saying is so similar. Aurelia and Jamie are speaking different languages and the whole point of these two exchanges is that their utterances could be interpreted (and analysed) as turns within the same Conversation, although the result would be a rather odd exchange, indeed. Another interpretation is that each one is having a conversation with him or herself, oblivious of what the other is saying; each set of each speaker’s turns add up to a coherent “monologue”. From this point of view, they are having different, but simultaneous, conversations. Here is a clear example that what unifies the two characters’ utterances into a single conversation (if thus analysed) is only from the point of view of the film communicating to its audience, who has the aid of subtitles to enable full comprehension of the intended irony. Examples (9) and (10) are analysed as two Conversations although they manifest the same audiovisual-rhetorical device because there is a change of scene between them. Within these two examples, though, we can establish smaller Conversation units: in example (9), lines 1–4 both react to situation, lines 5–7 are a monologue, and lines 8–19 are two separate monologues, that can also be analysed ironically as a single Conversation from the audience’s position. In example (10), lines 1–12 talk about the novel, lines 13–15 is a farewell.

As far as cultural references are concerned, it is worth mentioning that they play an important role in AVT, which might be expressed through verbal and nonverbal communication. Thousands of instances of cultural signs expressed both verbally and nonverbally can be found in films; Santamaria (2001), among
others, has discussed at length how the cultural elements, which are commonly used in AV texts, help people to identify themselves as well as their interlocutors. An AV text is a product of the source culture, and this might be perceived (nonverbally) by the target culture audiences through its images: landscapes, vehicles, the characters themselves, clothes, colours, food, etc., and (verbally) through the language or languages in the text. For example, in the film *Spanglish* (James L. Brooks 2004)—a film that makes its plurality and linguistic variety the reason for its existence (Díaz-Cintas 2014: 145)—there is a constant clash between the cultures of Mexico and the USA, perceivable nonverbally right from the beginning with the images of Mexico, but also through verbal language as the note Flor leaves in the kitchen to her daughter that reads: “Cristina, pon queso en la tortilla y ponlo en el microondas por un minuto” (Cristina, put cheese in the tortilla and put it in the microwave for a minute). Tortilla is a cultural reference, a round, thin Mexican bread, usually eaten hot with a filling of meat, cheese, etc. Someone from Spain might misunderstand the word by mistaking it for an omelette, which is also named tortilla; and somebody with no knowledge of Spanish or Mexican culture might not have a clue what tortilla is. Furthermore, the food mentioned here contrasts with the (very American) sandwich eaten by the US-citizen, Mr. Clasky (in minute 58); this is another instance of the cultural signs that distinguish the two clearly marked identities included in the film. In audiovisual communication, verbal language co-occurs with the non-verbal elements that embody a set of signs, values, customs, etc., as part of a particular culture. As Di Giovani (2003: 210) herself has written “the interaction of verbal language with other audiovisual signs, even if conventional, is therefore all the more important in shaping cultural representations”.

Cultural references are a well-known translation problem, especially when they are instrumental in conveying elements of personal and cultural identity. For example, characters are portrayed linguistically through the languages and expressions they use, and non-linguistically, in their wardrobe or the tunes they whistle, for instance. In any case, there is a close relationship between L3 and cultural issues. Fictional AV speakers tend to use their language to produce and reflect certain identities and cultures, and, in turn, identity helps to understand the relationship between language and culture. The topic deserves more research, though (Bucholtz & Hall 2004). Cultural references might be studied and analysed for a particular purpose, and therefore there would be a need (as in the study of L3-instances) to use a unit of analysis like the Conversation, which could be used as a tool to fulfil this purpose.
After examining the practical and theoretical consistency of the concept of Conversation for analysing multilingual AV texts, it seems not only logical but almost necessary for the concept to work for other elements of AV texts, and not only L3, for the simple reason that different conversational features are not always separable. As a unit of analysis, the Conversation is capable of providing a communicatively meaningful context. It provides the exact location and duration of each L3-instance and other features that can help define and classify Conversations. Thus, Conversations can be defined and classified according to an array of different variables, while carefully distinguishing feature types (e.g., L3, metaphor or polysemy) and conversational types (e.g., interview, argument, small talk, cross-examination), as explained in the Trafilm Guide (Zabalbeascoa & Sokoli 2017). The features and types most likely to yield productive analyses are the ones most closely linked to pragmatics, such as humour, persuasion, irony and politeness, and discourse, too, in features such as identity and ideology. Authors such as Chaume (2004) have called for improved collaboration between translation studies and film studies and we believe that proposals like the one presented here for the concept of Conversation are a step forward in precisely that direction. Authors like Hatim and Mason (Hatim & Mason 2000; Mason 2005) have linked the importance of context, discourse and conversation analysis to a notion of Conversation, which they refer to as discourse or dialogue. Victor Raskin (1987: 12) is a key scholar in linking humour studies to linguistics and pragmatic units like conversations, his contribution being that of linking conversations to the notion of scripts, defining script as “a large chunk of semantic information surrounding the word or evoked by it”, which is meant to include routines and common ways to do things and go about activities. Scripts are understood as patterns of behaviour, and, interestingly, the word, adopted by linguistics, is taken from film scriptwriting, where actions and behaviours are established on the basis of analogous events recorded or remembered from the past. Thus, a polite conversation would include one script, and a humorous conversation (by Raskin’s script-opposition theory for humour) by two opposing scripts, conflating within a single conversation (i.e., the joke). From this point of view, too, then, a joke can be seen as a Conversation type; and, of course, some jokes include the presence of L3-instances. Last but not least, we recall the authors first cited herein, Bakhtin, Grice and Sidnell, whose theories revolve around notions that overlap considerably, when not entirely, with the concept of Conversation to explain, among others, phenomena such as irony, sarcasm and paradox.
5. Conclusions

After considering several options for unit of analysis from a practical and a theoretical point of view, the most convincing option is to adapt and apply the concept of Conversation to the specific needs of research in multilingualism as a phenomenon of AVT. When Conversation is defined as fully and flexibly as it has been here, it can also be used as the criterion for making clips with a view to building an audiovisual corpus or database; i.e., each clip can contain one Conversation, sharing with it the metadata of start time and duration within a given audiovisual text such as a film. Of course, L3-instances (utterances or groupings of utterances that differ only in some small feature) are essential units too, but sometimes too small or isolated to provide a full enough analysis, so the information they provide needs to be supplemented with a proper description of the Conversation they belong to even when the Conversation is made up solely and entirely of a single L3-utterance (examples 2 and 3). Thus, the Conversation is seen as the smallest possible unit that contains instances of L3 with enough pragmatic and contextual information.

A functional approach also helps to see the importance of having Conversation as a separate unit. If we only analyse L3-instances as they appear in a film (figure 1) we can focus on the functions of L3 presence at different points in a film or TV show. But several of the examples presented here have shown that the functions of L3, on the one hand, and the functions of the Conversation that they are a part of, on the other, are two different things. For example, L3 presence may wish to convey otherness and identity as a part of character portrayal (examples 9 and 10) and, thus, signalling the out-group, which is deemed foreign, but the function of the Conversation is to be comical and also serve plot progress as the two characters become more intimate. L3 presence may sometimes have a comical function (sometimes simply by the fact of resorting to a foreign language at a given moment), but in this case it is not the presence of L3 that produces humour but rather the way the Conversation works. This means that one way of gaining insight into film language, L3 presence, and how it is rendered in foreign versions, is to study how and when L3 appears in Conversations depending on their functions as well as other typologies (e.g., Conversations as jokes, small talk, introducing new acquaintances, exclamations, racist discourse, etc.) to discover correlations between L3 presence and types of Conversations.

Despite our insistence here on the verbal dimension of L3 and its Conversations, there is still a need to see how foreignness is conveyed by paralinguistic and non-verbal elements (props and gestures) often in close combination with the words. Hence, cultural signs represented both verbally
and non-verbally also play an important role when featuring identities. The different cultural signs that help characterise different identities might not always be easily decoded by the audience; it depends on how distant the foreign culture is from the audience. Another important non-verbal consideration is to show that, even in a study like this, scenes are made up of silent moments as well as verbally expressed dialogue and both are audiovisual-text constituents that interact with each other.

Our closing remark is to say that we trust that our findings regarding the potential usefulness of “Conversation” as a unit of analysis, within scenes (figures 3 and 4) and sometimes combined with (verbal) silences, of conversational functions and types to provide a richer, more contextualised analysis of L3 can also be applied to research into other AV and AVT elements and features, especially the ones most closely related to pragmatics and discourse (e.g., forms of address, turn taking, expletives, onomatopoeia, puns and one-liners).

References


Conversation as a unit of film analysis. Databases of L3 translation and...


Online Resources


Filmography

A Passage to India (David Lean 1984)
An American in Paris (Vincente Minnelli 1951)
Babel (Alejandro González Iñárritu 2006)
Blade Runner (Ridley Scott 1982)
Beauty and the Beast (Gary Trousdale & Kirk Wise 1991)
Bread and Roses (Ken Loach 2000)
Breaking Bad (AMC 2008-2013)
Chocolat (Lasse Hallström 2000)
Django Unchained (Quentin Tarantino 2012)
Game of Thrones (HBO 2011-)
Inglourious Basterds (Quentin Tarantino 2009)
Land and Freedom (Ken Loach 2004)
The Lord of the Rings: The Fellowship of the Ring (Peter Jackson 2001)
The Lord of the Rings: The Two towers (Peter Jackson 2002)
The Lord of the Rings: The Return of the King (Peter Jackson 2003)
Lost in Translation (Sofia Coppola 2003)
Love Actually (Richard Curtis 2003)
Spanglish (James L. Brooks 2004)
Todo sobre mi madre (Pedro Almodóvar 1999)
Um Filme Falado (Manoel de Oliveira 2000)

NOTAS BIOGRÁFICAS / BIONOTES

Patrick Zabalbeascoa Terran is Full Professor at the Universitat Pompeu Fabra. He lectures and researches in translation theory and audiovisual translation, and FLL through translation. He has published widely and internationally in these areas and played a leading role in several EU funded and Spanish research projects, such as ClipFlair. He is currently working on the Trafilm.net project. He is well-known for his research in the field of humour translation ever since his PhD (1993) proposed a typology of humour translation and theoretical model of Priorities and Restrictions based on a study of the
Conversation as a unit of film analysis. Databases of L3 translation and...

BBC sitcom Yes, Minister. For further details visit <http://upf.academia.edu/PatrickZabalbeascoa>

**Montse Corrius** holds a PhD in Translation and Interpreting from the Autonomous University of Barcelona and a degree in English Philology from the University of Barcelona. She is a Senior Lecturer at the University of Vic – Central University of Catalonia (UVic-UCC), where she teaches English for specific purposes at the undergraduate level and audiovisual translation at the postgraduate level. She is a member of the research group TRACTE (Traducció Audiovisual, Comunicació i Territori), where she leads the line of research on translation. Her main research interests include audiovisual translation (with a special focus on multilingual texts) as well as advertising translation, language learning and lexicography. She has published several articles and lectured on these areas of research. She is one of the main researchers of the funded project trafilm.net.