

# Chapter 8

## Between audiovisual translation and localization: The case of *Detroit: Become Human*

Laura Mejías-Climent

Universitat Jaume I

The concept of audiovisual translation (AVT) is changing continuously in the current technological landscape, in which localization has emerged as a key process to adapt different types of modern multimedia products. Whether AVT encompasses localization, or vice versa, or they can be conceived as different fields instead, remains unclear. It is not my intention to find a unique solution to this unanswered question, but rather to shed some light on the convergences of AVT and localization from the specific perspective of dubbing, in a product that, in turn, raises some questions about the genre it belongs to: the graphic adventure *Detroit: Become Human* (Quantic Dream, 2018). This chapter aims to highlight some of the differences and convergences between AVT and localization analyzing the dubbing synchronies applied in a video game belonging to a genre closer to traditional movies, compared to other adventure games, due to the strong presence of cinematic scenes and the lower level of interaction. The results will indicate that, even though clear differences can be found between the localization of a game and translation of non-interactive products, the convergences in terms of dubbing synchronies, particularly in the cinematographic scenes, are quite evident.

### 1 The current landscape of audiovisual translation

#### 1.1 Introduction

Within the landscape of translation studies, and as a professional practice too, audiovisual translation (AVT) represents an umbrella term encompassing different



translation modes. These modes depend on the nature of the original product and the translated one (Hurtado Albir 2011), as well as the technical methods used to transfer the linguistic message from an original audiovisual text to the target one (Chaume 2004).

Traditionally, AVT is divided into two broad groups (Chaume 2013): revoicing and captioning. The former refers to those modes in which an additional soundtrack is included in the original product: dubbing, voice-overs, simultaneous interpreting of movies, free commentary, fandubs and audiodescription. The latter, captioning, encompasses those modes based on text inserted or next to the screen in which the original product is shown: subtitling, surtitling, respoken, subtitling for the deaf and hard of hearing and fansubbing. Both lists keep broadening together with technological advancements, products and market preferences.

AVT has developed at a rapid pace in the last decades, not only as a professional practice, but also as a research field. “Technological developments have brought about new audiovisual transfer modes, or new combinations of the latter” (Chaume 2018a: 41). With the exponential increase in the audiovisual production during the last decade, the concept of AVT has faced the challenge of dealing with many different types of audiovisual products and translation modes, and with different passive and active forms of consumption. All of this has caused the emergence not only of additional AVT modes, but also of other terms that coexist with that of AVT, sometimes referring to the same concept, sometimes evidencing our technological changing reality more specifically. In all of them, translation is the underlying concept that accounts for the access to any audiovisual product to a different audience.

In the current and changing technological landscape of audiovisual production, some of the “characteristics of AVT that are expanding the borders of the concept of translation” are the following, as Chaume (2018b: 88) points out: not only interlingual and intralingual transfer of content takes place in AVT modes such as subtitling (interlingual) or subtitling for the deaf (intralingual), but also intersemiotic transfers, in the case of audio guides for museums, sign language, and audiodescription. Transadaptation (Neves 2005; Gambier 2003) “could encompass all AVT modes known to date” (Chaume 2018b: 92) and, for Pruys (2009), it consists of two variations of the same topic. Transcreation implies a high level of creativity to “tilt the balance towards the target audience” (Bernal-Merino 2015: 90) and can be understood as another form of semiotic adaptation dealing with many different types of audiovisual texts. Transmedia narratives, rewritings (Bernal-Merino 2015) and media adaptations such as remakes are common practice nowadays. Finally, localization has emerged either as a synonym of a broader

concept of AVT, or as a different professional field dealing with the adaptation of software, websites and video games to a target culture (see §2).

Regarding the concept of *localization*, the borders between AVT and localization are no longer clear, if they ever were, and the use of different AVT modes can be found in any multimodal product nowadays. Whether AVT encompasses localization, or vice versa, or they can be conceived as different fields instead, remains unclear. It is not my intention to find a unique solution to this unanswered question, but rather to shed some light on the convergences of AVT and localization, from the specific perspective of dubbing, in a multimedia product that, in turn, raises some questions about the genre it belongs to: the graphic adventure *Detroit: Become Human* (DBH), developed by Quantic Dream.

## 1.2 The aim of the present chapter

In order to trace some similarities, but also to point out some of the differences, that can be observed in the final localized version of a video game and that of any traditional movie, the results of a case study will be discussed in the following pages. The aim is to offer a concise analysis of the similarities between the dubbing of traditional movies and that of graphic adventures, focusing on the case of the video game *Detroit: Become Human* and, more specifically, in the types of synchronies used in the cinematic scenes of this video game. A total of 20 hours of gameplay of this interactive audiovisual product was analyzed, focusing on the different game situations and dubbing synchronies used in each of them, especially, in the cinematic scenes.

To contextualize the analysis, first, the definitions of AVT and localization, and their convergence, will be discussed, as introduced above. Secondly, the characteristics of the AVT mode of dubbing in movies will be reviewed. Thirdly, the same will be done for the features of dubbing in video games, followed by a discussion of the game situations in which the analyzed video game is articulated. Finally, the results of the empirical and descriptive analysis will be presented to determine the convergences between the dubbing of a traditional movie, a fully-interactive video game (Mejías-Climent 2017; 2019) and the case of the graphic adventure *Detroit: Become Human*.

## 2 Audiovisual translation within localization or vice versa?

As introduced above, AVT encompasses nowadays a wide range of translation modes (Chaume 2004: 31; Hurtado Albir 2011: 54) that challenge the traditional

concept of translation in the strictest sense of “linguistic transfer”. The ever-growing variety of multimodal products that keep breaking into the market require translation practices to adapt, in order to accommodate to these continuous changes in the technological configuration of the products and the way they are consumed. In fact, nowadays the term *localization* itself can encompass

both consolidated as well as new groundbreaking interlingual, intralingual and intersemiotic “audiovisual translation” practices, namely dubbing, subtitling, surtitling, respeaking, audiosubtitling, voice-over and partial dubbing, simultaneous interpreting in film festivals, free-commentary, subtitling for the deaf and the hard of hearing, audio description for the blind and visually-impaired, fansubbing and fandubbing. (Chaume 2018b: 84)

Nonetheless, the term localization emerged in the 1980s when software developers identified the need to adapt their products to expand their markets to other cultures (O’Hagan & Mangiron 2013: 87). With the rapid growth of the game industry, especially since the 1990s, the term localization has settled among professionals and it is generally understood as a more complex process of adaptation, beyond a mere linguistic translation (Bernal-Merino 2006). The author highlights the fact that, however, this term does not refer to anything new that the concept of translation did not include already. Since its use is widely spread among the industry, it seems necessary to accept it within translation studies but always preceded by “linguistic” to differentiate it from the whole adaptation and industrial process described by professionals such as Esselink (2000) or Maxwell-Chandler (2005).

Other scholars and professionals such as Muñoz Sánchez (2017), Granell et al. (2015) or Mangiron & O’Hagan (2006) emphasize the idea that localization includes the adaptation of the interactive product on many different levels (including the linguistic one) to make it meet the needs of the target market completely.

The debate is open about the link between the concept of localization and the field of AVT. As Vázquez Rodríguez (2018: 9–23) points out, a group of professionals defend the position of localization as a differentiated area, since they understand translation from a reductionist and linguistic perspective (cf. Cadieux & Esselink 2004). In addition, some other professionals and scholars also prefer to separate the idea of localization from any other translation mode because of the broad range of adaptation processes it implies, the particularities in the professional practice and the type of product being translated (Pym 2016; Méndez González 2015; Jiménez-Crespo 2013; Mata Pastor 2005).

On the other hand, the term localization does not add anything new to the concept of translation defended by scholars such as Bernal-Merino (2006; 2015) and

O'Hagan & Mangiron (2013). This is the same position that Vázquez Rodríguez (2018) takes himself, acknowledging that the term localization is widely used in professional spheres and, as such, it can be used in research as well, as a label for the translation practice that deals specifically with video games, software and websites. Localization and translation should not be separated, but rather, the first can be included in the latter as another translation mode with defining characteristics.

Having said that, is localization another mode within the field of AVT or is it a completely different area? In both cases, a multimodal product with very similar characteristics—except for the interactive dimension—is translated. Game localization might liken to AVT in the sense that it encompasses some other modalities itself, such as dubbing for cinematic scenes or subtitling dialogues, in addition to some other practices to modify legal, technical or external contents, or even accessibility practices.

However, it should be taken into account that AVT is a very broad field under which any multimodal text can find its particular translation practice. From my point of view, it all depends on the perspective from which both processes are considered. As a professional practice, game localization can be understood as the hypernym under which different translation modes can be gathered. Localization seems to seek for a clear differentiation within the industry, related to the concept of transcreation (O'Hagan & Mangiron 2013). Although no clear definition has been established so far with empirical studies to validate the term transcreation (Bernal-Merino 2015). But that lies outside the scope of these pages. In academic circles, AVT is conceived as the process of adapting any kind of multimodal product, among which video games are included. As O'Hagan & Mangiron (2013) rightly summarize, and as pointed out above,

The emergence of new media resulting from the convergence of technologies is seeing the previously separate domains of localization and AVT come together to cater for the new type of products needing to be prepared to go global. Whether AVT subsumes localization or vice versa remains to be seen, although it is now widely acknowledged that AVT is fast gaining a foothold within Translation Studies. (O'Hagan & Mangiron 2013: 106)

Given the ambiguity in setting one field or practice within the other, I follow Vázquez Rodríguez's (2018) approach, who proposes to adapt AVT research practices to include the interactive dimension and playability in an empirical study, in order to determine the repercussion that both might have in the translation of interactive audiovisual products. Thus, it is not necessary to establish a com-

pletely different paradigm for localization, but just to adapt the existing research approaches used in the AVT field.

When focusing on the analysis of video games, an additional semiotic channel needs to be taken into account in the configuration of the audiovisual product under study (Mejías-Climent 2017). Consequently, some differences in the translation of a movie and a video game can be identified with descriptive analysis. As the following case of study shows, there are clear differences in the product as a whole. But similarities in some specific areas are more prominent than differences, especially in the case of a graphic adventure, which is, ultimately, a movie offering choices to the viewer (thus, a type of narrative and audiovisual content including interaction).

### 3 Dubbing as an audiovisual translation mode within the process of localization

As is widely known, dubbing “consists of replacing the original track of a film’s (or any audiovisual text’s) source language dialogues with another track on which translated dialogues have been recorded in the target language” (Chaume 2012: 1). This translation mode is included in the localization of AAA video games, those with a large budget whose developer can afford a full localization including the adaptation of box and docs, in-game text and audio files and dialogues.

There are certain differences in the dubbing of a video game compared to that of a movie. In particular, the following aspects can be mentioned (Mejías-Climent 2019): There is no linear script, but numerous dialogue strings that can be grouped, depending on the character, the setting or some other criteria. Thus, there is no division in takes or loops, as it happens in countries such as Spain or Italy (*anelli*), where cinematographic scripts are traditionally divided to facilitate the dubbing actors’ task. No dubbing symbols are generally used, although they can be introduced later in the studio by the dubbing director; no TCRs are used, since there is no linear development of events in a video game. Finally, in most cases no images are used to perform the dubbing in the studio. Sound engineers use audio waves sometimes, but in the case of translators, they will never have access to images accompanying the dialogue strings.

Apart from these differences, the results in the dubbing of a modern video game and that of a movie seem to be very close—with some striking but rare exceptions, such as *Arizona Sunshine* (Vertigo Games, Jaywalkers Interactive, 2016) or *Age of Pirates* (Akella, 2006).

Dubbing is a historical and extended practice in countries such as Spain, France or Italy, among others, in which all dubbed products must meet a series of quality standards to be consumed successfully by what could be considered an ideal spectator (Chaume 2007). Among them, synchrony is one of the most prominent characteristics of dubbing.

### 3.1 Characteristics of cinematographic dubbing: Synchronies

Dubbing synchronies represent coherence between what is heard (in this case, a soundtrack containing dubbed dialogues) and seen on the screen. As Chaume (2007: 73) points out, “respect for mouth articulation (phonetic or lip-sync), body movements (kinetic synchrony) and the duration of the translation to match the lines spoken by the screen actors (isochrony), constitute one of the cornerstones of dubbing.”

As far as traditional audiovisual products are concerned (this is, those in which there is no interaction: movies, series, TV shows, etc.), the three types of synchronization above mentioned have been studied in depth: lip-sync, kinetic synchrony and isochrony (Chaume 2012: 68). In the case of dubbing into Spanish, the implementation of the three synchronies depends on the configuration and nature of the audiovisual product and the acoustic and visual codes involved, among other aspects (such as genres and historical conventions). Especially the paralinguistic code and on- and off-sound codes (acoustic channel), and the photographic and the kinetic codes, and the types of shots (visual channel) determine to a greater extent the level of precision with which each of these three types of synchronies needs to be applied (Chaume 2004).

For example, in an extreme close-up, the character’s lip movement is clearly visible onscreen. Thus, the translator and dialogue writer need to use a similar articulation in the translation with, at least, the same number of open vowels and bilabial and labio-dental consonants (Chaume 2012: 73). However, in a long shot, there is no need to use lip-sync, and even isochrony is irrelevant, since the characters’ faces and bodies might not be visible clearly.

This happens in a linear and pre-configured audiovisual product, in which the acoustic and visual codes are determined beforehand. In a video game, however, interaction opens the visual configuration of the audiovisual product to a wider number of options; hence, synchronization does not necessarily work in the same way.

### 3.2 Characteristics of dubbing in video games: Restrictions

Dubbing synchronies play an important role in the dubbing of a video game as well. Video games are the most complex example of a modern audiovisual product. As such, they are very close to traditional movies in some aspects, especially when it comes to cinematic scenes. However, the materials available for translation are not the same as in a movie: neither a traditional script nor a final video is facilitated (Mejías-Climent 2019). Consequently, the process takes place differently and synchronies are rather understood as a series of restrictions (Pujol Tubau 2015: 197).

These restrictions can be transmitted to translators in a maximum number of characters or words per string. They can also be indicated depending on the type of string, being dialogues and sound content more restrictive (thus, the translation needs to resemble the original length as much as possible), and in-game dialogues more flexible.

Nevertheless, many other factors come into play when determining restrictions in the dubbing of a video game: different localization vendors work differently, as shown in Mejías-Climent (2019). In addition, the different agents taking part in the whole process of dubbing play different roles when applying restrictions (thus, synchronies) to the translated text.

It is in the dubbing studio when up to five types of synchronies can be identified and applied to the translated text for the dubbing of a video game, in contrast with the three synchronies described for traditional audiovisual products (see §3.1). Dubbing actors and directors would typically apply the three types of synchronies described for traditional movies if videos were available. However, this is not usually the case, and only sound waves are available in some AAA projects, so they tend to imitate the original sound waves as much as possible, in order to assure a well synchronized dubbed dialogue. Here, up to five levels of restriction can be set, depending on the type of string they dub. It is the role of sound engineers to make the dubbed audio files resemble the original ones as much as possible, according to five levels of restriction that can be understood as the five types of synchrony used in the dubbing of a video game (Mejías-Climent 2017: 105).

*Wild*: No time restriction applies.

*Time constraint (TC)*: The translated utterances must be the same length as the originals, with a 10% or 20% margin.

*Strict time constraint (STC)*: The translated utterances must be exactly the same length as the original ones, ignoring any internal pauses or specific intonation.

*Sound-sync (SS)*: The translated utterances must be exactly the same length as the original ones, including internal pauses and intonation.

*Lip-sync*: The translated text must be exactly the same length as the original, including pauses, and must resemble the lip articulation.

These five synchronies can be associated with game situations (Mejías-Climent 2019: 90). Game situations alternate continuously in any video game (Pujol Tubau 2015: 150). They are a direct consequence of the interactive dimension (interactive channel) and imply different conditions for interaction, depending not only on the genre, but also on the nature of every single video game.

Typically, in action-adventure video games, cinematic scenes stop interaction completely, since they are closed video clips based on a cinematographic configuration. Game action implies full interaction, the dynamic moment in which the player makes the game develop fully. Dialogues are dialectical exchanges and can be considered a situation in between cinematics and action: they can stop interaction partially, restricting the player's action to a few camera movements, for example, or not interfering with interaction at all. Finally, tasks are instructions transmitted to the player that can also take place during full interaction or stopping it completely. In the following section, the relationship between game situations and game synchronies will be described, as well as the methodology used to analyze dubbing as one of the AVT modes included in the process of localizing a graphic adventure.

## 4 Methodology

In a previous project (Mejías-Climent 2019), a relationship between game situations and dubbing synchronies in action-adventure video games was established. The empirical analysis showed that in the three action-adventure video games that were analyzed,

- tasks are always dubbed applying no restriction (wild sync), since they are transmitted through off-screen voices;
- game action is a relatively flexible situation, since full interaction does not always permit the highest level of visibility of the characters onscreen.

Thus, TC usually applies, with a few cases of wild sync for off-screen voices;

- dialogues are a hybrid situation in terms of interaction (they vary greatly from one video game to another). The result is that the five types of synchrony can be found in dialogues, being TC the most frequent;
- cinematics tend to resemble movies as much as possible. This is also evident in the type of synchrony most commonly used in them: lip-sync. In addition, wild sync is always used for off-screen voices.

In this case, a different video game has been analyzed to determine if this relationship between game situations and types of synchrony is also present in a game subgenre closer to traditional movies: the graphic adventure *Detroit: Become Human*. More specifically, the dubbing of the cinematic scenes in this game will be discussed in terms of lip-synching, in order to determine if there is a clear distinction between the dubbing of a movie and the dubbing of the cinematic scenes in a graphic adventure.

To do so, an empirical analysis was carried out within the framework of descriptive translation studies, following the methodology used in Mejías-Climent (2017; 2019). This can be labeled as an exploratory study requiring further research, since, to the best of our knowledge, no previous empirical studies on the particularities of dubbing in video games have been conducted. To reduce the scope of the study, the specific phenomenon of dubbing synchronies was analyzed in each game situation identified in said video game. More specifically, the game was played classifying in an Excel sheet every game situation that alternated throughout the story and annotating the type of synchrony identified in each situation. The game was played first in Spanish until the main goal was achieved. The same path was reproduced again in English, looking for the types of synchrony used in both the translated segments and the original ones (Toury 1995). The game was played for ten hours both in Spanish and in English. A total of 20 hours of gameplay was analyzed.

This methodology reveals some limitations that were already acknowledged in Mejías-Climent (2019: 315–317). First of all, the validity of the game situations taxonomy could be questioned, as dialogues represent a completely heterogeneous situation in terms of interactive options. Nonetheless, while game action, cinematics and tasks have been analyzed, the characteristics of dialogues have been redefined in the particular case of the video game presented here to suit the features of a graphic adventure. Indeed, dialogues have been considered as

*dialogic quick-time events* (see §5), which creates a more adapted taxonomy of game situations to conduct this empirical analysis.

Secondly, an interactive product will always entail a certain level of arbitrariness when displaying its components, which might make slight differences arise if the study was replicated. However, in terms of percentages, the impact of these changes on the results obtained would mean unrepresentative variations, so the general conclusions should not be affected.

Finally, a strong limitation imposed by empirical analyses of video games is the difficult access to the linguistic components (such as written scripts of the plot used for translation purposes), as strong non-disclosure agreements prevent the video game assets from being freely accessible or distributed. To overcome this, the corpus was analyzed while playing, since this is the only way to access the full game, and only some transcriptions of representative examples were made. The gameplay, nonetheless, was recorded and the time codes of the recording were annotated in the Excel sheet to track down all the game situations easily. Closer collaboration between the industry and academia could enrich empirical analysis like this in the future.

## 5 The case of *Detroit: Become Human*

The video game *Detroit: Become Human* was released in 2018 for Play Station 4, and in 2019 for PC. Its director, David Cage, is also the founder of Quantic Dream, a studio specializing in interactive storytelling and in which this game was developed.

The work by David Cage has always created controversy among the most purist players because of the high level of narrative all his games contain, at the expense of a fully interactive and ludic experience. The storytelling seems to be more important than mechanics based on a rapid response and quick reflexes from the (active) player (Altozano Dayo 2017). With his games, Cage exhorts the player to “play the story” combining continuous cinematic scenes with interactive dialogues and quick time events. Nevertheless, despite what some players claim, Cage states that he aims to achieve a full and realistic immersion rather than the narrative display (Altozano Dayo 2017).

Quick-time events (QTEs) are one of the most defining features in David Cage’s work. A QTE represents an action that is completed automatically after pressing a certain button in a limited time. They typically occur during cinematics and, if the specific button is pressed as indicated by the game, the scene continues developing successfully (Altozano Dayo 2017: 131). QTEs are a useful tool to make the

story develop combining action and cinematics (thus, interaction and movies). On the one hand, a QTE is like watching a movie, except for certain buttons that the spectator is required to press if he/she wants the story to continue. On the other hand, the most purist players also see it as an interruption of pure playability (Altozano Dayo 2017: 132).

Be it as it may, the truth is that QTEs are a recurrent tool included by Cage in all of his games to make the story develop making the player participate in an unpredictable but limited way. In this analysis, QTEs are considered as dialogues: they usually introduce a question or an answer that the player has to choose in a limited time as part of a longer dialogue. Pure QTE are not very common in *Detroit: Become Human*, neither are they considered a game situation to be analyzed, because they do not imply any sort of dubbing – in the middle of a fight, the player needs to press  $\times$ ,  $\triangle$  or  $\circ$  to hit the opponent or to cover him/herself, but no linguistic content is related to the QTE in such cases (thus no dubbing sync applies).

This video game, as any other by David Cage and Quantic Dream, is intended to make the player become part of the story through very simple mechanics based on basic movements, constant dialogues and numerous cinematic scenes and dialogical QTEs (as opposed to “action” QTEs, which are not analyzed here, as mentioned). The story revolves around a dystopian Detroit City in which androids start feeling and behaving beyond what machines are expected to do, causing some divergent androids to rebel against humans and fight for their rights. The player controls three characters alternatively. The story is divided into sequences. After each one, a blank tree diagram shows the different possibilities that the player could have considered with his/her choices. The 10 hours played in each language add up a total of 35 sequences.

## 5.1 Game situations in Detroit

As stated in §4, the game was played for 10 hours in each language, obtaining 696 registers distributed in 35 sequences (See Table 8.1). The most frequent game situation, as the numbers show, are cinematic scenes, followed by QTE-dialogues, game action, and finally, tasks. This represents a clear illustration of the nature of the game: a graphic adventure focuses on the narrative content. The story is the core of the game and those game situations that carry the narrative weight are more frequent than fully interactive moments (game action). This contrasts with the number of game situations obtained in the aforementioned studies (Mejías-Climent 2017; 2019), in which action-adventure games were analyzed. In all of them, game action was the most frequent situation throughout the interactive

audiovisual product. We consider it *frequent* instead of *long* because the time span has not been measured. While in a movie time codes and the duration of particular phenomena can be traced, in a video game, interactivity make some situations last as much as the player wishes (game action), while other do have a pre-established length (cinematics). This is why alternation and repetition have been analyzed instead of time span.

Table 8.1: Game situations identified in 10 hours gameplay of *Detroit: Become Human*

Game situation	Num. of registers	
Tasks	58	8.33%
Game action	180	25.86%
QTE-Dialogues	212	30.46%
Cinematic scenes	246	35.34%

Regarding the types of synchrony found in each group of game situations, the following data was obtained in *DTB*:

- Tasks are transmitted through in-game (written) text exclusively. Therefore, their translation is never dubbed but always written onscreen, with a single exception: an introductory sequence in the main menu, during which a woman in a close-up shot talks directly to the player to guide him/her through the main settings of the game, before the actual story begins. This single task is dubbed lip-synching.
- Game action: most of the game action is dubbed using wild sync (42 cases both in Spanish and in English). STC and TC are also frequent (31 and 22 cases respectively for the Spanish version, and 31 and 21 in English). Lip-sync is only used in 2 cases in Spanish, but in 6 cases in English. There are 80 moments in which no dialogues are heard during game action.
- QTE-Dialogues: with a few exceptions, most dialogical QTEs are dubbed using lip-sync, thus resembling the dubbing used in traditional movies. Only 1 example of TC, 4 of STC and 2 examples of SS have been found both in Spanish and English. There are 5 moments using off-voices (wild sync).
- Cinematics: as well as dialogical QTEs, cinematic scenes are dubbed using lip-sync almost exclusively (221 cases in Spanish; 222 in English), with only

1 example of wild, TC, STC and 5 of SS in Spanish; and 1 example of wild, TC, STC and only 4 of SS in English.

These results illustrate the nature of the subgenre of a graphic adventure, at least in the case of *DBH*: dialogues using QTEs and cinematic scenes are the most frequent game situations in a video game aimed at recreating the atmosphere of a movie, but always reminding the player that he/she is the main character in charge of taking all the decisions to make the story develop. Additionally, the use of lip-sync as the most frequent type of synchrony in the dubbing of such a game – 424 cases in Spanish and 429 in English out of 696 registers – shows the closeness of the final configuration of dubbing in a graphic adventure and in a traditional movie, in terms of the types of synchrony applied. As in any non-interactive audiovisual material, wild sync is also frequent—54 registers both in Spanish and English, because off-voices are used repeatedly to make the story develop.

A significant difference with non-interactive movies, however, is the constant use of text written on the screen, especially, to give instructions to the player, replacing dubbing. A total of 148 out of 696 cases, both in Spanish and English, have been registered as containing no dialogues; 80 took place during game action, 11 represented silent cinematic scenes and 58 were tasks in which the instructions were transmitted through in-game text. This in-game instructional text will never be used in a non-interactive movie, as no action is required from the viewer.

## 5.2 The dubbing of cinematic scenes

As described in §5.1, the highest level of restriction, lip-sync, is the most frequently applied in the dubbing of the graphic adventure *DBH*. This is the most complex type of synchrony, since the reproduction of the articulatory movements in the translated text is not always compatible with an accurate translation. In movies, it is reserved almost exclusively for close-ups and extreme close-ups (Chaume 2012).

In the case of *DBH*, the most restrictive type of synchrony is noticeable in almost all the cinematics and the dialogical QTEs. Here, a strong similarity with the dubbing of any non-interactive movie that should be noted is that the characters' lips in *DBH* are animated in English using the articulatory lip movements of the actors dubbing the original version. In a movie, we can see real actors speaking onscreen, whose utterances are replaced later by the translated utterances performed by the dubbing actors in the target language. In a video game, it has

been argued (Méndez González 2015: 106) that the “original version” might not be strictly the first one, since all video game characters need a human actor/actress to provide them with his/her voice. Nonetheless, in AAA modern video games such as *DBH*, characters’ movements and even speech articulation are recreated with the motion capture technique (Turnes 2020), using sensors on a human actor’s body to capture his/her movements and recreate them in a digital model to animate a virtual character (Kines 2000). In such a way, the characters’ lips in *DBH* accurately reproduce the articulatory movements of the words they utter in English, as real actors in a non-interactive movie do. In this case, the slight differences between the exact lip movements of the original version in English and the lip-synched Spanish sentences are visible in *DBH* as it happens in any movie.

Regarding the level of accuracy in lip-synching, it should be noted that, as explained above, in most cases dubbing actors do not have images to support their performances. This might result in a slightly less accurate lip-synching than in a traditional movie that, nonetheless, remains unnoticed by most spectators, although further reception studies in the field of video game dubbing would be quite revealing. In the case of *DBH*, lip-synchrony appears quite accurate, even though some open vowels, and bilabial and labio-dental consonants do not coincide exactly in some cases.

To determine how accurate lip-sync is achieved in *DTB*, the following examples reproduce the dubbed and original dialogues in a cinematic scene, a dialogical QTE and a second cinematic scene dubbed using lip-sync. Those paralinguistic features included in the acting (coughing, onomatopoeic expressions, etc.) have been reproduced with the dubbing symbol (G) used in Spanish dubbing scripts. Pauses are represented with “/” and whenever a non-dialogical QTE is required to make the action continue, the symbol (QTE) has been introduced.

(1) Cinematic scene 298

Connor:	¿Teniente? / [QTE]	Lieutenant? / [QTE]
Anderson:	(G)	(G)
Connor:	Despierte, teniente.	Wake up, lieutenant!
Anderson:	(G)	(G)
	[QTE]	[QTE]
Connor:	¡Soy yo, Connor!	It’s me, Connor!
Anderson:	(G)	(G)
Connor:	Por su bien, haré que vuelva a estar sobrio. Se lo advierto: será poco agradable.	I’m going to sober you up, for your own safety. I have to warn you, this may be unpleasant.

- Anderson: (A LA VEZ) (G) ¡Déjame en paz, puto **androide**! ¡Sal de mi casa, joder!
- (SIMULTANEOUSLY) Hey! Leave **me** alone, you **fuckin' android**! Get the fuck outta **my house**!
- Connor: Lo siento, **teniente**, pero lo **necesito**. [QTE] Le agradezco de **antemano** su **colaboración**.
- I'm sorry, **lieutenant**, but I **need you**. [QTE] Thank **you** in **advance** for your **cooperation**.
- Anderson: (P) Eh, ¡lárgate de aquí, joder!
- (P) Hey! Get the fuck outta **here**! [QTE]

(2) QTE-Dialogue 302

- Connor: Lo **entiendo**... **Tampoco** creo que tuviera mucho **interés**... **Han encontrado** el **cadáver** de un hombre en un **burdel del centro**... Ya **resolverán** el **caso** sin nosotros...
- I **understand**... It **probably** wasn't **interesting** **anyway**... **A man** found dead in a **sex club downtown**... **Guess** they'll **have to solve** the **case** without us...
- Anderson: **Oye**, no **me vendrá mal** tomar un **poco** el **aire**. **En el armario** de la **habitación** hay **ropa**.
- You know, **probably** wouldn't do **me** any **harm** to get **some air**... There're **some clothes** in the **bedroom** there.
- Connor: Iré **a por** ella.
- I'll go get them.

(3) Cinematic scene 305

- Connor: ¿**Qué se quiere poner**?
- What do you **want to wear**?
- Anderson: (G) (Tose)
- (G) (Coughing)
- Connor: ¿**Se encuentra bien, teniente**?
- Are you **alright, lieutenant**?
- Anderson: (G) Sí... sí... (G) De maravilla... **solo**... dame cinco **minutos**, ¿vale?
- (G) Yeah... yeah... (G) Wonderful... **Just a...** Give me five **minutes**, okay?
- Connor: Claro.
- Sure.
- Anderson: (G)
- (G)

In the three examples above, all the elements included in video game lip-synching are reproduced precisely, as in the case of a non-interactive movie: the

length of the translated utterances is the same as the original ones; intonation and paralinguistic features are reproduced exactly too. Also open vowels (*a*, *e*, *o*, marked in red) and most labial and bi-labial consonants (marked in blue) are reproduced as much as possible. This makes this level of restriction quite close to how it is done in a non-interactive movie, even though translators and most probably dubbing actors did not have access to the final videos. Most times, videos have not been produced yet when dubbing takes place, although this could not be confirmed in the particular case of *DBH*.

The use of the highest level of restriction in the translated text of a video game implies that all previous levels have been taken into account as well: length of the utterances (TC, STC) and intonation and pauses within utterances (SS). Thus, isochrony is applied, as it is always the case in a non-interactive movie, being this the most valued quality standard among Spanish spectators (Chaume 2007).

Some cases of cinematographic kinetic synchrony are also found in the dubbing of *DBH*. This is most probably not done intentionally in the dubbing studio, since videos are not available. Some deictic expressions have been either omitted, as in (2), or reproduced literally, in the same position within the sentence, to assure the correspondence with the image.

In (2), *There're some clothes in the bedroom there* has been translated as '*En el armario de la habitación hay ropa*'. The second occurrence of *there* has been omitted, but the reference to the bedroom should be enough to make the possible body movements of the character coherent with a reference to it, wherever it is located in the house.

Another example was found in the game situation 23 (see example (4)), a limited game action during which the player's actions were restricted to camera movements. The seller refers to the android in front of him with the following utterance:

(4) Game action 23

Seller:	En este momento tenemos una oferta especial para <b>esta</b> gama de 7999 dólares, con financiación sin intereses a 48 meses.	At the moment we're doing a special promotion on <b>this</b> entire range at \$7999, with a 48-months interest free credit.
---------	---	---

The deictic *this* coincides with the character onscreen pointing at the android in front of him. Since it has been translated reproducing the same length and structure of the sentence, kinetic synchrony has been applied here as well.

These are just a few examples that illustrate the accuracy with which lip-sync has been applied in the analyzed video game, similar to what Chaume (2016)

did in a descriptive and “qualitative analysis according to the episode’s adherence to a checklist of dubbing standards” in the TV series *The West Wing* (1999) (Chaume 2016). His work, focusing on quality standards in dubbing (Chaume 2007), as well as some other studies on the language of dubbing such as those by Piñero (2009); Ferriol (2013); Freddi & Pavesi (2009) and Sierra (2008), among many others, could serve as useful models for further research in the field of video game dubbing, compared to dubbing in non-interactive material.

Empirical and reception studies would especially benefit greatly from this pilot analysis on the characteristics of dubbing synchronies in video games, which represents a revealing starting point to trace considerable similarities in the dubbing of interactive and non-interactive material, in spite of the different materials available when translating, and the differences in the translation and localization processes.

## 6 Final remarks

The main aim of this chapter was to review the most prominent similarities and differences between cinematographic dubbing and that of a video game belonging to the subgenre of graphic adventures, within the main genre of adventure video games.

Video games represent the most complex example of a multimodal text and, as such, they share many of the characteristics of a non-interactive audiovisual product. The particularity of the added interactive dimension in video games makes them a specific case, which undergoes a process of localization when exported to other cultures, beyond a mere linguistic transfer.

Within the industry, professionals support the idea of localization being an independent field, different from AVT, since it implies some other adaptation processes including a more creative approach for translators and the modification of linguistic and non-linguistic contents of the localized product. Within academia, however, there is no clear evidence that justifies the need to separate localization from AVT. Both fields deal with the translation of multimedia and multimodal products, and both encompass a series of translation modes such as dubbing or subtitling, among many others, required for the different types of audiovisual products and their particularities.

Further research is needed on the convergences and differences of the process of dubbing a movie and a video game, and on localization and AVT modes in general. Notwithstanding this, the aim of these pages was to point out that the result in the dubbing of a graphic adventure and a traditional movie is very similar in terms of dubbing synchronies.

Three types of synchronies have been described in the dubbing of non-interactive audiovisual products (Chaume 2004; 2007; 2012). A new taxonomy of five dubbing synchronies is also necessary in video games (Mejías-Climent 2017; 2019), given the idiosyncrasy of the interactive audiovisual text. However, in the specific case of cinematic scenes and dialogical QTEs of a graphic adventure such as *Detroit: Become Human*, the most restrictive dubbing synchrony in video games, lip-sync, is the one used most frequently and encompasses all the three dubbing synchronies described for non-interactive products.

It should be acknowledged that a case study will never be enough to identify translation tendencies (Toury 1995), neither to close the debate of the link between AVT and localization. Nonetheless, this study can be understood as no more than a starting point after which further research is needed, first, enlarging the corpus of adventure video games. Other game genres would need to be explored as well, to look for translation tendencies in the use of dubbing synchronies, and to determine the most characteristic game situations for each genre. Audiovisual translation and localization are two convergent fields and professional practices, which seem to share more similarities than differences, although much is yet to be done with further research.

## Abbreviations

AVT	audiovisual translation	TC	Time constraint
DBH	<i>Detroit: Become Human</i>	STC	Strict time constraint
QTE	Quick-Time Event	SS	Sound-sync

## References

- Altozano Dayo, José. 2017. *El videojuego a través de David cage*. Sevilla: Héroes de Papel.
- Bernal-Merino, Miguel Ángel. 2006. On the translation of video games. *JoSTrans: The Journal of Specialised Translation* 6. 22–36.
- Bernal-Merino, Miguel Ángel. 2015. *Translation and localization in video games: Making entertainment software global*. London: Routledge.
- Cadieux, Pierre & Bert Esselink. 2004. GILT: Globalization, internationalization, localization, translation. *Globalization Insider* 11(1.5). 1–5.
- Chaume, Frederic. 2004. *Cine y traducción*. Madrid: Cátedra.
- Chaume, Frederic. 2007. Quality standards in dubbing: A proposal. *Tradterm* 13. 71–89.

- Chaume, Frederic. 2012. La traducción audiovisual: Nuevas tecnologías, nuevas audiencias. In Cristina Bosisio & Stefania Cavagnoli (eds.), *12° congresso dell'associazione Italiana di linguistica applicata*, 143–159. Perugia: Guerra Edizioni.
- Chaume, Frederic. 2013. The turn of audiovisual translation: New audiences and new technologies. *Translation Spaces* 2. 107–125.
- Chaume, Frederic. 2016. Dubbing a TV drama series: The case of *The West Wing*. *inTRAlinea* 18. 1–12.
- Chaume, Frederic. 2018a. An overview of audiovisual translation: Four methodological turns in a mature discipline. *Journal of Audiovisual Translation* 1(1). 40–63.
- Chaume, Frederic. 2018b. Is audiovisual translation putting the concept of translation up against the ropes? *JoSTrans: The Journal of Specialised Translation* 30. 84–104.
- Esselink, Bert. 2000. *A practical guide to localization*. Amsterdam: John Benjamins.
- Ferriol, José Luis Martí. 2013. *El método de traducción*. Castelló de la Plana: Universitat Jaume I.
- Freddi, Maria & Maria Pavesi. 2009. *Analysing audiovisual dialogue: Linguistic and translational insights*. Bologna: Clueb.
- Gambier, Yves. 2003. Screen transadaptation: Perception and reception. *The Translator* 9(2). 171–189.
- Granell, Ximo, Carme Mangiron & Núria Vidal. 2015. *La traducción de videojuegos*. Sevilla: Bienza.
- Hurtado Albir, Amparo. 2011. *Traducción y traductología: Introducción a la traductología*. Madrid: Cátedra.
- Jiménez-Crespo, Miguel. 2013. *Translation and web localization*. London: Routledge.
- Kines, Melianthe. 2000. *Planning and directing motion capture for games*. [https://www.gamasutra.com/view/feature/131827/planning\\_and\\_directing\\_motion\\_.php](https://www.gamasutra.com/view/feature/131827/planning_and_directing_motion_.php) (6 July, 2020).
- Mangiron, Carme & Minako O'Hagan. 2006. Game localization: Unleashing imagination with “restricted” translation. *JoSTrans: The Journal of Specialised Translation* 6. 10–21.
- Mata Pastor, Manuel. 2005. Localización y traducción de contenido web. In Detlef Reineke (ed.), *Traducción y localización: Mercado, gestión y tecnologías*, 187–252. Las Palmas de Gran Canaria: Anroart.
- Maxwell-Chandler, Heather. 2005. *The game localization handbook*. Boston, Massachusetts: Charles River Media.

- Mejías-Climent, Laura. 2017. Multimodality and dubbing in video games: A research approach. *Linguistica Antverpiensia, New Series: Themes in Translation Studies* 17. 99–113.
- Mejías-Climent, Laura. 2019. *La sincronización en el doblaje de videojuegos: Análisis empírico y descriptivo de los videojuegos de acción-aventura*. Universitat Jaume I. (Doctoral dissertation).
- Méndez González, Ramón. 2015. *Localización de videojuegos: Fundamentos traductológicos innovadores para nuevas prácticas profesionales*. Vigo: Servizo de Publicacións Universidade de Vigo.
- Muñoz Sánchez, Pablo. 2017. *Localización de videojuegos*. Madrid: Síntesis.
- Neves, Joselia. 2005. *Audiovisual translation: Subtitling for the deaf and hard-of-hearing*. University of Roehampton. (Doctoral dissertation).
- O'Hagan, Minako & Carme Mangiron. 2013. *Game localization: Translating for the global digital entertainment industry*. Amsterdam: John Benjamins.
- Piñero, Rocío Baños. 2009. *La oralidad prefabricada en la traducción para el doblaje: Estudio descriptivo-contrastivo del español de dos comedias de situación: Siete vidas y Friends*. Universidad de Granada. (Doctoral dissertation).
- Pruys, Guido Marc. 2009. *Die Rhetorik der Filmsynchronisation*. Cologne: Self-published.
- Pujol Tubau, Miquel. 2015. *La representació de personatges a través del doblatge en narratives transmèdia: Estudi descriptiu de pel·lícules i videojocs basats en El senyor dels anells*. Universitat de Vic. (Doctoral dissertation).
- Pym, Anthony. 2016. *Teorías contemporáneas de la traducción: Materiales para un curso universitario*. 2nd edn. Tarragona: Intercultural Studies Group.
- Sierra, Juan José Martínez. 2008. *Humor y traducción: Los simpson cruzan la frontera*. Castelló de la Plana: Universitat Jaume I.
- Toury, Gideon. 1995. *Descriptive translation studies and beyond*. Amsterdam: John Benjamins.
- Turnes, Yova. 2020. *GamerDic*. <http://www.gamerdic.es/web/sobre-la-web/> (6 July, 2020).
- Vázquez Rodríguez, Arturo. 2018. *El error de traducción en la localización de videojuegos: Estudio descriptivo y comparativo entre videojuegos indie y no indie*. Universitat de València. (Doctoral dissertation).

