

Multimodality and dubbing in video games: A research approach

Laura Mejías-Climent

Universitat Jaume I, Spain
lmejias@uji.es

This article discusses a new taxonomy of dubbing synchronies in the Spanish localized version of the video game Batman: Arkham Knight. To do so, the concept of the video game is first reviewed as the most sophisticated example of audiovisual text in terms of its multimodal nature. The article also describes how the message is conveyed through the different communication channels that form a part of its semiotic structure. Next, the research approach proposal is presented, based on the multimodal configuration of video games: game situations will be the unit of analysis used in this approach, since they are a basic unit with which to organize the content of a game and they distinguish video games as a unique type of audiovisual text. These game situations are used to analyse the specific features of dubbing and types of synchrony. Finally, some results are offered to confirm that the taxonomy of dubbing in video games discussed here applies to this particular case.

1. Video games as a modern object of study

Since video games represent one of the richest examples of multimodal texts, given their complex semiotic construction (Maietti, 2004), they have been taken as the object of study of this article. The aim of this study is to explore a new taxonomy of dubbing synchronies related to the multimodal nature of video games. In order to do so, a research approach will be proposed within the field of Audiovisual Translation (AVT) that takes into account the semiotic construction of video games. The focus is therefore on just one of the phases through which video games pass during the localization process: the dubbing phase. The article explores an audiovisual *translation mode* (Chaume, 2012, p. 1) that has already been researched extensively (see references below), although in this case from the unique perspective of video games.

The origins of video games date back no further than to the early 1960s (Juul, 2005; Levis, 1997; López Redondo, 2014). However, video games have developed rapidly, especially since the 1990s. Nowadays they represent the leading entertainment industry worldwide (Bernal Merino, 2015, pp. 9–11; Granell, Mangiron, & Vidal, 2015, p. 31), having generated an income close to US\$100 billion in 2016 (AEVI, 2017; Wijman, 2017). In the United States, the “total video game software sales exceeded US\$24.5 billion in 2016” (Siwek, 2017). In Spain, the figures are equally significant: in 2016, around 15 million game players were estimated to generate an income of €1.2 billion for the video-game industry. This far exceeds the 2016 revenues produced by the cinematographic industry in the same country, at almost €602 million (Belinchón, 2017), and by the music industry, which brought in approximately €164 million in 2016 (ICEX, 2017).

As the numbers show, video games are an immensely popular entertainment product. Although research on video game localization has increased in the past decade (Mangiron, 2017, p. 74), it is a relatively recent area of study that can still be explored further in academic research (Bernal Merino, 2015, p. 4; Mangiron, O’Hagan, & Orero, 2014, p. 10; O’Hagan & Chandler, 2016, p. 309). Research is needed in particular to explore specific practices in the localization process, such as dubbing, which has been

covered only partially by some authors (Le Dour, 2007; Méndez González, 2015; Sioli, Minazzi, & Ballista, 2007).

Video games are a highly complex—and interactive—audiovisual medium. Their adaptation to other cultures therefore suggests a localization process beyond mere linguistic translation, one that includes adapting the product's cultural, legal, visual and functional aspects (Bernal Merino, 2015, p. 106; Granell et al., 2015, pp. 19–27). In this process, AVT is just one of the phases. It shares remarkable similarities with the AVT in films and TV productions, but there are also some important differences, such as the unavailability of the videos when dubbing or the variability of the original text. (Translators work with relatively unstable text: the batches they receive may not be final because the localization process usually takes place simultaneously with the development of the video game. Consequently, the translated text may have to undergo continuous changes.¹)

To begin to identify the features of the specific phase of dubbing in the localization process, a definition will first be provided of what video games are and how their users interact with them. Next, a proposal will be outlined for a descriptive analysis based on the different *game situations*, which represent the unit of analysis. This concept is introduced and explained below (see section 2.2). To narrow the analysis, the proposal focuses on synchrony, offering a different taxonomy of dubbing synchronies: while only two channels are taken into consideration when creating the translated text in a film, a third, interactive dimension (i.e., an additional semiotic channel) intervenes when a video game is configured. This gives a different structure to the development of on-screen events (compared to non-interactive products), which do not follow a single, linear path. In the case of video games, the contents (ludic, narrative and thus linguistic) are distributed in different *game situations*. These require specific forms of synchronization, which is the focus of this study (see section 3.2).

2. Multimodal structure of interactive audiovisual texts in video games

The combination of computer and television media is the starting point for the object of this study. The complexity of this combination can also be analysed from a communication and an entertainment perspective:

Video games comprise interactive elements in text and graphics as well as multimodal user input mechanisms, and in some cases also include non-interactive audiovisual elements in the form of pre-rendered movies (O'Hagan, 2007, p. 2).

It is also pertinent to refer to the *audiovisual* and *interactive* aspects that characterize video games, as they are fundamental references when differentiating video games from other entertainment or professional products (Pérez Latorre, 2010, p. 35). Furthermore, video games are different from other types of audiovisual experience, such as watching a movie, because of the determining influence of the *user's participation* in the development of the story.

Video games can therefore be classified as the most complex example of audiovisual text, given their multimodal and interactive features (Bernal Merino, 2015; Levis, 1997). In fact, video games fit the definition of a traditional *audiovisual text* given by Chaume (2012): “An audiovisual text is a semiotic construct woven by a series of signifying codes that operate simultaneously to produce meaning” (p. 100). This semiotic construct is based on the different codes that are transmitted through acoustic and visual channels (Chaume, 2012, p. 172).

Although these two channels of communication have been defined for traditional audiovisual texts and also apply to video games, the latter feature up to three different channels, a characteristic that differentiates them from other traditional audiovisual texts

such as movies and TV series. Besides the acoustic and visual channels already mentioned, a third *tactile channel* operates in video games (Pujol Tubau, 2015, p. 134), which makes them a unique audiovisual text. Only a few attempts have been made to include more channels in other multimodal texts, such as *Earthquake* (1974), a film in which the audience could feel the ground trembling; the smell-o-vision films of the 1950s (Chaume, 2003, p. 88); or the most recent film, *XMILE* (2016), an accessible production by M. A. Font in which the sense of smell was also included. But these are exceptional examples in which other channels of communication have been added to the visual and acoustic.

In the case of video games, a third interactive component is always involved: interactivity is most frequently transmitted by the *tactile channel* referred to above in the form of a controller the user touches (pressing buttons or using a joystick). However, there are also examples in which a kinetic sensor is included in the hardware and interaction takes place as a result through the visual channel, replacing the tactile channel (Kinect for Xbox 360, 2010; Nintendo Switch, 2017). This is the case with video games from the seventh generation onwards, although this article will refer to video games in which interaction takes place only through a controller, that is, the tactile channel. Adding this interactive process results in a unique example of audiovisual text in which different modes—verbal language, images, sounds and touch—coexist to create meaning (Kokkola & Ketola, 2015).

Regardless of the type of hardware, interactivity is always a basic component of video games. The characteristic of this semiotic configuration, unlike that in traditional films, is that the transmission of meaning takes place bi-directionally (through interaction): (1) the player receives messages from the acoustic and visual channels (but also from the tactile one, when the controller vibrates); at the same time, (2) the player sends messages back to the game system in the form of either normative codes (each button on the controller represents a sign for the game) or normative and kinetic codes if there's no controller (each movement of the player's body represents a sign signifying something for the game, which is received by kinetic sensors).

López Redondo (2014, p. 32) insists on this distinctive feature (interactivity) and considers it essential to the success of this modern form of entertainment: video games are the first medium in which multiple stimuli and visual dynamism are combined with the user's active participation. This makes them the most complex example of multimodal texts in which a multiplicity of semiotic resources combine to create meaning (Jewitt, 2014, p. 17).

Maietti (2004, p. 117) also describes the configuration of an interactive audiovisual text as being based on the player's enjoyment. The player is presented with a series of possibilities through the acoustic and visual channels, but these are regulated by the normative codes and transmitted back through the tactile channel. Consequently, the player becomes the main character in the story and has the means to control it; it is a bi-directional creation of meaning.

Together with the visual and acoustic codes transmitted through their respective channels (Chaume, 2004, 2012), a series of rules within the video game intertwine with the rest of the codes and establish the different choices or options that a player has when acting in the virtual world (Maietti, 2004, p. 119). These codes constitute the *normative code* transmitted through the tactile channel, or through the visual one if no controller is involved. Each movement produced by the player (pressing buttons or moving the levers) is codified as a sign that represents a specific action within the game world. The next section describes how this normative code works.

2.1 Rules in the multimodal configuration of a video game

As Pérez Latorre (2012, p. 1906) states, rules are the emblematic sign of video games design. He distinguishes two sets of rules: those related to performance competences (rules involving the player) and those related to the behaviour of the non-playable characters (NPCs) and the dynamics of the phenomena that occur within the world of the game (the rules of the environment of the video game).

Salen and Zimmerman (2005, p. 65) also understand rules to be the basic signals that structure the message in a video game: “rules and play create carefully orchestrated instances of designed interaction” (Salen & Zimmerman, 2005, p. 66). Along these lines, Juul (2005, p. 58) describes rules as signs that channel interaction between the game and the user: they impose limitations and offer choices for the players that create meaning. Juul also distinguishes between the rules that govern players’ actions and those that determine how the game reacts.

In this semiotic system, all the components are interconnected and interrelated to create a common multimodal meaning. As happens with the acoustic and visual codes creating the semantic structure in a traditional audiovisual text (Chaume, 2004, p. 19), in video games the structure is based in part on a series of internal codes (rules) that govern the reactions and interactions between the NPCs and the environment and in part also on a series of external rules that provide interaction with the player (press X to jump, press B to hit your opponent, etc.). In other words, the code transmitted by the tactile channel includes both an *external normative code* (referring to the player) and an *internal normative code* (referring to the internal environment of the video game).

It could be said that only two channels and a structure of rules construct the meaning in the video game. Nonetheless, those rules act as a semiotic resource (Kress, 2014, p. 65) channelled by the tactile and kinetic activity (Poyatos, 2012) that involves the player in creating meaning. The rules are transmitted and codified as a system of signs (the external normative code) that the player has to follow; each of the player’s movements (using the controller or their own body kinetically) represents something, a sign that implies a specific reaction within the game world (internal normative code). In *Batman: Arkham Knight*, for instance, pressing the Δ or A button on the controller makes the player counter-attack the enemy (external normative code). Within the game world, the action “counter-attacking” means avoiding injury to the main character if the action is performed at the exact moment the enemy is attacking² (internal normative code). Therefore, the rules represent an additional system of signs transmitted through the aforementioned channels that regulate and create the meaning in such an interactive audiovisual text.

2.2 Game situations as units of analysis in an interactive audiovisual text

Given the interactive features of video games, it is natural that, when one is playing, the events do not develop linearly but rather as a series of situations that alternate throughout the game. Unlike films or TV shows, video games imply different conditions for interactivity: these *game situations* (Pujol Tubau, 2015, p. 150) can be understood as the *different forms of interaction* (or absence of it) in a video game. The interactive audiovisual text is shown on the interface according to the player’s actions; the player becomes an active part of the narrative discourse through the ludic action (following the normative codes transmitted through the tactile channel). There are four types of game situation:



- *Tasks*: acoustic or written instructions for the player. They can be transmitted to stop the interaction completely or not, depending on the type of video game.
- *Game action*: a dynamic moment, the player has full control of the activity (full interaction).

- *Dialogues*: dialectical exchanges among characters. They can take place to restrict the action partially (e.g., limiting the action options to just a few camera changes or to walking) or not to restrict it at all (partial or full interaction).
- *Cinematics*: video clips based on cinematographic resources. They stop the player’s action completely.

Game situations are a direct result of the presence of the normative code (which regulates interactivity) in a video game. When participating, the player creates a continuous alternation of situations regulated by the external system of signs (external normative code) and also by the internal normative code: the game responds to the player’s action (interaction), alternating these situations according to the rules. The tactile channel therefore adds this new system of signs which represent specific actions and reactions within the game world that regulate the alternation of game situations. For example, if the player defeats an enemy during a certain game action, the game action stops and a cinematic is activated. In this way, apart from the information transmitted through the acoustic and visual channels, the tactile channel organizes the content of the video game in different units or situations because the player touches the mouse or the keyboard and this action is reflected on-screen, making the events develop.

The resulting intangible structure of game situations can, in turn, be used as the basis of a descriptive analysis. This article proposes using game situations as translation units to analyse not only the linguistic dimension of the text, but also the whole unit in all its multimodal dimensions (Kaindl, 2013, p. 266).

Table 1: Semiotic structures of audiovisual texts and interactive audiovisual texts, and possible units of analysis in DTS

Semiotic structure of multimodal texts		Analysis
- Visual codes (visual channel) - Acoustic codes (acoustic channel)	 AV text (films, TV shows, etc.)	Syntactic code or montage (integration of the visual and acoustic channels) - Linear development (shots, scenes) - Fixed time codes (TCR)
- Visual codes (visual channel) - Acoustic codes (acoustic channel) - Normative codes (tactile channel)	 Interactive AV text (video games)	Syntactic code or montage (integration of the visual and acoustic channels) + tactile channel - Alternation of game situations (active player participation): - Non-linearity - Variable time codes

Game situations can be used as the *unit of analysis* in research. In some recent case studies (such as Van Oers & Müller Galhardi, in Mangiron et al., 2014, and Vázquez Rodríguez’s 2018 dissertation), descriptive methodology was based on selecting source and translated segments from the video games analysed. However, to date, no other research proposals have been found that take into account the non-linear and multimodal nature of interactive audiovisual texts, with the exception of Pujol Tubau (2015, pp. 150, 211), who also used game situations in his transmedia analysis, although with a completely different objective.

In order to carry out an empirical analysis such as that proposed below, the different game situations will be useful in classifying the translated material included in the game. With this classification, different aspects of the translatable assets in the video game can

be analysed, such as subtitling parameters or, in this case, the types of synchrony used in dubbing. These have been chosen as the feature to be analysed because of the lack of in-depth studies found thus far dealing with this specific topic that combine localization and audiovisual text (AVT) perspectives. Any other studies on different aspects such as subtitling or accessibility will be very helpful in confirming the usefulness of this analytical model.

Having described this unit of analysis that derives from the multimodal configuration of video games, my intention now is to start mapping out the characteristics of the different forms of audiovisual translation carried out within the broader video-game localization process. In order to do so, my analysis focuses on the audiovisual translation mode of dubbing. More specifically, I focus on one typical feature of dubbing: synchronization. This narrowing down responds to the challenges of space and feasibility.

3. Case of dubbing

3.1 Perspective of traditional dubbing

Turning to the specific audiovisual practice to be analysed, dubbing must meet certain quality standards in order to satisfy the audience's expectations (Chaume, 2005). One of these standards is synchronization, which includes matching characters' articulatory movements (*lip-sync*), body movements (*kinesic synchrony*) and the length of their utterances (*isochrony*) (Chaume, 2012, p. 68). This quality standard is one of the many aspects that can be studied and compared between cinematographic dubbing and dubbing in video games.

As has been described for traditional audiovisual texts (films, TV series, etc.), "Synchronization is the process of matching a target language translation to the screen actors' body and articulatory movement in a recording made in a dubbing studio" (Chaume, 2012, p. 67). The different semiotic codes shape the audiovisual text in such a way that it might require a higher or lower level of synchronization. For example, depending on how closely a character is seen or how bright the lights are in a shot, synchronization in dubbing will be applied either more strictly or more loosely.

In cinema and TV, the types of shot, photography, mobility code, paralinguistic code and off- and on-screen sound are especially important (Chaume, 2004, pp. 163-165; Pujol Tubau, 2015, pp. 189-193), since all of these have a direct impact on the linguistic code. Among them, the types of shot have a direct effect on the inclusion of the three types of synchronization mentioned above (see Chaume, 2004, p. 256). However, taking into consideration the multimodal nature of video games, the more complex localization process (including translating and dubbing the text) and the additional normative codes, it is necessary to adapt the traditional audiovisual text model and define a different taxonomy. Consequently, game situations have been used (see section 2.2) with the aim of structuring the dubbed contents of the interactive corpus (instead of using TCR) and defining the different levels of restriction that require other forms of synchronization.

3.2 Different taxonomy of synchronization in video games

In the case of dubbing video games, the synchronization is not done in the same way as in a film (Pujol Tubau, 2015, p. 197); it is instead understood as "a gradation of constraints" that the translator has to take into account. It could be argued that the same semiotic codes described for films (especially the types of shot) determine the types of synchrony applied when dubbing a video game. Nonetheless, professionals involved in the localization process note that videos are rarely available when translating the text, not even in the dubbing studio, where they instead work with the audio files recorded in the

original language and use the audio waves as their main reference (LaPS4, 2016; Xbox España, 2015). This renders it impossible for the types of shot to be the determining factor in synchronizing the dubbed audio. Restrictions are therefore applied depending on the game situation in which the translated audio will be reproduced. Sometimes these restrictions are transmitted to the translator simply as a certain number of characters (not exceeding the original number of characters by more than 10%), but translators and dubbing professionals tend to develop a certain intuition about how restrictive the text should be depending on the game situation in which those lines will be reproduced.

From a multimodal perspective, in each game situation (tasks, game action, dialogues and cinematics) the gradation of constraints mentioned above actually depends on the number of semiotic codes that have a direct impact on the linguistic one (see Table 2). In action-adventure video games, tasks are usually transmitted by voice-over, so there will be no specific restrictions on dubbing them.

Game *action* displays specific gameplay in which the players are performing and moving continuously, so restrictions will not be very stringent, mainly depending on where the voices come from (on-screen characters, thoughts, walkie-talkies) or the photography of the scene (perspective and lighting). The normative code also plays an important role in the case of game action: depending on the mechanics of the game and the perspectives and movements available to the player, synchronization will need to be more or less carefully applied (e.g., a game that allows only a third-person perspective will always show other characters' faces with a lower level of detail than a game in which other characters can be seen more closely).

The normative code and the possible camera and character movements for the player are also important in *dialogues*. They always differ depending on the game but, generally speaking, a possible closer perspective of the characters involved in the dialogue and their expressiveness are taken into account (mobility and paralinguistic).

Finally, *cinematics* tend to imitate the dubbing in movies. Although no videos are usually available, they tend to be dubbed so as to imitate the original sound wave as much as possible, since the more the two audio waves resemble each other, the more probable it is that all the semiotic codes will fit together when the final dubbed version is completed.

The same is true of the other three game situations: the more semiotic codes there are that are expected to come into play when the final version is ready, the stricter the type of synchrony that will be applied. This is related to another aspect that differentiates dubbing in films from that in video games: whereas in the former the dubbing studio produces the final version of the dubbed product, in the latter, neither the translators nor the dubbing professionals see the final results until the developer has received all the translated assets and integrated them to produce the final version of the game—which can vary greatly from the expectations of the professionals involved in the localization process.

To sum up, depending on the gradation of constraints expected in each situation and the number of semiotic codes involved, up to five types of synchrony³ can be found in the dubbing of a video game:

- *Wild (VO)*: 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 (not taking into account 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.

The manner in which these five types of synchrony are applied depends on how restrictive each game situation is. Restrictions can be associated with the different semiotic codes, as explained above. These ideas are also illustrated in the table below (adapted from Pujol Tubau, 2015, p. 198). The impact of the semiotic codes creating the different levels of constraints when translating and dubbing allows a comparison to be established between the taxonomy of synchronization in video games and the three synchronies already seen in films and TV: the more restrictive the game situation is expected to be, the more the dubbing will resemble the types of synchrony used in films.

Table 2: Possible types of synchronization in video games compared to traditional audiovisual texts, depending on the game situations and their relationship with semiotic codes

<i>Game situation (alternated by the player's actions)</i>	<i>Type of synchrony</i>	<i>Acoustic, visual and tactile channels</i>	<i>Corresponding cinematographic type of synchrony</i>
		<i>Semiotic code that determines the level of restriction</i>	
Tasks	▪ VO	▪ Normative	▪ None
Game action	▪ TC	▪ Normative ▪ Off- and on-screen sound ▪ Photography	▪ Isochrony +/-10%
Dialogues	▪ STC ▪ SS	▪ Normative ▪ Off- and on-screen sound ▪ Photography ▪ Mobility codes ▪ Paralinguistic codes	▪ Isochrony ▪ Isochrony and kinesic synchrony
Cinematics	▪ Lip-sync	▪ Normative (only triggering cinematics) ▪ Off- and on-screen sound ▪ Photography ▪ Mobility codes ▪ Paralinguistic codes ▪ Types of shot	▪ Isochrony, kinesic synchrony and lip-sync

Source: Adapted from Pujol Tubau, 2015

The different levels of constraint mentioned by Pujol Tubau (2015) are a direct consequence of the higher number of semiotic codes that affect and influence the linguistic one: the more restrictive the situation is, the more the codes that apply. The normative code applies to all situations, since it regulates their alternation (internal normative code), depending on the signs (movements) that the player makes (external normative code). Because of the semiotic characteristics of the product and the particular work process when localizing it, it would not make sense to determine the types of synchrony based solely on the traditional audio and visual (cinematographic) codes. This is so because the video game is not a final product responding to a clear cinematographic configuration until the end of the localization process, when all the translated assets are finally integrated. As previously mentioned, videos are rarely made available to translators and dubbing professionals, so their tasks are carried out based on the characteristics of each game situation rather than on the final visual configuration of the product.

It is also interesting to compare the different types of synchrony described in video games with those applied in cinematographic dubbing: the semiotic codes that come into play (acoustic and visual) are the same in both cases, although synchronization in video games includes more options (up to five types). This is because not only the acoustic and

visual codes are active; there is also a continuous alternation of game situations caused by the normative codes (transmitted by the intervention of the tactile channel, see section 2.1). The structure of the audiovisual text in different game situations is a direct consequence of player participation through interaction.

In Table 2, therefore, depicting the types of synchrony expected in a descriptive study, game situations are the basic unit of analysis, since they imply different levels of restriction and derive from the various forms of interaction that the player can have with the video game.

4. Analysis and results

4.1 Approaching the method

The empirical analysis proposed is based on the action-adventure video game *Batman: Arkham-Knight (BAK)* (Rocksteady Studios, 2015). The different types of synchrony employed in its dubbing are analysed, using the four game situations as the basic unit of analysis (tasks, game action, dialogues and cinematics). As previously explained, in this empirical analysis, the aim is to reflect how the multimodal configuration of a video game motivates a different approach to translation. In this case, having described how the transmission of the message takes place between the player and the video game (see section 2), game situations become the unit of analysis with which to structure a video game's translatable assets, as their particular configuration does not allow an analysis to be done based on TCR.

This game was played first in its dubbed version (Spanish) and later in the source language (SL) (English). The same path was covered in each case, following Toury's 1995 model of analysis, in which the translated segments (in this particular case, game situations) are analysed and then compared with the originals. It must be noted that the comparison was not made between the linguistic translation but between the types of synchrony applied in both the original and the translated segments. The game was played only until the main goal was accomplished, ignoring any secondary missions, in order to limit the time spent playing. While a movie usually lasts around 90 minutes—a pre-established and fixed length—a game can be extended for as long as the player wishes. This criterion was therefore set to limit the corpus, and so it took 12 hours to complete the main mission in each language.

While playing, the types of synchrony observed in the different game situations were noted in a chart (see the examples in Tables 4 and 5). These game situations alternated continuously throughout the game, depending on the player's action: every time a new game situation was triggered, a new record was added to the chart and the type of synchrony observed in it was registered. The types of synchrony observed in *BAK* were classified based on the definitions given by professionals and academics (see section 3.1) about the characteristics of each type of synchrony. This procedure allowed the types of synchrony found in this video game to be classified empirically, its correspondence to the different constraints expected being observed (see section 3.2) and represented in Table 2.

It should be acknowledged that the descriptive analysis of only the product, without analysing the process, is incomplete. Analysing the process will require another step in this research that I intend to carry out in the future. This article constitutes a mere first step in the descriptive analysis of the product that requires further exploration; such exploration should include the localization process.

4.2 Results

It took 12 hours to complete the main goal in the Spanish version, the same time it took to play in English. In the process, 553 records were obtained for the game situations found in the Spanish-dubbed version of *BAK*, as opposed to 535 records for the English version.

The following two examples represent two different game situations (task and dialogue, respectively); each indicates the storyline moment and observes a specific time constraint for each game situation.

Table 3: Task with wild synchrony in *BAK*

<i>BAK</i>	<i>Storyline moment:</i>	<i>Mission: Use the antenna in Panessa Studios to determine Scarecrow's position.</i>		
EN	Batman: The generator's damaged. I need to bypass it and power the antenna directly.			
ES	Batman: El generador está dañado. Tendré que darle energía a la antena directamente.			
Game situation	Task			
Synchrony EN	VO	Synchrony ES	VO	
Comments				

Table 4: Dialogue with sound-sync in *BAK*

<i>BAK</i>	<i>Storyline moment:</i>	<i>Mission: Investigate the Clock Tower for clues to helping find Oracle.</i>		
EN	Joker: Nothing like a trip down ol' memory lane, eh Bats? And now a front row seat to the main event! You would've thought old Jim Gordon had been through enough. First some handsome young maniac cripples his daughter. Now he's blaming himself for her being taken. You really should tell him this is your fault, Bats. I'm sure he'll understand. I mean, it's not like you went and got his daughter killed. Well, not yet anyway. The night's still young.			
ES	Joker: Nada como una pequeña visita al baúl de los recuerdos. ¡Un asiento en primera fila para el espectáculo! Creerás que el bueno de Jim ya habrá tragado bastante. Primero un joven y apuesto maniaco deja tullida a su niña. Y ahora se culpa a sí mismo por su secuestro. Deberías decirle que eres el responsable, Batman, seguro que lo entiende. Ni que fueras allí y consiguieras que la matasen... Bueno, nunca se sabe. La noche es joven.			
Game situation	Dialogue			
Synchrony EN	SS	Synchrony ES	SS	
Comments				

In the first case (Table 3), a task is introduced in the form of the character's thoughts. Wild synchrony has been applied in both versions since the tasks in *BAK* are always transmitted through a voice-over (VO). In the second case (Table 4), the dialogue is dubbed applying SS: the expressiveness in both versions resembles the audio, and the pauses and length have been reproduced exactly. However, the lip movement is not absolutely precise; as a result, the level of restriction is not as strict as it could be in a cinematic, where the player would not be allowed to move the camera.

With respect to the way in which each game situation is synchronized in both versions, the following instances were observed:

Table 5: Types of synchrony in *BAK*

	<i>ES</i>	<i>EN</i>
<i>VO</i>	272 – 49.19%	256 – 47.85%
<i>TC</i>	76 – 13.74%	72 – 13.46%
<i>STC</i>	12 – 2.19%	7 – 1.31%
<i>SS</i>	67 – 12.12%	50 – 9.36%
<i>Lip-sync</i>	125 – 22.6%	149 – 27.85%
<i>TOTAL</i>	553 records	535 records

The two versions display similar percentages for each type of synchrony, although the Spanish version tends to be slightly less restrictive because there are more cases of VO, TC and STC (relaxed types of synchrony). There are also more cases of SS in Spanish than in English, but the reason is that lip-sync is more frequent in English, making the audio in the original text a bit more restrictive, although these differences are very small. While the semiotic codes operating in both versions should be the same (see Table 2), the small changes in the linguistic code during the localization process might have caused a slight loss of accuracy in the audio synchronization.

Table 6: Types of synchrony according to game situation in *BAK* (percentages)

	<i>Tasks</i>		<i>Game action</i>		<i>Dialogues</i>		<i>Cinematics</i>	
	<i>ES</i>	<i>EN</i>	<i>ES</i>	<i>EN</i>	<i>ES</i>	<i>EN</i>	<i>ES</i>	<i>EN</i>
<i>VO</i>	100	100	62.6	62	27.6	26.6	14.3	14.3
<i>TC</i>	0	0	36.3	37	2.9	0	3.1	3.1
<i>STC</i>	0	0	0	0	5.2	2.4	3.1	3.1
<i>SS</i>	0	0	0.5	0.5	32.8	23.7	9.2	9.2
<i>Lip-Snc.</i>	0	0	0.5	0.5	32.2	47.3	69.4	69.4

The types of synchrony distributed according to the game situations (Table 6) are displayed differently in each situation: tasks are always transmitted by an off-screen voice, therefore applying VO synchrony; game action switches from VO synchrony to TC, and only a very small number of units present as SS and lip-sync. Dialogues are the most complex situation: sometimes they partially restrict the player's action, forcing them to see the dialogue (lip-sync, SS or STC) (although camera movements and walking are permitted); and sometimes they do not interfere with the player's action at all (TC or VO). Finally, cinematics tend to use lip-sync combined with VO and a few instances of SS, STC and TC.

A comparison of the results for the English and the Spanish versions indicates that the clearest difference in the types of synchrony applied is in the dialogues: in Spanish, lip-sync is used in a smaller percentage because other types of less-restrictive synchrony have been applied instead (STC and SS); in English, the biggest percentage is lip-sync. In tasks, game action and cinematics, however, the results for both versions are either identical or almost identical.

In view of these results, it can be concluded that, in *BAK*, game situations do suggest a different level of restriction:

- tasks are always dubbed VO;
- game action combines mainly VO and TC;
- in dialogues, STC, SS and lip-sync are the most commonly applied types of synchrony; and, finally,
- cinematics represent the most restrictive situation, with the majority being lip-sync.

This relates to the ideas described in section 2: the more restrictive the situation is (because more semiotic codes operate and have an impact on the linguistic code), the more strictly the type of synchronization will be applied.

This analysis shows how an additional dimension (i.e., interaction transmitted through the tactile channel), one not yet included in traditional audiovisual texts (see section 2), motivates a new research approach based on game situations instead of the temporal distribution of the material (see section 2.2, Table 1). The analysis further identifies different typologies of dubbing synchrony (up to five: VO, TC, STC, SS and lip-sync) in an action-adventure video game.

7. Concluding remarks

As discussed, video games are a technological and multimodal space that creates different forms of interaction between the on-screen events (acoustic and visual codes) and the player (normative codes). Consequently, their localization process encompasses a wide range of technical and functional aspects that need to be adapted in order to maintain the gameplay created by the multiple dimensions (Muñoz, 2017, p. 12) of a video game.

Bearing these features in mind, this article has offered a review of the product's different semiotic dimensions. More specifically, it has discussed how the player interacts with the video game through three different channels: those described for traditional audiovisual texts, that is, the visual and acoustic channels, and a third, interactive (bidirectional) dimension transmitted through a tactile channel when using a controller. In the case of the game analysed, the normative codes that regulate the game action are transmitted through the tactile channel. In this system of codes, each movement made by the player is codified as a sign (external normative code) that represents a specific action within the game world (internal normative code).

As a result of this specific configuration of video-game contents, a research approach has been proposed based on the distribution of the multimodal contents in different game situations, which are a consequence of interactivity (and which do not exist in traditional audiovisual texts). Each game situation is triggered by a sign, that is, an action carried out by the player, which is also one of the signs contained in the normative code system that regulates interaction.

To narrow down the analysis, synchronization was chosen as the feature to be described using this game-situation structure. The aim was to study how the multimodal configuration of a video game creates the need to design a new taxonomy of the types of synchronization.

In view of the results, it can be concluded that in the action-adventure game *Batman: Arkham Knight* a new taxonomy of synchronies can be used, depending on how

restrictive each game situation is. A game situation can be more or less restrictive depending on the different codes involved to create meaning: the more the codes involved, the more restrictive the situation is, and therefore the stricter the synchronization.

For future research, game situations could be used to map out the characteristics of dubbing, subtitling or any other aspect of video-game localization. However, the corpus of analysis should first be expanded in order to confirm whether this distribution of game situations and these types of synchrony apply to different video games and genres or, on the contrary, whether there are no clear patterns (norms) in their dubbing.

Regarding the limitations of the study, one has just been mentioned: the need to expand the corpus to confirm these tendencies in a broader range of titles. Moreover, the comparison between the translated text and the replaced segments in the original version (Toury, 1995) could pose some problems owing to the interactive dimension of the product: depending on the player's actions, situations could alternate differently. Therefore, the exact same content might not be displayed when playing both versions, since any slight variation in the player's actions could trigger different reactions in the game.

Finally, it should be pointed out that research on video-game localization, particularly regarding specific practices in the localization process, such as dubbing or subtitling, is still scarce and only recent (Granell et al., 2015), and it should be expanded in order not only to benefit professional translators and localization companies (Méndez González & Calvo-Ferrer, 2017), but also to contribute to the improved training of future professionals. As Bernal Merino states (2015),

new knowledge and training are necessary to tackle MIES [Multimedia Interactive Entertainment Software] creations because they are some of the most theoretically and technically complex products translation professionals have to work with. (P. 248)

Therefore, the game industry as a whole would benefit from further research on the localization process, and Translation Studies would also be enriched with new research approaches which consider the different domains that form as complex an audiovisual product as video games.

References

- AEVI. (2017). *Anuario de la industria del videojuego*. Retrieved from http://www.aevi.org.es/web/wp-content/uploads/2017/06/ANUARIO_AEVI_2016.pdf.
- Belinchón, G. (2017, January). El cine en España tuvo más de 100 millones de espectadores en 2016. *El País*. Retrieved from https://elpais.com/cultura/2017/01/02/actualidad/1483350472_132303.html.
- Bernal Merino, M. Á. (2015). *Translation and localisation in video games: making entertainment software global*. New York, NY: Routledge.
- Chaume, F. (2003). *Doblatge i subtitulació per a la TV*. Vic, Spain: Eumo.
- Chaume, F. (2004). *Cine y traducción*. Madrid, Spain: Cátedra.
- Chaume, F. (2005). Los estándares de calidad y la recepción de la traducción audiovisual. *Puentes*, 6, 5–12.
- Chaume, F. (2012). *Audiovisual translation: Dubbing*. Manchester, UK: St. Jerome.
- Granell, X., Mangiron, C. & Vidal, N. (2015). *La traducción de videojuegos*. Seville, Spain: Bienza.
- ICEX (2017, September). La industria musical Española sale de gira. *El Exportador*. Retrieved from <https://www.icex.es/icex/es/Navegacion-zona-contacto/revista-el-exportador/observatorio2/REP2017726165.html#>.
- Jewitt, C. (2014). *The Routledge handbook of multimodal analysis* (2nd ed.). Oxon, UK: Routledge.
- Juul, J. (2005). *Half-real: videogames between real rules and fictional worlds*. Cambridge, MA: MIT Press.

- Kaindl, K. (2013). Multimodality and translation. In C. Millán & F. Bartrina. (Eds.), *The Routledge handbook of translation studies* (pp. 257–269). London, UK: Routledge.
- Kokkola, S., & Ketola, A. (2015). Thinking outside the “Methods Box”: New avenues for research in multimodal translation. *Vakki publications*, 4, 219–228.
- Kress, G. & Van Leeuwen, T. (2001). *Multimodal discourse. The modes and media of contemporary communication*. London, UK/New York, NY: Oxford University Press.
- LaPS4 (2016, October 6). *Entrevista a Guiomar Alburquerque, la VOZ de LARA CROFT en España* [Video file]. Retrieved from <https://www.youtube.com/watch?v=Xbk3GXgHPC8>.
- Levis, D. (1997). *Los videojuegos, un fenómeno de masas*. Barcelona, Spain: Paidós.
- Le Dour, C. (2007, February). Surviving audio localization. *Gamasutra*. Retrieved from http://www.gamasutra.com/view/feature/1713/surviving_audio_localization.php#5
- López Redondo, I. (2014). *¿Qué es un videojuego?* Seville, Spain: Héroes de Papel.
- Maietti, M. (2004). *Semiotica dei videogiochi*. Milan, Italy: Unicopli.
- Mangiron, C. (2017). Research in game localization: An overview. In X. Zhang & S. Strong. (Eds.), *The Journal of Internationalization and Localisation* 4(2), 74–99.
- Mangiron, C., O’Hagan, M. & Orero, P. (Eds.). (2014). *Fun for all: translation and accessibility practices in video games*. Bern, Switzerland: Peter Lang.
- Mangiron, C., & O’Hagan, M. (2013). *Game localization*. Amsterdam, Netherlands: John Benjamins.
- Maxwell-Chandler, H. (2012). *The Game localization handbook* (2nd ed.). Sudbury, ON: Jones & Bartlett Learning.
- Méndez González, R., & Calvo-Ferrer, J. R. (2017). *Videojuegos y [para]traducción: aproximación a la práctica localizadora*. Granada, Spain: Comares.
- Méndez González, R. (2015). Localización de videojuegos: Fundamentos traductológicos innovadores para nuevas prácticas profesionales. Vigo, Spain: Servizo de Publicacións UV.
- Muñoz Sánchez, P. (2017). *Localización de videojuegos*. Madrid, Spain: Síntesis.
- O’Hagan, M. (2007). Video games as a new domain for translation research: From translating text to translating experience. *Revista Tradumàtica*, 5. Retrieved from <http://www.fti.uab.es/tradumatica/revista/num5/articulos/09/09.pdf>.
- O’Hagan, M., & Chandler, H. (2016). Game localisation research and translation studies: Loss and gain under an interdisciplinary lens. In v. Y. Gambier (Ed.), *Border crossings: Translation studies and other disciplines* (pp. 309–330). Amsterdam, Netherlands: John Benjamins.
- Pérez Latorre, Ó. (2012). El significado en juego: Sobre la dimensión lúdica del discurso del videojuego. *Proceedings of the 10th World Congress of the IASS/AIS, 1901–1910*. University of A Coruña, Spain.
- Pérez Latorre, Ó. (2010). Análisis de la significación del videojuego: Fundamentos teóricos del juego, el mundo narrativo y la enunciación interactiva como perspectivas de estudio del discurso. (Doctoral dissertation). Pompeu Fabra University, Barcelona, Spain.
- Poyatos, F. (2002). *Nonverbal communication across disciplines*, 2. Amsterdam, Netherlands: John Benjamins.
- Pujol Tubau, M. (2015). *La representació de personatges a través del doblatge en narratives transmèdia* (Doctoral Dissertation). University of Vic – Central University of Catalonia, Barcelona, Spain.
- Salen, K., & Zimmerman, E. (2005). Game design and meaningful play. In J. Raessens & J. Goldstein. (Eds.), *Handbook of computer game studies* (pp. 59–79). Cambridge, MA: MIT Press.
- Sioli, F., Minazzi, F. & Ballista, A. (2007, October–November). Audio localization for language service providers. *Multilingual Localization: Getting Started Guide*, 18–23.
- Siwek, S. E. (2017). *Video games in the 21st century: The 2017 Report*. ESA. Retrieved from <http://www.theesa.com/category/around-the-industry/business-economy/>.
- Scholand, M. (2002). Localización de videojuegos. *Revista Tradumàtica*, 1. Retrieved from <http://www.fti.uab.es/tradumatica/revista/articulos/mscholand/art.htm>.
- Vázquez Rodríguez, A. (2018). El error de traducción en la localización de videojuegos: Estudio descriptivo y comparativo entre videojuegos indie y no indie (Doctoral dissertation). University of Valencia, Spain.
- Wijman, T. (2017, November 28). New gaming book: Newzoo ups its 2017 global games market estimate to \$116.0Bn growing to \$143.5Bn in 2010. *Newzoo*. Retrieved from

<https://newzoo.com/insights/articles/new-gaming-boom-newzoo-ups-its-2017-global-games-market-estimate-to-116-0bn-growing-to-143-5bn-in-2020/>.

Xbox España (2015, November 5). *FanFestXbox – Tertulia Doblaje en Videojuegos con David García y Guiomar Alburquerque* [Video file]. Retrieved from <https://www.youtube.com/watch?v=Tzyrqin7nwc>.

-
- 1 For example, an Excel spreadsheet listing names for different weapons could be updated with new weapons and functions, or some may be deleted, even after the document has been translated. This requires the translator to re-translate the text several times.
 - 2 The normative code for the main character in this video game can be seen in the following YouTube video: EK2 (2015, June 23). *Batman Arkham Knight – Controls Dualshock 4 Wireless Controller (All Features)*. [Video file]. Retrieved from <https://www.youtube.com/watch?v=IGTvJbwJSt4>.
 - 3 These five types of synchrony in video games have been described by professionals such as Maxwell-Chandler (2012), Le Dour (2007), Sioli et al. (2007), and by the academics Mangiron and O'Hagan (2013), Bernal Merino (2015) and Pujol Tubau (2015).