

The aesthetics of isochrony and literal synchrony in voice-over translation

Alfonso Carlos Rodríguez Fernández-Peña



Alfonso Carlos
Rodríguez Fernández-
Peña
Universidad de Oviedo
rodriguezalfonso@uniovi.es;
ORCID:
[0000-0002-0273-8442](https://orcid.org/0000-0002-0273-8442)

Abstract

Voice-over translation is characterised by some technical synchronic features (isochrony, literal synchrony, action synchrony, kinetic synchrony, content synchrony and character synchrony). From these, isochrony and literal synchrony contribute to the illusion of authenticity and realism with what is called sound bites (a time span in the target version in which we only hear the original voice, and which can occur at the beginning and/or at the end of the speaker's intervention). In our study, after analysing a corpus made up of different voiced-over programmes using speech analysis software and a spreadsheet, we have seen that the average duration of sound bites differs from that stated by the scholarly tradition both in terms of seconds and number of words. In addition, we also analysed samples that show no literal synchrony to see how and whether the rendition success of those parts could be affected. The results confirm that sound bites and literal synchrony are aesthetic enhancers which provide voice-over with an authenticity feel that makes it, for some scholars, the most faithful and reliable audiovisual translation mode.

Keywords: synchrony, isochrony, literal synchrony, voice-over, sound bites.

Resum

La traducció de veus superposades es caracteritza per diversos requeriments tècnics, com ara la isocronia, la sincronia literal, cinètica, d'acció, de contingut i de personatge. De totes, les dues primeres contribueixen a la il·lusió d'autenticitat i realisme gràcies a allò que anomenem sound bites (un desfasament d'isocronia que permet sentir el text original uns segons abans i/o després de la intervenció de l'interlocutor). En aquest estudi, després d'analitzar un corpus format per diversos programes amb veu en off per mitjà d'un software d'anàlisi de la parla i una fulla de càlcul, hem comprovat que la durada mitjana dels fragments de so difereix de l'establerta per la major part dels acadèmics tant en segons com en nombre de paraules. També hem analitzat mostres que no presenten sincronia literal a fi d'observar si la traducció d'aquestes parts podria veure's afectada i de quina manera. Això confirmaria que, en cas que n'hi hagin, els sound bites i la

sincronia literal tenen una funció purament estètica en la traducció de veus superposades.

Paraules clau: sincronia, isocronia, sincronia literal, veus superposades, sound bites

Resumen

La traducción de voces superpuestas se caracteriza por diferentes requerimientos técnicos como la isocronía, la sincronía literal, cinética, de acción, de contenido y de personaje. De todas estas, las dos primeras contribuyen a la ilusión de autenticidad y realismo gracias a lo que se denomina *sound bites* (un desfase de isocronía que permite oír el texto original unos segundos antes y/o después de la intervención del interlocutor).. En nuestro estudio, tras analizar un corpus formado por distintos programas con voz en off mediante un software de análisis del habla y una hoja de cálculo, hemos comprobado que la duración media de los fragmentos de sonido difiere de la establecida por la mayoría de los académicos tanto en segundos como en número de palabras.. También hemos analizado muestras que no presentan sincronía literal para observar si la traducción de estas partes podría verse afectada y de qué modo. Esto confirmaría que, en caso de haberlos, los *sound bites* y la sincronía literal sirven un propósito meramente estético en la traducción de voces superpuestas.

Palabras clave: sincronía, isocronía, sincronía literal, voces superpuestas, sound bites.

1. Introduction

Voice-over has become one of the dominant translation modes in recent years in the Spanish audiovisual market. It can be found in programmes such as documentaries, reality shows, corporate web videos, and even in filmed interviews broadcast on TV channels as diverse as news companies like Euronews or the TVE (Televisión Española – the Spanish public TV company –), and sports networks like DAZN and Eurosport. What once was labelled as the ‘ugly duckling’ of audiovisual translation modes (Orero, 2006c), seems to be the favourite translation method for most of the audiovisual unscripted content both on TV, video on demand (VOD) platforms and web company videos, although there are exceptions. Thanks to its cost-cutting and cost-effective nature, voice-over translation takes less studio and script translation time since there is no need for lip synchronisation and phonetic adaptation (Barzdevics, 2012: 67, Gorska, 2015: 68). Moreover, voice-over is regarded, a priori, as a faithful and authentic representation of the original text according to the definitions attempted to describe this AVT mode (Franco et al. 2010: 26) since, amongst other things, the original audio track can be heard at low level while the translation is being provided orally. This gives voice-over translation a feel of realism and faithfulness that other AVT modes do not manage to transmit (Chaume, 2004a: 35, Orero, 2006b: 175).

In this study we will provide a state of the art concerning voice-over isochrony which shows how the scholarly tradition has measured (in seconds and words) this synchronic feature. Our work aims to complement that of Baños (2019) in which she researched the isochrony and literal synchrony of voiced-over reality shows. However, our study includes more categories of voiced-over shows and provides, in our view, a wider scope. For this purpose, we will analyse and measure, by means of the audio analysis software

Praat, two synchronic features of voice-over translation (isochrony and literal synchrony), that will help us understand how these are used concerning the delay effect, the literal translation provided (if any) and the way they affect the final translated product.

2. Technical characteristics of voice-over translation

Voice-over is regarded as technically less complex than dubbing (Diaz Cintas & Orero, 2006: 478; Diaz Cintas & Orero, 2010: 441; Rica Peromingo, 2016: 143), and, therefore, a more economical translation mode (Schwarz 2011: 402). The most noticeable technical features that any viewer of a voiced-over programme can notice are that the original voice is not deleted (no need to technically remove the original audio track when mixing) and that there is a time gap, or break, at the beginning and the end of the utterances during which one can hear for a few seconds the original source text voice, known as *sound bites*. This means that the translated voice starts a few seconds after the original one and ends a few seconds before the original, and is, for scholars such as Orero (2006b: 175) “voice-over’s most characteristic feature”.

There is no given or specific pattern as to how long this delay should last or as to whether this should change according to the category or programme in which voice-over is used, since there are examples within the researched literature that show different possibilities. Accordingly, Baños (2019), in her study of voiced-over reality shows, finds that “the general practice of leaving a few seconds at the beginning of the utterance is disregarded in the voice-over of reality TV programmes” (2019: 272). This scholar found that in Spanish reality shows the duration of the translation hardly matches that of the original, thus giving voice actors and sound engineers enough discretion to deliver the text in synchrony with the original or to use a slight asynchrony. This issue is quite controversial (Franco et al. 2010: 80) since there is no agreement among scholars and practitioners as to how long this gap between the original and target texts should be or if there should be a gap at all. The existing sound bite recommendations according to the audiovisual scholarly tradition are shown in the table below.

AUTHOR	SOUND BITE LENGTH
Luyken, Herbst, Langham-Brown and Spinhof (1991: 80)	“allow the original sound to be heard for several seconds at the onset of speech [...]”
Mayoral (2001: 35)	“La sincronización exige el inicio tras una pausa de tres o cuatro palabras [...]” “Synchronization requires starting after a pause of three or four words [...]”
Orero (2004: 83)	“Some two seconds before and after.”
Espasa (2004: 189)	“Two seconds before and after, or three or four words.”

Chaume (2004a: 35)	“The actor should wait a few seconds (not more than 2 or 3) until the character or narrator starts talking [...]”
Chaume (2004b: 45)	“[...] the translation often coming in two or three seconds after the narrator or screen character has started to speak.”
Orero (2005: 219)	“[...] esta suele comenzar dos segundos después [...] y suele terminar dos segundos antes [...]” “[...] it usually starts two seconds after [...] and ends two seconds before [...]” (my translation)
Orero (2006a: 261)	“A final constraint, which could also be called a literal synchronicity, is the slight delay of the delivery of literal translation of first utterances [...]”
Orero (2006b: 175)	“[...] hearing a few seconds of the original recording.”
Darwish & Orero (2006: 136)	“While first Arabic few words of the excerpted segment (below) are audible [...] Usually in news items there are no seconds left at the end, so it is almost impossible to check if the speaker finished the discourse at the same time as the voice-over.”
Orero (2009: 132)	“some seconds into the beginning of the original utterance – and sometimes finishes before the actual person on screen – allowing the viewer to hear part of the original although, this practice is not universal.”
Franco, Matamala and Orero (2010: 80)	“A few seconds might be left at the beginning and even at the end.”
Diaz Cintas & Orero (2010: 447)	“[...] a few seconds at the onset of the speech [...] The translation typically finishes a couple of seconds before the foreign language speech does [...]”
Schwarz (2011: 402)	“It is common practice to allow several seconds before the TL speech takes over. Quite often the SL voice can be heard again at the end of the utterance.”
Wozniak (2012: 216)	“[...] trailing the original soundtrack by a second or two or even covering it completely.”

Chaume (2012: 03)	“The translation is heard a few seconds after the original voices.”
Sepielak (2013: 95)	“[...] leaving the original speech audible in the foreign language for a few seconds at the onset and at the end of the segment.”
Rica Peromingo (2016: 144)	“the translation usually starts some 1-3 seconds after the original and ends some 1-3 seconds before the original.”
Sepielak (2016: 1057)	“[...] leaving some words of the original soundtrack audible at the beginning and end of the utterance.”
Ameri & Khoshsaligheh (2018: 12)	“[...] as the common practice is that translation begins a few seconds after the original utterance and it finishes a few seconds earlier.”
Matamala (2019: 64)	“[...] the translation usually begins some words after the original and finishes some words before the latter ends.”

Table 1. Duration of sound bites according to different scholars.

As can be seen from the statements shown in these descriptions, the view on the duration of this time gap differs among the different authors. Some scholars recommend just a few seconds, others 1 to 3 seconds, a few words, three or four words (to be precise), or even imply that the target voice covers the original completely.

Nonetheless, the time gap or sound bite is not the only technical feature of voice-over translation. Unlike dubbing, in which lip synchrony is paramount, the degree of synchrony in voice-over is less demanding and just requires matching the original audio in time length, which is known among professionals as *phrase synchrony*. While this feature makes voice-over considerably cheaper than dubbing, this economic factor, together with the relative easiness of synchrony, might, at the same time, explain the boom of voiced-over products in the audiovisual industry in Spain in recent years.

However, the idea that voice-over does not impose any demands for synchrony and that the sole requirement is just to produce a text translation which is similar to the original in terms of temporal length, is a sweeping generalisation, to say the least. The only actual technical difference from dubbing is that lip sync is replaced with phrase sync; words do not have to match the exact movement of the lips in voice-over, instead phrases, intonational phrases¹ (IPs) to be precise, have to match the original ones to a certain extent and, given that in English into Spanish translations target texts tend to be longer than the originals (Orero, 2006a: 260; Reynolds & Sizemore, 2010: 473), this poses a serious problem.

¹ An intonational phrase stands for the chunks of spoken material that speakers utter when speaking and can consist of one word or multiple words (Wells 2006: 06)

Apart from this technical feature, voice-over translation is mainly characterised by four synchronic features, as stated by Orero (2006a) and Matamala (2019): isochrony, literal synchrony, kinetic synchrony, and action synchrony. However, other scholars, such as Fodor (1976), Schwarz (2011) and Rodríguez Fernández-Peña (2020), also include character synchrony and content synchrony.

What follows is a brief definition of these features before delving into those under study in this article (isochrony and literal synchrony):

- Kinetic synchrony: the body language of the person on screen has to be “in sync” with the translation that is being provided orally.
- Action synchrony: there should be an agreement between what the voice delivering the translation is saying and what is shown on screen in terms of content and image.
- Content synchrony: this exists when the semantic content of the source text matches closely the semantic content of the target text.
- Character synchrony: the voice providing the translation has to be similar in terms of age, gender, and pitch to the voice delivering the source text.
- Isochrony: the duration of the voice delivering the translation has to match the original voice.
- Literal synchrony: a literal translation is provided at the beginning and at the end of the utterances.

From these characteristics, our study focuses on isochrony and literal synchrony, and we will analyse their relationship with respect to sound bite length. Therefore, we will now cover in more depth the concepts of isochrony and literal synchrony and how these affect the translation process.

2.1 Isochrony

Isochrony, from Greek *isos* (same) and *khronos* (time), in AVT implies that the duration of the translated voice has to match the original one. Schwarz (2011: 399) considers that the most essential synchrony is quantitative rather than qualitative (lip synchrony), adding that “To an audience, the divergence of visually and acoustically perceived utterances is extremely distracting.” This is really important when video editing in post-production is not possible, since there may be utterances for which the translation could be longer than the originals, as is often the case in translation from English into Spanish (30% according to Orero 2006a: 260), and there is no chance to edit the video according to the time lengths of each language. Normally, the video track is untouchable, and translators have to manage to make their lines fit in the given time codes. This is one of the main challenges of voice-over isochrony for Orero, for whom, in these cases, the translated texts have to be reduced or condensed, and she suggests three “strategies” (sic.) (Orero, 2006a: 260):

1. In interviews, she suggests eliminating all the *fatic* (sic.) elements of discourse; repetitions, question tags, mistakes, corrections, etc. proper of colloquial language.

2. In some language pairs like English/Spanish, scientific terminology is used more in the target text than in the source text; and in interviews it is often the case that the post-production team can edit the video to fit the voice track.
3. In interviews, the production team can help by editing the video file and adding more images so the translated text can fit and match the original.

Sepielak (2016) understands isochrony in a way that goes against the literal meaning of the term (*iso chronos* = same time) and distinguishes three different types of isochrony: full isochrony, initial isochrony and final isochrony. She states that *full isochrony* happens in voice-over when at least one word is heard at the beginning and the end of the utterance (2016: 1061). This means that there is a delay or sound bite to hear that word which, in our view, goes against the literal definition of isochrony. *Initial isochrony* implies that at least one word is heard only at the beginning of the utterance, and *final isochrony* that one word is heard at the end. Again, for us, these terms go against the meaning of isochrony since the duration of the target text is not the same as the original text. We consider that isochrony implies that the target text has to match the original one in terms of timing and that it is not longer and out of sync with the original video.

2.2 Literal synchrony

This consists of a literal or semiotic translation of the source text at the beginning and the end of the utterances, that is, the sound bites, which is the moment when the audience can perceive the original and its translation consecutively. Although this feature is widely known as literal synchrony, some scholars such as Chaume (2012: 70) understand that there are no temporal nuances and, therefore, there is no synchrony but rather coherence. Although we agree with Chaume's understanding of this voice-over feature (2012), and *coherence* seems a more precise term than *synchrony*, in this article we will use the traditional label (synchrony) since recent studies on voice-over translation keep using it (Matamala 2019: 68; Rodríguez Fernández-Peña 2020: 48) and it seems to be consolidated in the AVT terminology.

This literal synchrony or coherence provides a sense of authenticity and realism, as the audience can then understand the source language (in the case of English) and will be able to match what they hear first (in English) with what is translated and delivered a few seconds later (in Spanish in our case). If there is a semiotic match, the audience will probably be happy with the translation and trust it; if, on the other hand, there is no semiotic match, the trust bond might break since the audience may start thinking that the version in the voice-over does not really reflect what the original speaker is actually saying. Franco et al (2010: 81) point out that literal synchrony is mostly done with proper nouns, which can be spotted by the audience even if they are not fully understood, and according to Luyken et al (1991: 141), even a well-considered semantic translation will occasionally be insufficient, and a literal translation will be required to create the effect of a simultaneous interpretation, which contributes to the authenticity effect.

Franco et al. (2010) provide relevant information about literal synchrony and sound bite length with a survey conducted on voice-over. The authors discovered that

although the general practice is to leave a few seconds at the beginning, there seems to be some variation, and the same happens with the seconds left at the end, since sometimes this time gap disappears (2010: 172-173). In any case, the fact that this gap exists makes the translation process more challenging, because the time available for the voice actor to deliver the translation is shorter.

3. Corpus description

Considering that voice-over is a translation mode that may be found in a wide range of audiovisual programmes, eleven videos were chosen to provide a representative sample, ranging from documentaries to reality shows, TV news interviews and web videos. These programmes include a wide range of topics, were produced and broadcast during a 12-year period, and were translated by freelance and in-house translators, as well as journalists in the case of TV news interviews (Matamala, 2008: 118). Furthermore, our corpus for analysis is comparable to others utilized for similar objectives, such as Baños (2019), who studied voice-over synchronization in three episodes of reality shows, and Sanchez-Mompean (2019), who researched tonal patterns in dubbing in six episodes of a TV sitcom.

Three documentaries, three short TV interviews, three reality shows and two web videos make up our corpus. Table 2 below shows the original English names of the programmes and their translations into Spanish, the duration, the year they were created and the broadcast medium which aired them in Spain.

DOCUMENTARIES				
Original Name	Spanish Name	Duration	Broadcast	Year
The last days of Anne Boleyn	Los últimos días de Ana Bolena	59 minutes	RTVE/ Youtube	2013
Jamie's comfort food	La comida reconfortante de Jamie	24 minutes	RTVE/Youtube	2014
Tasteology: Chill & Experience	Saborología: Frío y Experiencia	13 & 19 minutes	Youtube	2016
TV NEWS INTERVIEWS				
Original Name	Spanish Name	Duration	Broadcast	Year
Le Mag (Bruce Dickinson interview)	Le Mag (entrevista a Bruce Dickinson)	2:33 minutes	Euronews	2015

Reporter: What's it like for EU migrants in Britain amid UKIP's success?	Reporter: Turismo social, ¿un problema en Gran Bretaña?	8:21 minutes	Euronews	2014
I-Talk: Is it time to scrap the CAP?	I-Talk: (¿Debería el Reino Unido quedarse en la Unión Europea?	9:48 minutes	Euronews	2013
REALITY SHOWS				
Original Name	Spanish Name	Duration	Broadcast	Year
Cupcake Wars - Season 6 Episode 13	Guerra de Cupcakes - Temporada 6 Episodio 13	36:04 minutes	Divinity/Youtube	2010
Ramsay's Kitchen Nightmare's - Season 1 Episode 1	Pesadilla en la Cocina UK - Temporada 1 Episodio 1 "Bonaparte's"	47:47 minutes	Nova/Youtube	2004
Teen Mom OG - Season 6 Episode 6 (online trailers 1 and 2).	Teen Mom OG - Temporada 6 Episodio 6 (tráilers online 1 & 2)	4 minutes	MTV	2016
CORPORATE WEB VIDEOS				
Original Name	Spanish Name	Duration	Broadcast	Year
Columbia Threadneedle Update January 2017	Boletín sobre Renta Variable Europea 2017	5 minutes	Company website	2017
Columbia Threadneedle Investment Philosophy	Filosofía de Inversión	10 minutes	Company website	2016

Table 2. Corpus Selection

4. Objectives and Analysis

The main goal of this study is to understand how isochrony and literal synchrony are applied in Spanish voice-over translation. Do translators and voice actors have to comply

with them to contribute to the desired authenticity effect? Or are these mere enhancers of the audio-visual experience? Do sound bites behave in the same way as they did in Baños’s (2019) study?

The analysis of these texts has been conducted using two tools: a spreadsheet and Praat, a speech analysis software. With the Excel spreadsheet we compared both texts (source and target) and included their respective time codes for each scene. We resorted to filters so that we could track examples of no sync (literal and isochrony). This gave us a clear picture in terms of quantity and quality. We were able to quantify the number of examples in which isochrony and literal synchrony were not respected, and also analyse these instances to assess whether the lack of synchronicity affects the successful rendition of the source text.

With Praat we analysed sound bites from the target texts and measured the time lapse and word count for each of them. This would provide us with an idea of whether the academic trend is followed or not in terms of sound bite length (seconds and wordcount).

4. 1 Isochrony

After watching the voiced-over translated versions of the programmes in our corpus, we identified 5 scenes in which isochrony is out of sync; 1 belongs to one documentary and 4 to a TV news interview programme (all of these are found in Euronews’ *iTalk: Is it time to scrap the CAP?*). There are no isochronic issues in reality shows and web videos in the programmes analysed here.

In the following, we will comment on the examples found that show no isochrony and how that can affect the target text. These are listed in the tables below and include the following information: the time frame of the scene is displayed in the columns labelled “In” (when the scene begins) and “Out” (when the scene ends); we can identify the person speaking on screen in the column labelled “Character”; the English original text is displayed in the column “English Original,” and the Spanish translation is provided in the column “Spanish Translation.”

In	Out	Character	English Original	Spanish Translation
00:10:20	00:10:25	Interviewer, female 3	...food, for weeks, and it will. I just think if you guys...	(none)

Table 3. No isochrony example 1 from *Tasteology: Chill*

The line belongs to a woman who is standing on a stage and talking to the audience about food. It lasts just five seconds, and it is not really relevant to the story of the documentary. However, considering that she is the only speaker on screen, we believe

that her lines should have been translated. Leaving this utterance untranslated implies, for us, that this scene is out of sync.

The remaining 4 cases in which isochrony is out of sync were found in *Euronews: is it time to scrap the CAP?*

In	Out	Character	English Original	Spanish Translation
00:35	00:36	Daniel Hannan, male	Thank you.	(none)

Table 4. No isochrony example 2 from *Euronews Is it time to scrap the CAP?*

As happened in the previous case, there is no translation for this utterance, so there is no sync. Although we think that the audience of a channel such as Euronews understands what “thank you” means, since it is a very common English expression, we believe that a translation for this passage is necessary.

In	Out	Character	English Original	Spanish Translation
02:16	02:35	Daniel Hannan, male	Part of the problem in our whole European project is being that different countries, with different visions of how society and the economy should work, have reached compromises that satisfy nobody. And that's why I'd be much happier with a situation where Britain lives under its own laws and made no attempt to advise any neighbouring country on how they should run their affairs...	Parte del problema del proyecto europeo es que diferentes países con diferentes puntos de vistas sobre el funcionamiento de la sociedad y de la economía han firmado acuerdos que no satisfacen a nadie. Por eso, yo estaría mucho más satisfecho si el Reino Unido pudiese vivir aplicando sus propias leyes sin pedirle consejo a ningún país vecino.

Table 5. No isochrony example 3 from *Euronews Is it time to scrap the CAP?*

In this time frame, when Mr. Hannan stops talking, the interviewer interrupts him and both speak at the same time. There is no translation for that final part, probably because it is already confusing in the original version since both are speaking simultaneously and their voices overlap. There is a translation for what the interviewer says (see below) but not for Mr. Hannan's words, who is the person being interviewed.

In	Out	Character	English Original	Spanish Translation
02:35	02:42	Chris Burns	But what about making compromises in order to be part of a larger whole? What about making compromises to be part of a larger whole or a more powerful whole?	¿Por qué no encontrar compromisos para así formar parte de una Unión más fuerte?

Table 6. No isochrony example 4 from Euronews *Is it time to scrap the CAP?*

These are the interviewer's words when he interrupts Mr Hannan. We can see and hear that the interviewer has to ask the question twice because the first time he is interrupting Mr Hannan. His first question is left untranslated since the interruption implies that there are two people talking (so a third voice would probably be too much). Therefore, the translation is only provided for the second question.

In	Out	Character	English Original	Spanish Translation
05:12	05:15	Chris Burns	So, how much are you worried? Is the door going to slam shut on all the Brits working on the continent?	¿Le preocupa que las oportunidades laborales se cierren a los británicos si abandonan la UE?

Table 7. No isochrony example 5 from Euronews *Is it time to scrap the CAP?*

In this scene the translation (the target voice) ends around one second after the original, and therefore it is out of sync.

These examples show no isochrony due to the fact that there is no translation provided for a particular portion of the source text and, therefore, we only get the original voice in the translated version of the programme. In addition, there is one example in which the target version lasts more than the original.

4. 2 Sound bite length

Now that we have seen the examples that show no isochrony, we will analyse those that are in sync from an isochronic perspective. Here we will study the duration of the sound bites, and for this purpose we have used Praat, a speech analysis software, to provide graphical evidence of how long these bites are and how many seconds and words these include. The image below shows how we measured this using the software.

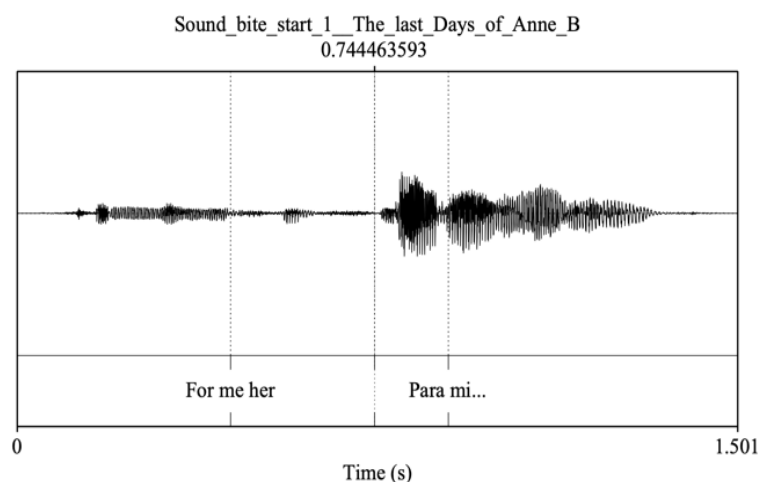
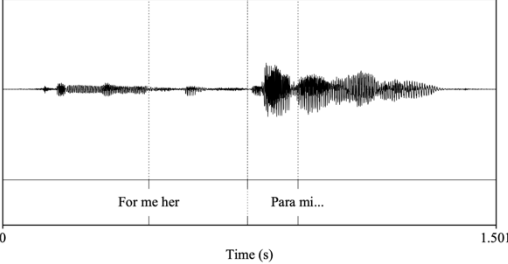
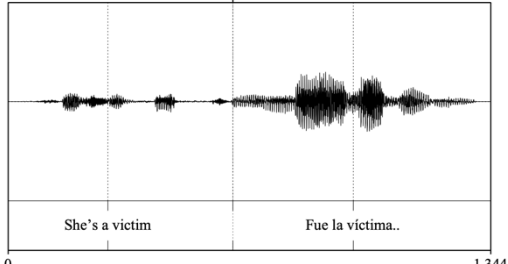
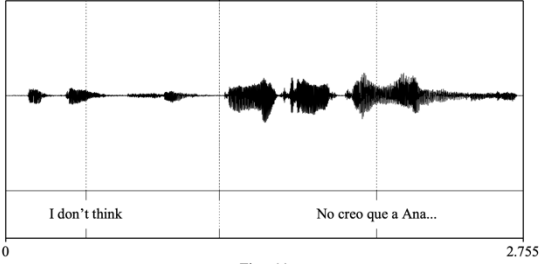


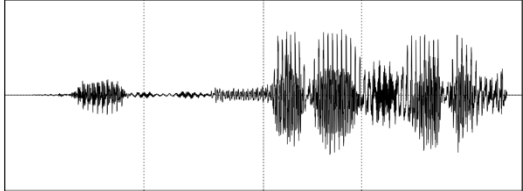
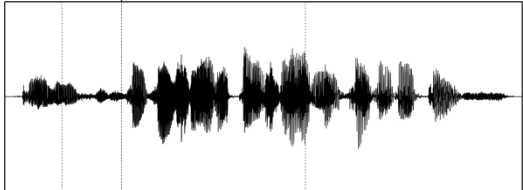
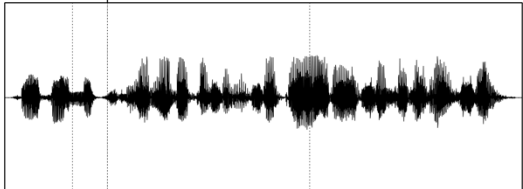
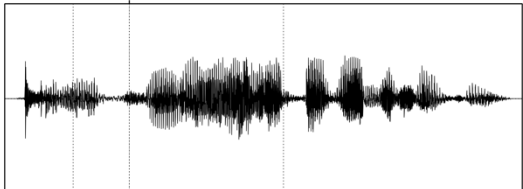
Image 1. Sample analysis using Praat

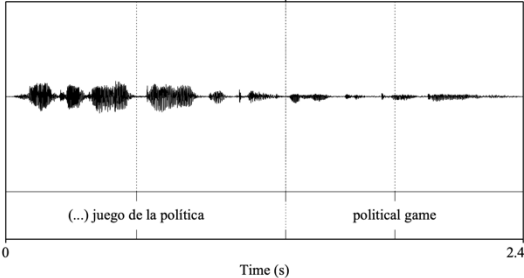
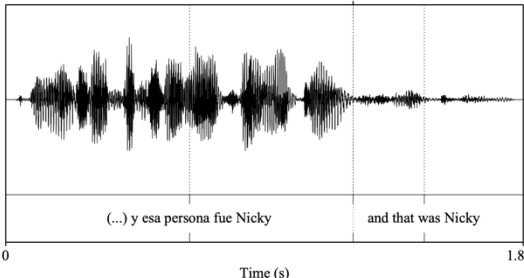
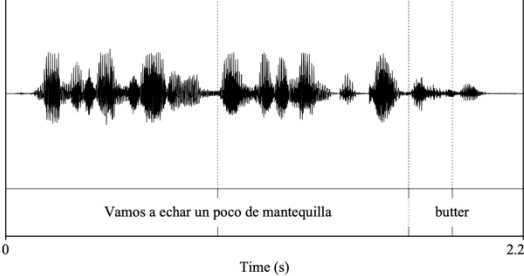
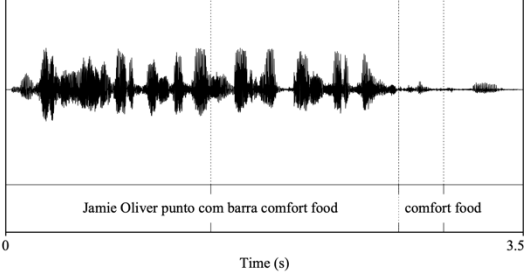
The image above shows the audio wave of that passage. We can see that the audio wave on the left is thinner than the one on the right. This is because the left wave belongs to the original English text, which is reduced in volume, and the one on the right is wider, as it shows the Spanish translation and is louder. Below the audio wave we can see the transcription of the passage (English source text on the left side and Spanish on the right). Finally in the horizontal axis we have the time frame for this example, which is 1.501 seconds. All the samples taken for analysis are studied in this way, showing the audio wave, the transcription of the passage (so that we know the exact number of words) and the duration in seconds of the scene.

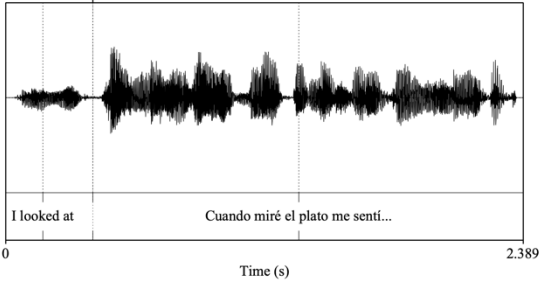
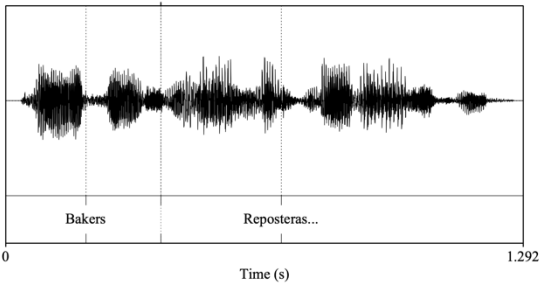
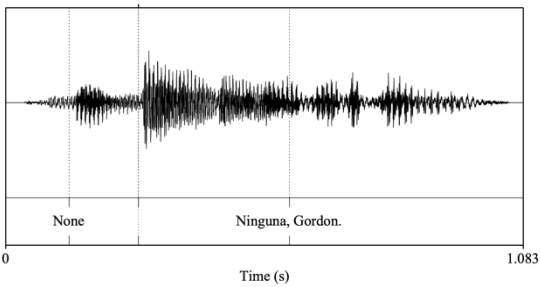
In terms of sound bite length, we have noticed that the duration suggested by some scholars, shown in the table 1 above, is not present and/or consistent all the time in the programmes of our corpus. There seems to be no rule of thumb regarding the duration of the sound bites, which may depend on two aspects: the length of the target script to be voiced, and the pace used by the original speaker. Thus, if we take, for example, the documentary *The last days of Anne Boleyn*, we can see that sound bite length differs almost every time depending on who speaks on camera. This is in line with Sepielak's (2013) analysis of English-Spanish and English-Polish documentaries, for whom "this fundamental assumption of voice over is rather neglected [...] Less than one third of the segments could be categorized as keeping full isochrony" (2013: 95). Here Sepielak (2013) understands «full isochrony» as when we can hear the original voice at the beginning and at the end of the utterance for a few seconds. However, we do not share this idea of full isochrony since, in our view, isochrony implies that the target text has to match the source text in length .


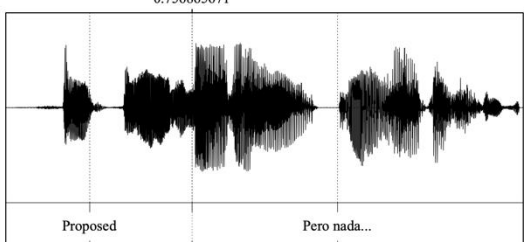
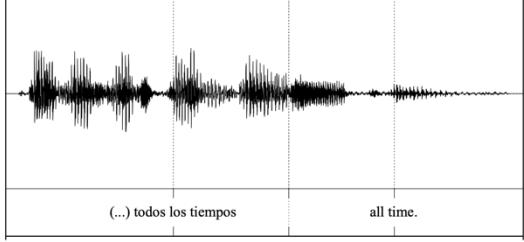
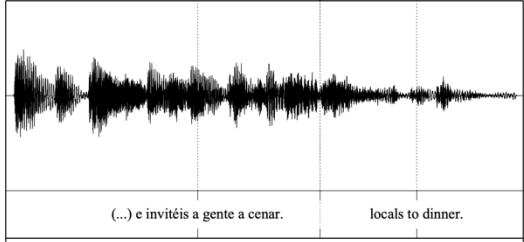
The table below shows some audio bites (at the start and at the end of certain phrases) in terms of audio wavelength and the number of words used for each one.

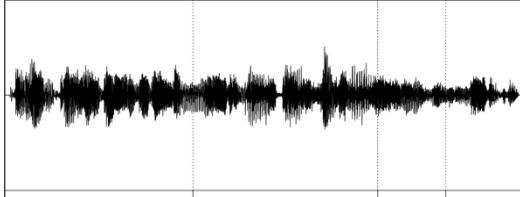
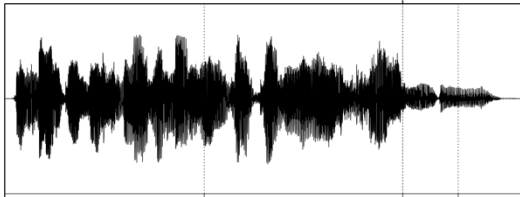
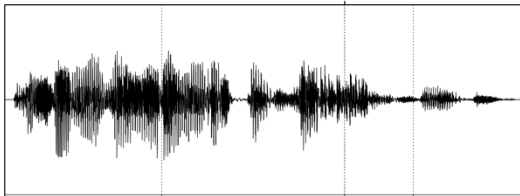
SOUND BITES		
Documentaries		
Starting Sound Bites		
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
<p>Sound_bite_start_1_The_last_Days_of_Anne_B 0.744463593</p>  <p>0 1.501 Time (s)</p> <p>For me her Para mi...</p>	0.75 SECS.	3 WORDS
<p>Sound_bite_start_2_The_last_days_of_Anne_B 0.62587614</p>  <p>0 1.344 Time (s)</p> <p>She's a victim Fue la victima..</p>	0.62 SECS.	4 WORDS
<p>Sound_Bite_start_3_The_last_days_of_Anne_B 1.13728284</p>  <p>0 2.755 Time (s)</p> <p>I don't think No creo que a Ana...</p>	1.13 SECS.	3 WORDS

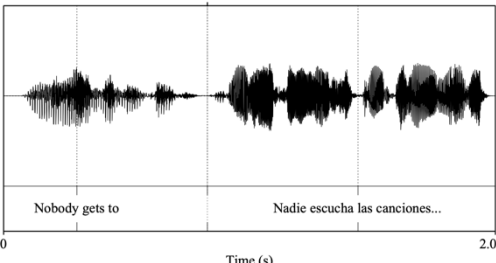
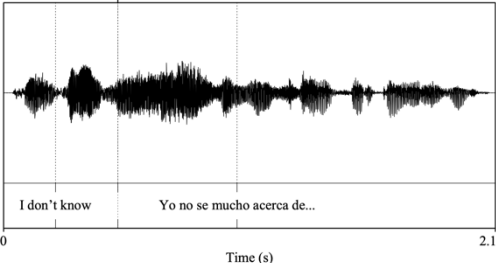
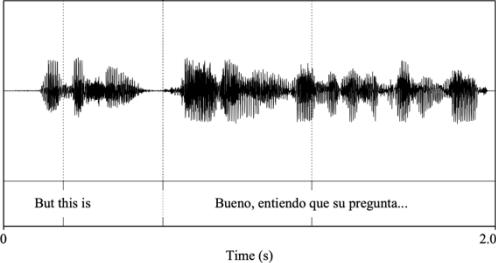
<p>Sound_bite_start_4_Tasteology_Chill 0.371689342</p>  <p>After Tras organizar..</p> <p>0 0.7434 Time (s)</p>	<p>0.37 SECS.</p>	<p>1 WORD</p>
<p>Sound_bite_start_5_Tasteology_Chill 0.539247932</p>  <p>Maybe this Quizás este sea mi favorito</p> <p>0 2.389 Time (s)</p>	<p>0.53 SECS.</p>	<p>2 WORDS</p>
<p>Sound_bite_start_6_Jamies_Comfort_Food 0.635040187</p>  <p>So if I knock this out De modo que si la saco ahora se quedará así</p> <p>0 3.199 Time (s)</p>	<p>0.63 SECS.</p>	<p>6 WORDS</p>
<p>Sound_bite_start_7_Jamies_Comfort_Food 0.405331829</p>  <p>I'm gonna turn it off Y apagamos el fuego</p> <p>0 1.684 Time (s)</p>	<p>0.40 SECS.</p>	<p>5 WORDS</p>
<p>Ending Sound Bites</p>		
<p>AUDIO WAVE</p>	<p>TIME FRAME</p>	<p>NUMBER OF WORDS</p>

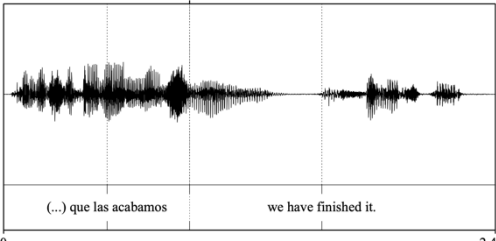
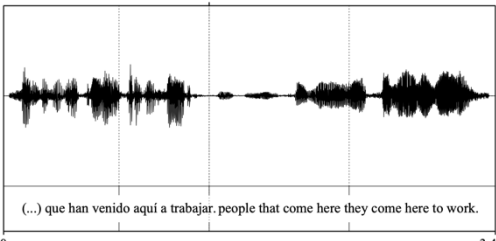
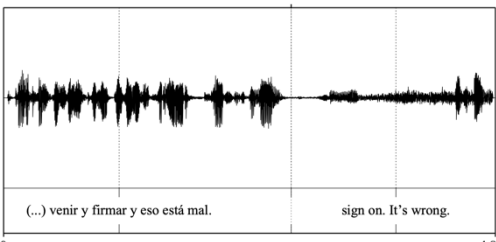
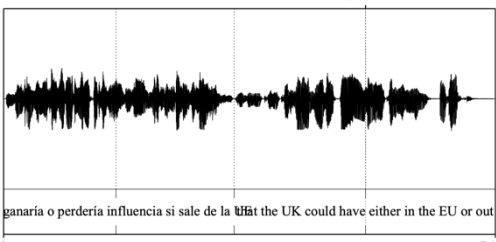
<p>Sound_bite_end_1_The_last_days_of_Anne_B 1.34903181</p>  <p>0 2.494 Time (s)</p> <p>(...) juego de la política political game</p>	<p>1.14 SECS.</p>	<p>2 WORDS</p>
<p>Sound_bite_end_2_Tasteology_Chill 1.25296838</p>  <p>0 1.867 Time (s)</p> <p>(...) y esa persona fue Nicky and that was Nicky</p>	<p>0.61 SECS.</p>	<p>4 WORDS</p>
<p>Sound_bite_end_3_Jamies_Comfort_food 1.71770562</p>  <p>0 2.206 Time (s)</p> <p>Vamos a echar un poco de mantequilla butter</p>	<p>0.49 SECS.</p>	<p>1 WORD</p>
<p>Sound_bite_end_4_Jamies_Comfort_Food 2.68667525</p>  <p>0 3.538 Time (s)</p> <p>Jamie Oliver punto com barra comfort food comfort food</p>	<p>0.85 SECS.</p>	<p>2 WORDS</p>
<p>Average Start Sound Bite in Documentaries</p>		
	<p>0.63 SECS.</p>	<p>3.4 WORDS</p>

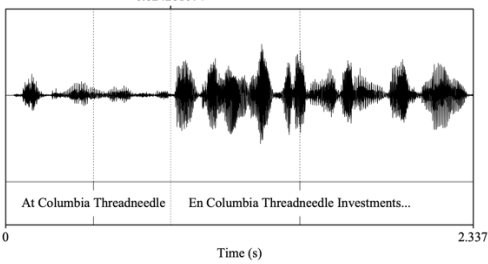
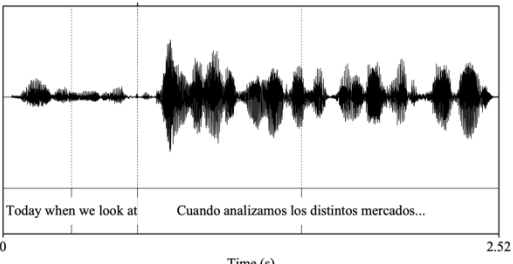
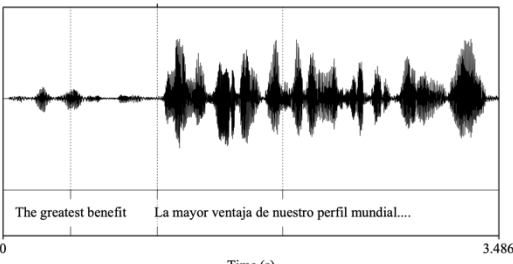
Average Ending Sound Bite in Documentaries		
	0.77 SECS.	2.2 WORDS
B) Reality Shows		
Starting Sound Bites		
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
<p style="text-align: center;">Sound_bite_start_8__Cupcake_Wars</p> <p style="text-align: center;">0.402253704</p>  <p style="text-align: center;">Time (s)</p>	0.40 SECS.	3 WORDS
<p style="text-align: center;">Sound_bite_start_9__Cupcake_Wars</p> <p style="text-align: center;">0.387348101</p>  <p style="text-align: center;">Time (s)</p>	0.38 SECS.	1 WORD
<p style="text-align: center;">Sound_bite_start_10__Kitchen_Nightmares</p> <p style="text-align: center;">0.277677262</p>  <p style="text-align: center;">Time (s)</p>	2.27 SECS.	1 WORD

<p>Sound_bite_start_11_Teen_Mom_clip_1 0.891883354</p>  <p>0 1.814 Time (s)</p> <p>I mean No quiero que Simon...</p>	<p>0.89 SECS.</p>	<p>2 WORDS</p>
<p>Sound_bite_start_12_Teen_Mom_clip_1 0.756865671</p>  <p>0 2.102 Time (s)</p> <p>Proposed Pero nada...</p>	<p>0.75 SECS.</p>	<p>1 WORD</p>
<p>Ending Sound Bites</p>		
<p>AUDIO WAVE</p>	<p>TIME FRAME</p>	<p>NUMBER OF WORDS</p>
<p>Sound_bite_end_5_Cupcake_Wars 0.849634715</p>  <p>0 1.553 Time (s)</p> <p>(...) todos los tiempos all time.</p>	<p>0.70 SECS.</p>	<p>2 WORDS</p>
<p>Sound_bite_end_6_Kitchen_Nightmares 1.48269099</p>  <p>0 2.441 Time (s)</p> <p>(...) e invitáis a gente a cenar. locals to dinner.</p>	<p>0.96 SECS.</p>	<p>3 WORDS</p>

<p style="text-align: center;">Sound_Bite_end_7_Kitchen_Nightmares 2.68190903</p>  <p style="text-align: center;">Time (s)</p> <p style="text-align: center;">0 3.721</p> <p style="text-align: center;">(...) están hundidos en mierda hasta las cejas. eyeballs in shitter</p>	1.04 SECS.	3 WORDS
<p style="text-align: center;">Sound_bite_end_8_Teen_Mom_clip_1 2.01779882</p>  <p style="text-align: center;">Time (s)</p> <p style="text-align: center;">0 2.624</p> <p style="text-align: center;">(...) voy a diseñar un anillo con mi hija daughter</p>	0.61 SECS.	1 WORD
<p style="text-align: center;">Sound_bite_end_9_Teen_Mom_clip_1 1.43183334</p>  <p style="text-align: center;">Time (s)</p> <p style="text-align: center;">0 2.18</p> <p style="text-align: center;">Ese anillo era un despilfarro. any sense.</p>	0.75 SECS.	2 WORDS
Average Starting Sound Bites in Reality Shows		
	0.93 SECS.	1.6 WORDS
Average Ending Sound Bites in Reality Shows		
	0.81 SECS.	2.2 WORDS
C) TV News Interview Programmes		
Starting Sound Bites		

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
<p style="text-align: center;">Sound_Bite_start_13_Euronews_Iron_Maiden 0.86009564</p>  <p style="text-align: center;">Time (s)</p>	0.86 SECS.	3 WORDS
<p style="text-align: center;">Sound_bite_start_15_Euronews_What's_it_like_for_EU_migrants_in_Britain_amid_UKIP's_success 0.488203787</p>  <p style="text-align: center;">Time (s)</p>	0.48 SECS.	3 WORDS
<p style="text-align: center;">Sound_Bite_start_16_Euronews_Is_it_time_to_scrap_the_CAP 0.672367311</p>  <p style="text-align: center;">Time (s)</p>	0.67 SECS.	3 WORDS
Ending Sound Bites		
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS

<p>Sound_bite_end_10_Euronews_Iron_Maiden 0.933197616</p>  <p>0 2.467 Time (s)</p>	1.53 SECS.	4 WORDS
<p>Sound_bite_end_11_Euronews_What's_it_like_for_EU_migrants_in_Britain_amid_UKIP's_success 1.43539847</p>  <p>0 3.434 Time (s)</p>	2 SECS.	9 WORDS
<p>Sound_bite_end_12_Euronews_What's_it_like_for_EU_migrants_in_Britain_amid_UKIP's_success 2.84811137</p>  <p>0 4.871 Time (s)</p>	2.03 SECS.	4 WORDS
<p>Sound_bite_end_13_Euronews_Is_it_time_to_scrap_the_CAP 5.29666734</p>  <p>0 7.196 Time (s)</p>	1.9 SECS.	10 WORDS
Average Starting Sound Bites in TV News Interview Programmes		
	0.67 SECS.	3 WORDS
Average Ending Sound Bites in TV News Interview Programmes		

	1.86 SECS.	6.7 WORDS
D) Web Videos		
Starting Sound Bites		
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
<p style="text-align: center;">Sound_bite_start_17__TN_Philosophy 0.824261074</p>  <p style="text-align: center;">Time (s)</p>	0.82 SECS.	3 WORDS
<p style="text-align: center;">Sound_bite_18__TN_Philosophy 0.683230346</p>  <p style="text-align: center;">Time (s)</p>	0.68 SECS.	5 WORDS
<p style="text-align: center;">Sound_bite_19__TN_Philosophy 1.08433381</p>  <p style="text-align: center;">Time (s)</p>	1.08 SECS.	3 WORDS

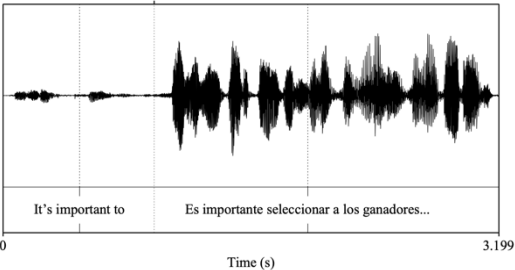
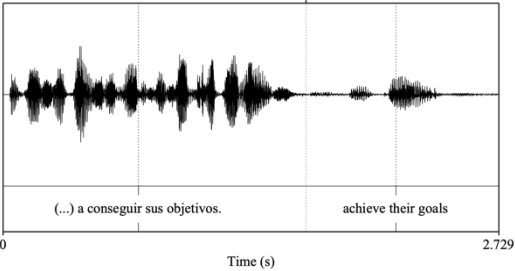
	0.97 SECS.	3 WORDS
Ending Sound Bites		
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
	1.06 SECS.	3 WORDS
Average Starting Sound Bites in Web Videos		
	0.88 SECS.	3.5 WORDS
Average Ending Sound Bites in Web Videos		
	1.06 SECS.	3 WORDS
Average General Starting Sound Bite		
TOTAL	0.77 SECS.	2.8 WORDS
Average General Ending Sound Bite		
TOTAL	1.12 SEC	3.5 WORDS

Table 8. Sound bite analysis

The information provided in the table above is summarised in the following table.

Category	Opening Bite		Closing bite	
	Seconds	Words	Seconds	Words
Documentaries	0.63	3.4	0.77	2.2
Reality Shows	0.93	1.6	0.81	2.2
TV news interviews	0.67	3	1.86	6.7
Web Videos	0.88	3.5	1.06	3
Average	0.77	2.8	1.12	3.5

Table 9. Sound bite length summary

As can be seen in the table above, the sound bites analysed from our corpus show that the average starting bite is 0.77 seconds, with 2.8 words, while the ending sound bite is 1.12 seconds, with 3.5 words. This information clashes with the time frame and number of words specified by most of the scholars, as stated in table 1, since the opening sound bite does not reach one second and the closing bite lasts an average of less than two seconds. This result contrasts with the two or three seconds recommended by most of the researchers. In terms of word count, the opening bite has an average of 2.8 words and the closing bite has 3.5, while the literature suggests two or three words, or a few words. The way sound bites are measured with this speech analysis software shows detailed time frames for the sound bites and, since we do not round up the results, the differences compared to the academic tradition are evident.

What the table does show is that the time frame allocated to each time is decided at random. In other words, there is no compulsory time frame to use. In fact, all the videos analysed in our corpus (with the exception of one web video) contain times with no sound bites because the target voice starts and finishes at the same time as the original. This coincides with Baños's (2019: 273) analysis of sound bites in Spanish reality TV programmes, for which there is a "lack of clear guidelines as far as the implementation of this type of synchrony (isochrony) is concerned". However, this scholar finds differences between reality TV show and documentaries as regards voice-over isochrony and literal synchrony (2019: 280), and states that, "The analysis carried out reveals that the general practice of leaving a few seconds at the beginning of the utterance is disregarded in the voice-over translation of reality tv programmes." (2019: 272). In our case, after analysing both documentaries and reality TV programmes, together with TV news interviews and web videos, we find no difference concerning how isochrony and literal synchrony are applied, as will be shown in the next section. Thus, we detected three patterns in all voice-over categories or genres: 1) both voices start and finish at the same time; 2) there is a sound bite at the beginning but not at the end because both

versions end together; and, 3) vice versa, that both versions start at the very same time and there is a sound bite at the end.

An interesting factor is that the Web Video category contains almost no sound bites. This is due to the fact that one of the videos (TN Europe Jan 17) has been edited in the target version and the target script is so long that it does not match the original video. The only sound bites available are at the beginning and at the end of the video for the first and last words uttered by the presenter, being the original audio track removed in the rest of the video. Therefore, we consider these sound bites are not relevant since they have been produced and probably edited in the postproduction stage. On the other hand, the web video, TN Philosophy, does have sound bites, most of them being starting bites, and with only one final sound bite.

The videos and sound bites analysed for each programme have shown that sound bites are used when possible. Sometimes the original speaker's speech is pacy and there is not enough time frame to fit the target text without sounding rushed, so the sound bite is not used. In other cases, the target script might be too long to leave some time for sound bites and, consequently, these are not included.

4.3 Literal synchrony

Literal synchrony consists of a literal translation (mostly word for word) of the sound bites, which contributes to reinforcing the sense of authenticity and realism that characterises voice-over translation. This synchronic feature is a quality marker in voice-over translation of interviews (Orero, 2005: 219) and encourages translators to be “as literal as possible” in the parts belonging to the first and last sound bite. In our corpus, we have noticed that literal synchrony is present in almost all the videos analysed. Given the fact that the presence of sound bites can vary, as seen in the previous section, and may occur either at the beginning of the utterance only, or at the end, or both at the beginning and end, literal-sync is subordinated to the presence of these sound bites. This implies that literal-sync will be sacrificed in situations when synchronicity is hard to keep due to time constraints.

We have identified 45 cases in our corpus in which there is no literal synchrony; that is, when the source and target texts do not match literally at the beginning and at the end of the utterance. The table below shows the number of examples of no literal-sync found according to programme and category.

Literal Synchrony (cases of no sync)		
DOCUMENTARIES	<i>The last days of Anne Boleyn</i>	18
	<i>Jamie's Comfort Food</i>	9
	<i>Tasteology: Chill</i>	1

	<i>Tasteology: Experience</i>	0
	Total Documentaries	28
TV News Interview Videos	<i>Euronews: What's it like for EU migrants in Britain amid UKIP's success?</i>	3
	<i>Euronews: Is it time to scrap the CAP?</i>	6
	<i>Euronews: Iron Maiden interview</i>	1
	Total TV News Interview Videos	10
REALITY TV	<i>Kitchen Nightmares</i>	1
	<i>Cupcake Wars</i>	0
	<i>Teen Mom 1</i>	3
	<i>Teen Mom 2</i>	3
	Total Reality TV	7
WEB VIDEOS	<i>TN Philosophy</i>	6
	<i>TN Europe</i>	0
	Total Web videos	6
TOTAL		51

Table 10. List of examples of no literal-sync

We will now analyse some examples of no literal-sync. These examples are displayed in tables including the same information as the examples analysed in the section devoted to isochrony. Therefore, the information provided includes the time frame of the scene in the columns labelled “In” (when the scene begins) and “Out” (when the scene ends); the person speaking on screen in the column labelled “Character”; the English original text is displayed in the column “English Original”, in which the words that can be heard during the bite are underlined. The Spanish translation is provided in the column “Spanish Translation.” A literal word for word translation of the sound bite is provided in the column on the right-hand side labelled “Literal Translation”.

In	Out	Character	English Original	Spanish Translation	Literal Translation
00:13:47	00:13:49	Philippa Gregory	<u>The belief was</u> , in the medieval world,	En el mundo medieval, cuando una mujer daba a luz	La creencia era...
00:23:10	00:23:16	Hilary Mantel	<u>By 1536</u> , the conflict is ready to explode.	Dos años después, el conflicto estaba a punto de estallar.	Para 1536...
00:42:55	00:43:00	David Starkey	<u>The final driver</u> of everything under Henry, is Henry.	Durante su reinado, quien tomaba las decisiones en última instancia era Enrique.	El último responsable...

Table 11. No literal synchrony example 1 from *The last days of Anne Boleyn*

In	Out	Character	English Original	Spanish Translation	Literal Translation
03:31	03:42	Chris Burns	<u>There's a bit of a segue</u> to our next question. It's from a gentleman in Spain. Question of influence. How much influence could the UK have either in the European Union or out?	Perfecto, Daniel, pasemos a la siguiente pregunta sobre la influencia. ¿El Reino Unido ganaría o perdería influencia si sale de la UE?	Pasemos a la ...
05:12	05:15	Chris Burns, male	<u>So, how much are you worried?</u> Is the door going to slam shut on all the Brits working on the continent?	¿Le preocupa que las oportunidades laborales se cierren a los británicos si abandonan la UE?	¿Cuánto te preocupa...

Table 12. No literal synchrony example 2 from *Euronews: Is it time to scrap the CAP?*

In	Out	Character	English Original	Spanish Translation	Literal Translation
00:03	00:04	Female	<u>For how long is Kate going?</u> What's going on? How long is she.	¿Cómo va? ¿Cuánto tiempo estará?	¿Cuánto tiempo...
00:12	00:14	Female	<u>Yeah. And she went</u> 'cause she wanted to go.	Y... ¿ha ido por voluntad propia?	Vale. Y fue...
00:15	00:52	Male	She so wanted to go. She just said, «I don't wanna fake this anymore.» <u>I am like,</u> «OK, you know what, dude? Get your ass up!»	Quería ir. Dijo que ya no quería sentirse así más. (...) Tenía que insistirle para que se despertara.	Pensé, OK...

Table 13. No literal synchrony example 3 from MTV's Teen Mom OG trailer 2

In	Out	Character	English Original	Spanish Translation	Literal Translation
00:33	00:52	Bruce Dickinson	Nobody gets to hear the music, uuuh until we have finished it. Uhh, so, the manager doesn't get to come down to the studio. <u>He's banned.</u> The record company don't get to come down to the studio. Nobody gets to hear it until we've finished it.	Nadie escucha las canciones hasta que las acabamos. Nuestro manager no puede entrar en el estudio. Los responsables de la compañía discográfica tampoco. Nadie puede escuchar nuestros discos hasta que no están totalmente acabados,	Está vetado.

Table 14. No literal synchrony example 4 from Euronews' Le Mag (interview with Bruce Dickinson)

In	Out	Character	English Original	Spanish Translation	Literal Translation
00:00:53		Jim Cielinski	<u>Today</u> when we look across markets, we see central banks doing things they've never done before in history.	Cuando analizamos los distintos mercados, observamos que los bancos centrales toman medidas sin precedentes.	Hoy...
00:02:50		William Davies	<u>But as</u> we do the research, the assumptions we make are dependent on what is going on around the whole world, not just within that company.	A medida que profundizamos en el análisis, basamos nuestras hipótesis en lo que ocurre en el mundo y no solo en la empresa.	Pero mientras...

Table 15. No literal synchrony example 5 from TN Philosophy

If we now consider how these cases of no literal-sync relate to the successful transfer of meaning in the target text, we could say that this synchronic feature is irrelevant for that purpose. Literal synchrony is, in our view, irrelevant to achieve a successful rendition of the source text, since there may not be any sound bites in a voiced-over video in the first place. Therefore, the presence of sound bites and the literal synchronisation of the target text with the original bites are more a matter of time-related possibilities (the target text is short and there is time enough to have sound bites), and style, which contributes to the feeling of authenticity and faithfulness that, for some scholars, characterises this translation mode.

7. CONCLUSIONS

Voice-over is, as we know, a constrained translation method, with visual and temporal restrictions that affect the translation process, which means that the translated text has to be synchronised with the limits imposed by the source text. These synchronic features, though restrictive and demanding, can be exploited by translators to provide a sense of authenticity and realism to the target text, as is the case with literal, kinetic and character synchronisation. However, these synchronic aspects do not need to be present or applied all the time in Spanish voice-over translation since, as we have observed in our corpus, they are actually frequently missing with no serious consequences in the final effect from a translational perspective. However, the fact that certain utterances include sound bites and others do not could be disturbing for the audience of the programme. If the audience

is used to, or expects, sound bites and literal synchrony to reinforce the idea of faithfulness that is usually attributed to voice-over translation, the chance that these elements are not provided could jeopardise the tacit agreement in terms of reliability that is set between the audience and the translation mode. Therefore, we consider that further research concerning the reception of translated utterances with no sound bites should be conducted. This could shed light on a practice that seems to be commonplace in voice-over translation.

As regards isochrony and sound bites, we have noted that there is no set time duration for these, and that they are present only when possible, which means that these bites – which allow literal synchrony – are sacrificed when isochronic needs have to be met. Our understanding of sound bite length is more a matter of how long the translated script is and whether or not there is the chance to leave this gap at the beginning and the end. Here, Baños's (2019: 273) question "Do voiced-over reality programmes observe the same translation conventions and norms as voiced-over non-fictional products such as documentaries?" could be answered positively, at least in terms of isochrony and literal sync. This would contradict her conclusion that, "the voice-over of reality TV in Spain present differences if compared to those observed in documentaries, especially as regards voice-over isochrony and literal synchrony." (2019: 279). In her view, reality TV voice-over undergoes a "more flexible type of isochrony, where the duration of the translation has to correspond approximately with that of the original [...]" (ibid) but, after presenting the evidence obtained from analysing reality TV, documentaries, TV news interview programmes and web videos, we can state that this "more flexible type of isochrony" is not just specific to reality TV but to the other audiovisual products that use voice-over as a translation mode. Whether a new term should be coined to refer to this flexible synchronic practice, as suggested by Baños, in order to expand the current definitions of voice-over translation is not recommended, in our view, given the blurriness and chaos around the term *voice-over* that already exists, as claimed by scholars such as Franco (2000: 30; 2001: 294), Sanchez-Mompean (2015: 273) and Rodríguez Fernández-Peña (2020: 35), among others.

On the other hand, we agree with Baños on the irrelevance of literal synchrony when synchrony does not allow sound bites, which confirms that it only works as an enhancer that provides a sense of authenticity and realism for the target audience. Literal synchrony has no effect on the successful rendition of the source text. We have found examples of sound bites in this study which were long enough to include a literal translation, but the translator's decision was not to use this resource, and that did not encumber the final result (the conveyance of a translation).

This paper has focused mainly on two synchronic aspects of voice-over translation (isochrony and literal synchrony) and how these are treated in the different genres or programme categories that use voice-over as a translation mode. With this study we have attempted to continue Baños's work (2019) and answer the final remark in her study of synchronic features in voiced-over reality shows in Spain, for whom, "the situation is likely to be similar in comparable scenarios" (2019: 280). As illustrated, isochrony and literal synchrony act as aesthetic enhancers which make the end product

more faithful and realistic for the target audience and are treated similarly no matter the category or genre that resorts to voice-over to provide the translation.

References

- Ameri, S.; Khoshsaligeh, M. (2018). Voice-over translation of news programs on television: evidence from Iran. *Hikma: revista de traducció*, v. 17, pp. 9-32, <<https://doi.org/10.21071/hikma.v17i0.11029>>. [Accessed: 20220702].
- Baños, R. (2019). Translating reality TV into Spanish: When fast-food TV challenges AVT conventions. *Perspectives: Studies in Translation Theory and Practice*, v. 27, n. 2, pp. 265-282. <<https://doi.org/10.1080/0907676X.2018.1424221>>. [Accessed: 20220502]
- Barzdevics, I. A. (2012). Un recorrido por la voz superpuesta. In: Martínez Sierra, J. (ed.) *Reflexiones sobre la traducción audiovisual: tres espectros, tres momentos*. València: Universitat de València, pp. 57-68.
- Chaume, F. (2004a). *Cine y traducción*. Madrid: Cátedra. (Signo e imagen).
- Chaume, F. (2004b). Synchronization in dubbing: A translational approach. In: Orero, P. (ed.). *Topics in audiovisual translation*. Amsterdam/Philadelphia: John Benjamins, pp. 35-52. (Benjamin Translation Library; 56). <<https://doi.org/10.1075/btl.56.07cha>>. [Accessed: 20220502].
- Chaume, F. (2012). *Audiovisual translation: Dubbing*. Manchester: St. Jerome. (Translation Practices Explained).
- Darwish, A.; Orero, P. (2014). Rhetorical dissonance of unsynchronized voices: Issues of voice-over in news broadcasts. *Babel*, v. 60, n. 2, pp. 129-144. <<https://doi.org/10.1075/babel.60.2.01dar>>. [Accessed: 20220502].
- Díaz Cintas, J.; Orero, P. (2006). Voice-over. In: *Encyclopedia of Language & Linguistics*. 2nd ed. [S.l.]: Elsevier, pp. 477-479. <<https://doi.org/10.1016/B0-08-044854-2/00473-9>>. [Accessed: 20220502].
- Díaz Cintas, J.; Orero, P. (2010). Voiceover and dubbing. In: Gambier, Y.; Van Doorslaer L. (eds.). *Handbook of Translation Studies*, v. 1, pp. 441-445. <<https://doi.org/10.1075/hts.1.voi1>>. [Accessed: 20220502].
- Espasa, E. (2004). Myths about documentary translation. In: Orero, P. (ed.). *Topics in audiovisual translation*. Amsterdam/Philadelphia: John Benjamins, pp. 183-198. (Benjamins Translation Library; 56). <<https://doi.org/10.1075/btl.56.19esp>>. [Accessed: 20220502].
- Fodor, I. (1976). *Film Dubbing: Phonetic, Semiotic, Esthetic and Psychological Aspects*. Hamburg: Helmut Buske.
- Franco, E. (2000). *Revoicing the alien in documentaries: Cultural agency, norms and the translation of audiovisual reality*. Unpublished PhD Dissertation. Catholic University of Louvain.

- Franco, E. (2001). Voiced-over television documentaries: Terminological and conceptual issues for their research. *Target: International Journal of Translation Studies*, v. 13, n. 2, pp. 289-304. <<https://doi.org/10.1075/target.13.2.05fra>>. [Accessed: 20220502].
- Franco, E.; Matamala, A.; Orero, P. (2010). *Voice-over translation: An overview*. Bern [etc.]: Peter Lang.
- Górska, K. (2015). Ni doblaje ni subtitulación: el gran éxito del voice-over en la televisión polaca. *Quaderns de cine*, n. 10, pp. 63-72. <<https://doi.org/10.14198/QdCINE.2015.10.07>>. [Accessed: 20220502].
- Luyken, G. H.; Herbst, T.; et al. (1991). *Overcoming Language Barriers in television: Dubbing and Subtitling for the European Audience*. Manchester: The European Institute for the Media.
- Matamala, A. (2008). Teaching voice-over translation: A practical approach. In: Díaz Cintas, J. (ed.). *The Didactics of Audiovisual Translation*. Amsterdam: John Benjamins, pp. 115-127. (Benjamins Translation Library; 77). <<https://doi.org/10.1075/btl.77.12mat>>. [Accessed: 20220502].
- Mayoral, R. (2001). El espectador y la traducción audiovisual. In: Chaume, F.; Agost, R. (eds.). *La traducción en los medios audiovisuales*. Castellón: Universitat Jaume I, pp. 33-48. <<http://dx.doi.org/10.6035/EstudisTraduccio.2001.7>>. [Accessed: 20220702].
- Orero, P. (2004). The pretended easiness of voice-over translation of TV interviews. *JoSTrans: The Journal of Specialised Translation*, n. 2, pp. 76-96. <https://www.jostrans.org/issue02/art_orero.pdf>. [Accessed: 20220502].
- Orero, P. (2005). La traducción de entrevistas para voice-over. In: Zabalbeascoa, P.; Santamaría Guinot, L.; Chaume, F. (eds.). *La traducción audiovisual: investigación, enseñanza y profesión*. Granada: Comares, pp. 213-222.
- Orero, P. (2006a). Synchronization in voice-over. In: Bravo, J. M. (ed.). *A New Spectrum of Translation Studies*. Valladolid: Publicaciones de la Universidad de Valladolid, pp. 255-264.
- Orero, P. (2006b). Voice-over: A case of hyper-reality. In: *Proceedings of the Marie Curie Euroconferences MuTra 'Audiovisual Translation Scenarios' –Saarbrücken 1-5 May 2006*.
- Orero, P. (2006c). Voice-over: The ugly duckling of audiovisual translation. In: *Proceedings of the Marie Curie Euroconferences MuTra 'Audiovisual Translation Scenarios' –Saarbrücken 1-5 May 2006*.
- Orero, P. (2009). Voice-over in audiovisual translation. In: Díaz Cintas, J.; Anderman, G. (eds.). *Audiovisual translation: language transfer on screen*. London: Palgrave Macmillan, pp. 130-139. <https://doi.org/10.1057/9780230234581_10>. [Accessed: 20220702].
- Rabadán, R.; Fernández Nistal, P. (2002). *La traducción inglés-español: fundamentos, herramientas, aplicaciones*. León: Universidad de León. <<http://hdl.handle.net/10612/5868>>. [Accessed: 20220702].

- Reynolds-Díaz, M.; Sizemore, M. T. (2000). Spanish voice-over: An effective option for producing educational videotapes for the hispanic elderly. *Educational Gerontology*, v. 26, n. 5, pp. 467-474.
<<https://doi.org/10.1080/03601270050111878>> [Accessed: 20220702].
- Rica Peromingo, J. P. (2016). Aspectos lingüísticos y técnicos de la traducción audiovisual (TAV). Bern [etc.]: Peter Lang.
- Rodríguez Fernández-Peña, A.C. (2020). Translating intonation and prosody in English-Spanish voice-over: strategies, resources and professional practices. (Unpublished PhD. Dissertation), University of Oviedo.
- Sánchez Mompean, S. (2015). Dubbing Animation into Spanish: Behind the Voices of Animated Characters. *JoSTrans: The Journal of Specialised Translation*, n. 23, pp. 270-291. <https://www.jostrans.org/issue23/art_sanchez.pdf>. [Accessed: 20220702].
- Sánchez Mompean, S. (2019). Dubbing attitudes through tonal patterns: when tones speak louder than words. *VIAL Vigo International Journal of Applied Linguistics*, n. 16, pp. 135-155. <<https://doi.org/10.35869/vial.v0i16.96>>. [Accessed: 20220702].
- Schwarz, B. (2011). Translation for dubbing and voice-over. In: Malmkjaer, K.; Windle, K. (eds.). *The Oxford handbook of translation studies*. Oxford: Oxford University Press, pp. 394-409. <<https://doi.org/10.1093/oxfordhb/9780199239306.013.0027>>. [Accessed: 20220702].
- Sepielak, K. (2013). Voice-over documentaries: synchronization techniques in English-Spanish and English-Polish translation. Unpublished MA Dissertation. The University of Texas Rio Grande Valley.
<https://scholarworks.utrgv.edu/cgi/viewcontent.cgi?article=1048&context=leg_etd>. [Accessed: 20220402].
- Sepielak, K. (2016). Synchronization techniques in multilingual fiction: Voiced-over films in Poland. *International Journal of Communication*, n. 10, pp. 1054-1073.
<<https://ijoc.org/index.php/ijoc/article/viewFile/3559/1578>> [Accessed: 20220402]
- Sinclair, J. (1996). Preliminary recommendations on corpus typology. EAGLES Document TCWG-CTYP/P <<https://www.passeidireto.com/arquivo/10767974/preliminary-recomendation-in-corpus-tipology-sinclair-john-mc-hardy>>. [Accessed: 20220402].
- Wozniak, M. (2012). Voice-over or Voice-in-between? Some Considerations about the Voice-over Translation of Feature Films on Polish Television. In; Remael, A.; Orero, P.; Carrol, M. (eds.). *Audiovisual Translation and Media Accessibility at the Crossroads*. Leiden, The Netherlands: Brill, pp. 207-228.
<https://doi.org/10.1163/9789401207812_013>. [Accessed: 20220402].