Universidad de Oviedo

Programa de Doctorado
en Investigaciones Humanísticas

LA TRADUCCIÓN DE LA ENTONACIÓN Y LA PROSODIA EN LAS VOCES SUPERPUESTAS (*VOICE-OVER INGLÉS-ESPAÑOL*): ESTRATEGIAS, RECURSOS Y PRÁCTICAS PROFESIONALES

TRANSLATING INTONATION AND PROSODY IN ENGLISH-SPANISH VOICE-OVER: STRATEGIES, RESOURCES AND PROFESSIONAL PRACTICES

TESIS DOCTORAL

Alfonso Carlos Rodríguez Fernández-Peña

Oviedo, 2020

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

RESUMEN (en español)

Este trabajo ofrece un estudio exhaustivo de dos campos que no han recibido la atención suficiente en el ámbito académico, pese a ser, muy relevantes cada uno en su disciplina. Nos referimos a las voces superpuestas (*voice-over* en inglés) y la traducción de la entonación y la prosodia del inglés al español en el ámbito audiovisual.

Las voces superpuestas o *voice-over*, constituyen un modo de TAV (traducción audiovisual) en auge, que en algunos ámbitos parecen estar desplazando a otros modos más consolidados y académicamente más estudiados como son el doblaje y la subtitulación. Existen hoy en día numerosos productos audiovisuales ofertados en televisión y plataformas de contenido a la carta VOD (*video on demand* por sus siglas en inglés), como Netflix o HBO, que se sirven de este modo de traducción para emitir sus contenidos, por ejemplo: documentales, programas de tele-realidad, entrevistas para programas informativos, y vídeos de empresa en internet. No obstante, y pese a su crecimiento exponencial y a ser calificado con frecuencia como el modo de TAV más fiel y realista de todos (más incluso que el doblaje y la subtitulación), las voces superpuestas parecen estar al margen y recibir menos atención académica que otros modos de TAV.

Por otro lado, la entonación y los rasgos prosódicos que ofrece la lengua inglesa son increíblemente ricos desde un punto de vista pragmático, algo que, sin duda, es sumamente relevante en el ámbito traductológico. Cabe citar, por ejemplo, los acentos regionales, la calidad de la voz, y las diferentes funciones que cumple la entonación inglesa para expresar la fuerza ilocucionaria del mensaje (la intención del hablante). Sin embargo, la aplicación del componente fonético y fonológico en la traducción audiovisual (inglés-español) apenas ha contado con los rasgos suprasegmentales (prosódicos), siendo los aspectos segmentales (los fonemas) los más estudiados y aplicados en materia de sincronización labial (o *lip-sync*) en el doblaje.

En esta tesis doctoral se pone en relación, por tanto, la traducción de la entonación y las voces superpuestas, que, al componer un tipo de texto eminentemente oral y espontáneo, constituyen un campo de pruebas idóneo para este tipo de estudio. Nuestra hipótesis parte de la premisa de que los traductores de voces superpuestas (inglés-español) no son del todo conscientes de las múltiples funciones comunicativas y pragmáticas que ofrece la entonación inglesa, ni de los recursos que ofrece el español –a menudo muy distintos de los de esa lengua– para reflejar tales funciones. Además, dadas las características técnicas de este modo de TAV, los límites que el texto audiovisual impone, y el número de profesionales que participan en el proceso de traducción, es muy posible que el mensaje que llega al público objetivo con

frecuencia haya perdido valor pragmático y por tanto no refleje fielmente el mensaje original.

Para llevar a cabo este trabajo se han seleccionado diversos programas (documentales, programas de tele-realidad, entrevistas para programas informativos, y vídeos web) que han sido transcritos en los idiomas origen y meta. Para su análisis, tanto cualitativo como cuantitativo, se marcaron los siguientes objetivos:

- Identificar cómo se traducen al español las características prosódicas y entonativas del inglés concernientes a la acentuación, el timbre, la tonalidad, la tonicidad, y el tono, determinando así si la fuerza prosódico-ilocucionaria implícita en el texto origen se traslada al texto meta o, por el contrario, se pierde.
- Identificar cómo se tratan los marcadores de discurso (*discourse markers*) en la traducción para voces superpuestas.
- Comprobar cómo las características relacionadas con la sincronía de los textos pueden afectar, o verse afectadas, como resultado del proceso de traducción.
- Analizar cómo se trata la cualidad y calidad de la voz en la traducción para voces superpuestas.
- Comprobar cómo los diversos agentes que participan en el proceso de traducción para voces superpuestas (traductores, locutores, ajustadores y directores de casting) conciben los rasgos prosódicos en dicho proceso.

Los resultados obtenidos del análisis, así como de una encuesta realizada entre profesionales del sector audiovisual, ponen de relieve el cumplimiento de nuestra hipótesis de partida; es decir, que el valor pragmático comunicativo de los rasgos prosódicos consustanciales a la lengua inglesa no es mayoritariamente trasladado al texto meta español por medio de los múltiples recursos que esta lengua ofrece. Se propone, por tanto, como conclusión al estudio, y dado el marcado carácter oral de los textos origen y meta implicados en el proceso de *voice-ove*r así como la supuesta autenticidad que esta modalidad de traducción ofrece frente a otras, contemplar, un nuevo tipo de sincronía junto a los demás tipos comúnmente asociados con las voces superpuestas, que se puede denominar *sincronía prosódico-ilocucionaria*, de tal modo que la intención que el hablante original en lengua inglesa a menudo expresa mediante la entonación sea reflejada en el texto meta mediante alguno de los recursos (sintácticos, semánticos, discursivos, entonativos, etc.) que la lengua española pone a disposición durante el proceso de traducción.

RESUMEN (en Inglés)

This dissertation offers a thorough study of two fields that have not received sufficient attention in the academic field, despite being, very relevant in their respective disciplines. We refer to voice-over translation and the rendition of intonation and prosody from English to Spanish in the audiovisual field.

Voice-over constitutes a booming AVT (audiovisual translation) mode, which, in some fields, seems to be displacing other more consolidated and more academically studied AVT modes such as dubbing and subtitling. In Spain, numerous audiovisual products are nowadays shown on TV and VOD (video on demand) platforms, such as Netflix or HBO by means of voice-over, for example: documentaries, reality shows, news interviews programmes, and company videos on the Internet. However, despite its exponential growth and being considered by some scholars as the most faithful and realistic AVT mode (even more so than dubbing and subtitling), voice-over seems to receive less attention in terms of academic research than other AVT modes.

On the other hand, the intonation and prosodic features offered by the English language are incredibly rich from a pragmatic point of view, something that is undoubtedly highly relevant in the translational field. Regional accents, voice quality, and the different functions that English

intonation performs in order to convey the illocutionary force of the message (the speaker's intention) are good examples of this. However, the application of the phonetic and phonological component in (English-Spanish) audiovisual translation has barely included suprasegmental (prosodic) traits, with the segmental aspects (the phonemes) being the most commonly studied and applied in terms of lip sync in dubbing.

Therefore, this doctoral dissertation combines the translation of intonation and voice-over, which, as an eminently oral and spontaneous text type, constitutes an ideal testing ground for this type of study. Our hypothesis is based on the premise that voice-over (English-Spanish) translators are not entirely aware of the multiple communicative and pragmatic functions offered by English intonation, nor of the resources offered by Spanish to reflect such functions – commonly very different from those of the English language. In addition, given the technical characteristics of this AVT mode, the limits that the audiovisual text imposes, and the number of professionals involved in the translation process, the message that reaches the target audience may frequently have lost pragmatic value and, therefore, not faithfully reflect the original message.

In order to carry out this task, various programmes (documentaries, reality shows, news interviews, and web videos) have been selected and transcribed in the source and target languages. We have conducted a qualitative and quantitative study in which a number of objectives have been set:
- To identify the way prosodic features related to stress, pitch, tonality, tonicity and tone are translated into Spanish, and determine whether the original prosodic-illocutionary force implicit in the source text has been rendered or lost in the final translation.
- To identify the way discourse markers are considered in voice-over translation.
- To check how synchronic features can affect or be affected as a result of the translation process.
- To check how voice quality is treated in voice-over translation.
- To check how voice-over professionals (translators, actors, adjusters, directors) conceive prosody and prosodic traits in the process of a voice-over production.

The results obtained from the analysis, as well as from a survey done by audiovisual professionals, highlight the fulfilment of our starting hypothesis; namely, that the pragmatic communicative value of the prosodic features inherent in the English language are, mostly, not transferred to the Spanish target text by means of the multiple resources that this language offers. Thus, to conclude our study, and given the marked oral nature of the source and target texts involved in a voice-over translation process as well as the assumed authenticity that this AVT mode offers as compared to others, we propose a new type of synchronic feature to be considered with the other types commonly associated with voice-over: *prosodic-illocutionary synchrony*. This way, the original speaker's intention which is in English frequently expressed through intonational means may be reflected in the target text by one of the resources (syntactic, semantic, discursive, intonational, etc.) that the Spanish language offers during the translation process.

SR. PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO EN

Acknowledgements

This work would not have been possible without the help of numerous people, and, although I am sure I will inevitably forget some of them here, I hope they forgive me.

First, I must thank my supervisor and thesis director, Prof. Marta Mateo, for her support and guidance during this journey. Without her dedication, time, and invaluable feedback, this dissertation would not have been possible. It has been an honour to work with you.

I am truly grateful to the University of Angers for giving me the chance to use their facilities during 3 months in the early research stages of this project, to the staff of the Modern Languages Department and the 3L.AM Research Group, specially to Dr Sandra Contamina and Prof. Manuelle Peloille.

I am extremely obliged to Prof. Jorge Díaz-Cintas and Dr Richard Mansell for reviewing and assessing my work with such short notice and for doing so during the Covid-19 crisis.

I am indebted to all the voice-over professionals who answered the questionnaires. A big big thank you to: Mapi Lucchesi, Ana Ávila, Joanna Shiokka, Marianne Dubois, Christian Tillmanns, Krisitana Ivanova, Koen Smit-Amor, Poullet Sangbaran, Margaret Davies, Marta González, Bernard Roger, Anna Sticken, Brenda Viviani, Thomas Dormoy, Constantino de Miguel, Thierry Wieder, Davide Vocale, David Svarcer, Sonja Chiphorost, Annalisa Cazzola, Ricky Salmon, Paul Bridge, Elena Fernández, Marina Mollá, Alexander Rodríguez, Ana Viñuela, Federico Louhau, Vicky Tessio, Joel Valverde, Antonio Alfonso, Joanna Rubio, Elena Silva, Megui Cabrera, David García, Mari Luz Rodrigo, Nadia García y Rafael López.

Javier Fernández-Peña deserves a special place here for many reasons. He introduced me to voice-over back in 2006 and has always helped me in many different ways. Tío Javi, without your help, advice and love this would not have happened. Gracias por tanto.

Special thanks to the UCL (University College London) librarian and cataloguer Paul Amies for his help at the early stages of my research during the 2014 CenTras summer course.

List of examples in order within the analysis chapter

List of tables

List of figures

List of images

List of abbreviations

AD	Audio description
ADR	Audio Dialogue Replacement
AV	Audiovisual
AVT	Audiovisual Translation
Br En	British English
C.S. Ex	Cleft Sentence Example
D. Ex	Deaccenting Example
DTT	Digital Terrestrial Television
F.T. Ex	Falling Tone Example
Gen Am	General American
Hz	Hertz
I.S. Ex	Intonational Subordination Example
IP	Intonational Phrase
IVR	Interactive Voice Response
K.S. Ex	Kinetic Sync Example
L.S. Ex	Literal Synchrony Example
LED	Light Emitting Diode
MM	Multimedia
N.F. Ex	Narrow Focus Example
NDA	Non Disclosure Agreement
OED	Oxford English Dictionary
Q. Ex	Quotation Example
R. Ex	Reaccenting Example
R.T Ex	Rising Tone Example
RAE	Real Academia de la lengua Española
RP	Received Pronunciation
S.P.D. Ex	Subject-Predicate Division Example
SDH	Subtitling for the deaf and hard of hearing
SEC.	Seconds
ST	Source Text
T.Q. Ex	Tag Question Example

Top. Ex	Topic Example
TT	Target Text
VOD	Video On Demand
VOIP	Voice Over Internet Protocol

Introduction

Voice-over has become one of the dominant translation modes in recent years in the Spanish audiovisual market. We find it in documentaries, reality shows, internet videos and even in filmed interviews broadcast on TV channels as diverse as news companies like Euronews or the BBC, and sports networks like Eurosport or Teledeporte. In addition to this, the online AVT market has boomed in voice-over translation companies, which offer their translation solutions[1], from script translating to audio recording and mixing, taking advantage of the globalisation opportunities that the Internet offers nowadays. This `ugly duckling´ of audiovisual translation modes, as it was once considered (Orero 2006c), seems to be the favourite translation method for most of the interactive audiovisual unscripted content - although there are exceptions, thanks to its cost-cutting and cost-effective nature, given that it takes less studio and script translation time since there is no need for lip synchronisation and phonetic adaptation. In addition to this, voice-over is regarded as a faithful and authentic representation of the original text (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. This gives voice-over translation a feel of realism and faithfulness, that other AVT modes, such as dubbing or subtitling, do not get to transmit.

Given the assumed faithfulness that, *a priori*, characterises voice-over translation, and the fact that most of the content that is translated using this AVT mode is unscripted, resulting from spontaneous speech, and given the important role that intonation (or the suprasegmental features of language) play in this type of speech, we have decided to analyse how English prosodic and intonational features and their pragmatic essence are conveyed in Spanish voiced-over programmes.

English is a language characterised by a rich intonational nature. In fact, it resorts to prosodic phenomena, such as stress, rhythm, pitch, and tonality, tonicity and tone, to convey meaning in a way that Spanish does not (García-Lecumberri 1995: 181). On the other hand, Spanish makes great use of word order and semantic devices to compensate for its lack of

---

[1] These online companies offer an "all in one" solution. They hire freelance translators, voice actors, video editors and sound engineers to meet the audiovisual needs of any company, from dubbing and subtitling to voice-over.

intonational flexibility in order to deliver pragmatic meaning. Thus, given the fact that English voice-over texts are purely oral, rich in prosodic and intonational markers, and that English prosody has received "sporadic and very partial attention from translation researchers and practitioners" (Mateo 2014: 113) probably because there is not a special module on phonetics in most translation studies university syllabuses, as seen in Cerezo Merchán's PhD thesis (2012), we consider that it is possible that audiovisual translators (in particular English-Spanish voice-over translators) are unaware of the ample spectrum of pragmatic intonational mechanisms that English uses and the different means that Spanish often resorts to. The possibility that these prosodic traits may not have been satisfactorily identified and, therefore, rendered conveying an equivalent illocutionary force in the Spanish target text by means of the strategies offered by this language (syntactic, semantic, etc.), may have resulted in second-rate translations (Mateo Martínez 1996: 85), since the speaker's underlying intention will probably not have reached the target audience.

In this PhD dissertation, we have decided to analyse how English prosodic elements (those pertaining to the realm of intonation, and also to the human voice) are rendered in Spanish voice-over texts. The materials analysed are English-Spanish voice-over texts, ranging from documentaries and reality TV shows, to web videos and news programmes. We have conducted a qualitative and quantitative analysis of these texts following the translemic analysis model proposed by Rabadán (1992: 208), which is a variation of De Beaugrande & Dressler's model (1981). Thus, we have firstly analysed the English source texts looking for all the relevant prosodic features conveying illocutionary meaning. Secondly, we evaluated the Spanish target texts concerning those segments, so as to identify the strategies used in order to render the original illocutionary force, and to assess whether this rendition was achieved or not. In addition to this, a series of questionnaires have been provided to voice-over professionals, in order to understand the role that prosody plays in the voice-over translation process; the questions raised cover a range of topics, from how translators deal with intonation and translation, to the way voice actors and casting directors consider the use of the voice in their everyday job.

A set of objectives has been established in order to achieve our goal in this work:

a) To identify the way prosodic features related to stress, pitch, the 3Ts (Wells 2006: 06) (tonality, tonicity and tone) are translated into Spanish, and whether the original

prosodic-illocutionary force implicit in the source text have been rendered or lost in the final translation.

b) To identify the way discourse markers are considered in voice-over translation.

c) To check how synchronic features can affect or be affected as a result of the translation process.

d) To check how voice quality is treated in voice-over translation.

e) To check how voice-over professionals (translators, actors, adjusters, directors) conceive prosody and prosodic traits in the process of a voice-over production.

This dissertation has been structured in 6 chapters, and 7 annexes in digital format, which include the audiovisual materials analysed, the spreadsheets used to conduct the analysis, the examples analysed (including audio files and pitch contour images), the surveys answered by voice-over professionals, a video interview to a renown Spanish voice actor, and some email conversations with TV networks that offer voice-over content.

Chapter 1 provides a description of the audiovisual text and the different AVT modes considering the most relevant academic definitions and characteristics for each of them. Voice-over has a more in-depth portrayal within this chapter, in which we discuss the academic understanding of voice-over within the disciplines of Film Studies and Translation Studies, the blurriness and terminological chaos surrounding this term (*voice-over*) and how it differs from what the audiovisual industry understands as voice-over. In addition, the linguistic and technical features that make voice-over a transfer mode half-way between dubbing and interpreting are explained in detail.

Chapter 2 covers the relevance that English phonetics and phonology has in translation in general and AVT in particular. It provides a journey through the form and functions of English intonation, the prosodic traits that characterise English native speakers (socially and culturally) through their voice, and how these apply to convey pragmatic meaning in oral speech..

Chapter 3 focuses on voice-over translation in relation to English prosody is handled for Spanish target texts. Here we will describe the voice-over translation process in detail, and see the differences between translating voice-over for production and for post-production and the way these two forms can affect the effective rendition of prosodic-illocutionary elements.

We will deal with the issue of faithfulness and how it can be applied to a translation mode that is known for its authenticity and reliability, taking into consideration the restrictions imposed by the nature of the audiovisual text, the two voice-over translation possibilities (for production and post-production) and the different categories or sub-genres that can be found within voiced-over products. Finally, we will see how intonation can be rendered in voice-over during the delivery of the target text, taking dubbing as an example.

Chapter 4 offers a description of the audiovisual texts that make up our corpus of study, the documentation presented, including the transcripts of the texts, the spreadsheets with the analysis, the questionnaires answered by voice-over professionals, and other relevant information (interviews and emails with voice actors and TV networks). The hypothesis of the thesis is set at this point. We assume that English-Spanish voice-over translators are not fully aware of the communicative relevance that English prosody has and the strategies offered by the Spanish language to successfully render the illocutionary force implicit in them. To do so, the methodology chosen to conduct our analysis is based on Rabadán's translemic analysis model (1992: 208), with which we will try to reach a series of objectives, aiming to provide an answer to our hypothesis.

Chapter 5 includes the analysis and results (both qualitative and quantitative) of the corpus. It is divided into seven sections: stress, pitch, tonality, tonicity, tone, synchrony, and discourse markers. We will examine the pragmatic meaning of the English intonational variables, and will assess the success rate and form of transfer in the Spanish texts. The technical aspects relating to synchrony will be analysed taking as examples those segments which are out of sync so that we can check whether synchrony influences the rendition success of the target text.

Chapter 6 includes the conclusions of this study, summarising the results of the analysis section and answering the objectives set in chapter 4. Based on those results, we will answer the initial hypothesis and see whether voice-over professionals (translators, actors and directors) working in Spanish are fully aware of the importance that English prosody has to produce illocutionary force in oral speech and how successful they are in conveying an equivalent meaning in the target version of the programmes. To conclude the dissertation, a final remark on the future of AVT in general, and voice-over in particular, will be presented, based on my experience as a professional voice-over translator and voice actor for over a

decade in the audiovisual industry, and taking into consideration the recent technological changes which constitute a breakthrough and will inevitably shape the industry in the next decades, making AVT as we know it today a thing of the past.

1 AUDIOVISUAL MATTERS

1.1 The audiovisual text

Translation has played a major role in improving and facilitating communication and knowledge among cultures and civilisations through history, and it keeps playing a similar role in the 21st century whatever the format of the script or source text. If hundreds, and even thousands, of years ago the most common form source text was a tablet made of pottery or even stone, like the Rosetta stone, some kind of papyrus or leather sheet, nowadays texts for translation are mainly soft documents which can be translated, adapted, and revised on a computer. In fact, today, most professional translations are done using software like SDL Trados, MemoQ, Memsource, or Babylon, to cite a few, which facilitate the translation process with the help of cloud-based dictionaries, glossaries, translation memories and machine translation tools, which not only ease and speed up the translation process, but make it smarter and more accessible.

This doctoral thesis has as its core subject of study the translation of voice-over; therefore, this first chapter will be devoted to identifying and defining what voice-over texts are, their characteristics and how they fit in the text-type frame of translation studies.

Traditionally, translation has been studied from a linguistic, or literary, perspective and commonly as a branch of applied linguistics (Mayoral, Kelly, and Gallardo 1988), its major problem being "finding words in a target language which render the meanings stated in the source language" (1988: 356). However, this has radically changed since the introduction of audiovisual elements in 20th century communication. Audiovisual and multi-media formats like those TV spots, films, musicals, documentaries, news reports, video-games etc. do have a verbal text, or script, as in the same way as literary documents, but they differ in the fact that they have some *"space, time and language constraints that make them a group apart in the field of translation"* (Comitre 1997: 89). These time and space determining factors are, for instance, the length which a sentence takes to be uttered and the scene where it takes place in a film. Voice-over texts are part of this audiovisual text group, so we will now observe what characterises these texts.

Audiovisual texts are distinguished for communicating information by means of two channels: visual and acoustic (Sokoli 2005, Comitre 1997, Chaume 2004, Zabalbeascoa 2008). The visual information concerns everything the viewers can see, like images in a film, photos in a Power Point presentation, or subtitles; whereas the acoustic channel concerns what the viewers can hear, such as words in the form of dialogues or monologues in films, video-games, or news reports, and also the music and sound effects that enhance films, video-games and any kind of motion picture product.

According to Sokoli (2005: 182-3) there are five features that characterise and define the audiovisual text (my translation):

1- Reception via two channels: acoustic and visual
2- Significant presence of non-verbal elements
3- Synchronisation between verbal and non-verbal elements
4- Transfer on screen – Reproducibility
5- Predetermined sequence of images and sounds – Recorded material

Zabalbeascoa (2008: 21) and Sokoli (2005: 182) illustrate their concepts of the audiovisual text with the following charts, which show the relationship and synchronisation amongst its different elements:

	Audio	**Visual**
Verbal	Words heard	Words read
Non-Verbal	Music + Special effects	The picture Photography

Figure 1 The four components of the audiovisual text for Zabalbeascoa

Acoustic - verbal (i.e. sentences)	⟷	Visual – verbal (i.e. titles)
↕	⟷	↕
Acoustic – non verbal (i.e. sounds, music)	⟷	Visual – non verbal (i.e. images)
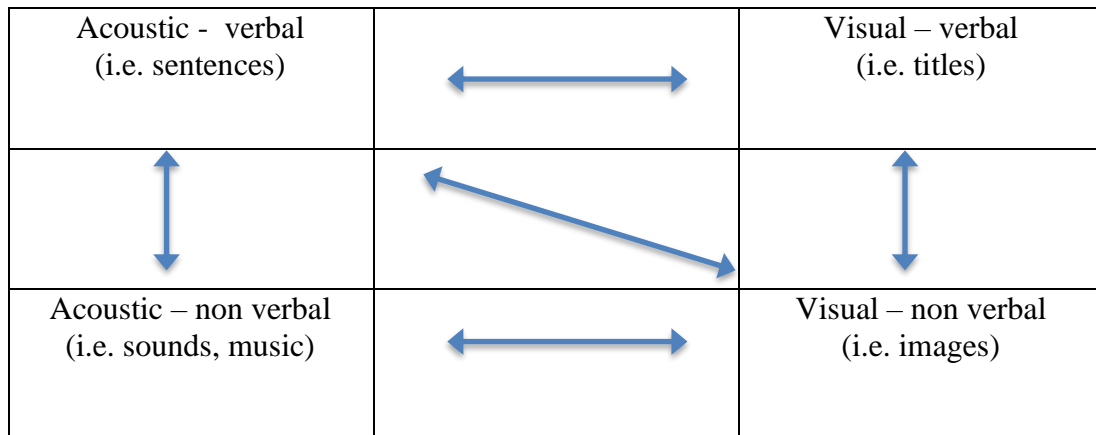
Figure 2 The synchronization among the elements of the audiovisual text for Sokoli

Audiovisual texts (mostly cinematic) have also been labelled as complex-mode texts (Rabadán 1991: 100) given that these types of texts demand a triple translation (sound, graphic, and sound). This applies only to dubbing and subtitling, but not to voice-over or any other audiovisual translation mode, surprisingly. The chart below shows the different kinds of texts for the different media (sound, spelling, and picture), and their mode (primary[2] and complex[3]) (Rabadán 1991: 101).

Medium / Mode	SOUND	SPELLING	IMAGE
PRIMARY MODE	Immediate oral texts	Written texts	Iconic Texts
COMPLEX MODE	Oral Mediate reception texts	Visual mediate reception texts	
	Cinematic texts		

Figure 3 Text types according to the mode and media for Rabadán (1991:101)

Rabadán, like Chaume, Zabalbeascoa, and Sokoli above, identifies the different elements that form the audiovisual text since cinematic texts fill the space for the three media shown in the headers in her table: sound, spelling and image. Nonetheless, it is worth considering her views on one type of primary mode texts: immediate oral texts. For Rabadán, one basic characteristic of the [primary mode] immediate oral text is *spontaneity* (1991: 100)

[2] The primary mode refers to immediate oral texts (*texto oral inmediato*) Rabadán 1991:100
[3] The complex mode refers to cinemtic textx (textos cinematográficos) Rabadán 1991:100

since the frame of negotiability, citing House 1986, is public and direct in the field of translation. These oral texts belong to what is known as *interpreting* and are subject to different strategies and methods from those usually required by written texts (as in dubbing and subtitling), which lack the spontaneity feature and thus follow a linguistic pattern of behaviour which is thought out in advance in a private negotiability frame; that is, the primary mode texts lack the [+ spontaneity] label. The oral features of the texts to be dramatized, such as theatre plays or dubbing scripts, are fake signs of orality despite the fact that these texts include all the characteristics of a spontaneous conversation. Dubbing texts are not spontaneous in spite of creating the illusion of immediate linguistic acting (1991: 101). The idea of orality and spontaneity in AVT, especially in voice-over and dubbing, will be developed in section 2.2 in the following chapter

With all this in mind, it is worth considering how voice-over texts fit in the above text-type divisions. As will be seen in section 1.3, voice-over differs from the other audiovisual modes precisely in the fact that its source text is spontaneous, despite some exceptions, and is commonly used to translate what is known as "unscripted programmes", that is, programmes without a script, which develop the plot spontaneously, such as *Master Chef*, *Lock Up* or *The Ultimate Survivor*. For this reason, on the one hand, voice-over texts, fits perfectly in the descriptions and charts provided by Zabalbeascoa (2008) and Sokoli (2005), since they require synchronisation of text, image and sound; on the other hand, in Rabadán's chart, a new box would be needed to include voice-over texts, given their immediate oral component, thus providing a broader representation of cinematic, texts, as I propose in the chart below:

Medium / Mode	SOUND	SPELLING	IMAGE
PRIMARY MODE	Immediate oral texts	Written texts	Iconic Texts
COMPLEX MODE	Oral Mediate reception texts	Visual mediate reception texts	
	Cinematic texts (dubbing & subtitling)		
	Oral Mediate & Immediate reception texts	Visual mediate reception texts	
	Cinematic texts (Voice-Over)		

Figure 4 Text types according to the mode and media for Rabadán (1991:101) including voice-over texts

Another relevant aspect when dealing with audiovisual texts is to consider whether these are going to be read aloud as in dubbing and voice-over, or silently as in the case of subtitling, since there are different stylistic norms that apply in each reception mode, and there is also a "further complication in the translation process when the text is dramatized." (Mayoral, Kelly and Gallardo 1988: 357). Translators have to take into consideration the verbal elements (the written text) and the image accompanying the text so that both components are in synchrony and there is no mismatch between the words e.g. uttered by the voice actors, or the subtitled lines, and the images shown on the screen. In addition to this, the dramatization provided by an actor in the translation has to be in synchrony with the original performance, as will be seen in section 1.2 in more detail.

Audiovisual texts are present in several genres (Agost 2001: 234) and can include all sorts of registers and dialects, such as the languages of specialisation in documentaries, colloquialisms, dialects and jargons in films, and a polite and formal language in news reports. This means that no audiovisual text is the same as there is great variation in audiovisual texts, which implies that the translator has to be aware of how to convey these linguistic features and the specificity of each genre. The oral features proper of voice-over texts, such as voice quality, gender, intonation, and accent pose a real challenge to the whole translation process, which not only applies to translators themselves, but also to voice talents, adjusters, casting directors, linguistic directors, and all the people involved in the process, which may vary according to the budget of the project as will be studied in chapter 3.

Another issue concerning audiovisual texts is the fact that new formats and genres may be generated (Agost 2001: 235), which is precisely what is intended by corporate enterprises, given the interest they create and the fact that they "tend to catch the eye of the audience". Agost's words have been proved right as new audiovisual and multi-media formats have appeared since 2001. Social networks like Facebook, Snap Chat, Twitter, Vine etc. are used by millions of people and companies willing to access and contact the general public for commercial purposes. These companies, and users, translate their messages into multiple languages thus creating new formats and genres, as can be the case with Twitter, for example,

whose limit on the number of characters per tweet (140 before, now 280) makes the translating act in this platform more challenging.
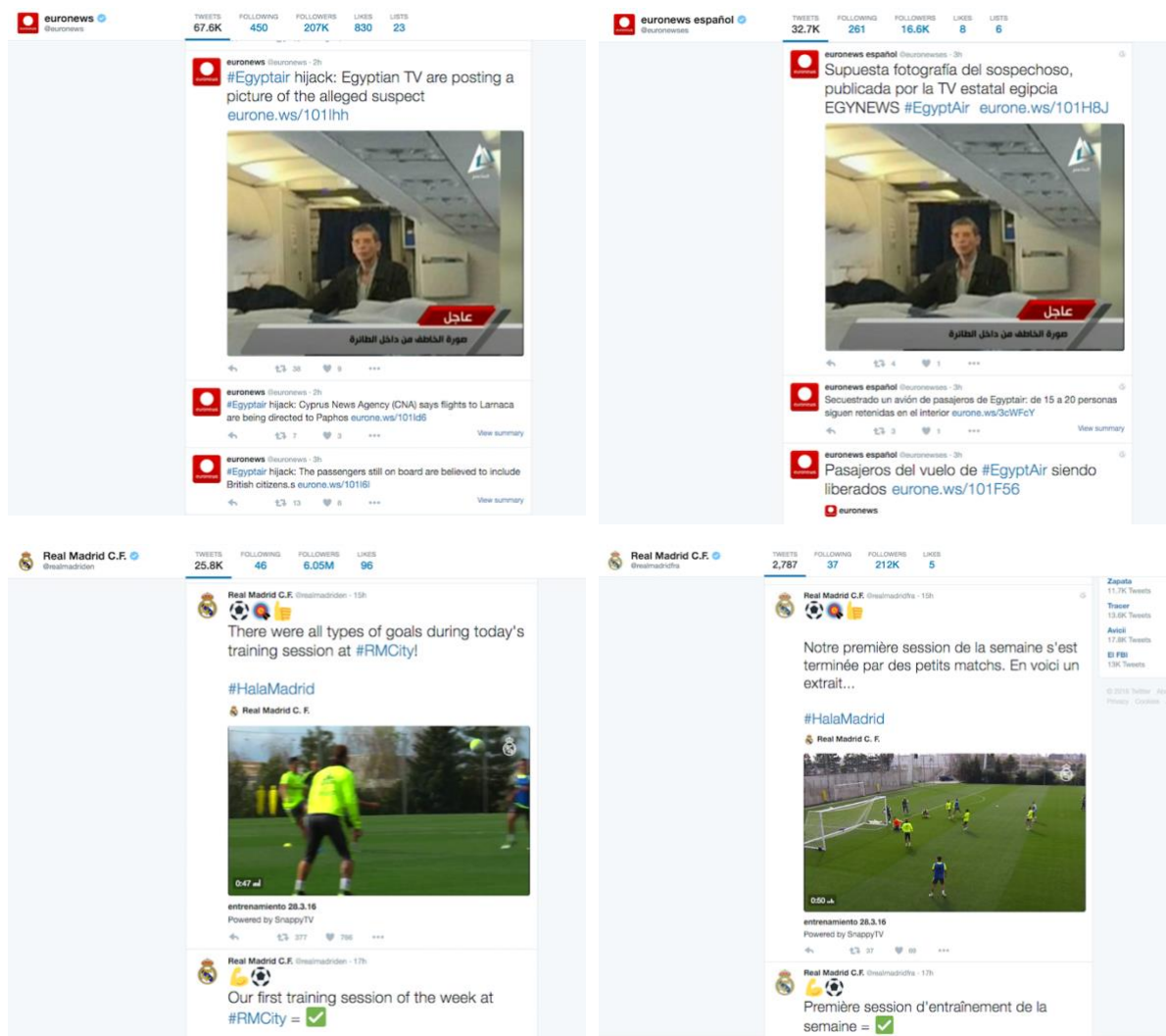
Figure 5 Translation of Tweets from the Euronews and Real Madrid's Twitter accounts

Once the characteristics of the audiovisual text have been considered, and the special oral spontaneous feature of voice-over texts acknowledged, let us close this first chapter with a reflexion on the complexity of translating these types of texts, bringing back the idea of "constrained translation" developed by Mayoral et al. (1998). The translator's role is complicated and constrained when translation not only implies a verbal text but texts which are in association with other communication media, like images, music, acoustic non-verbal sources, etc. For Mayoral et al. (1988), the fact of only considering the linguistic element of texts in the translation process, leaving behind those aspects which are specific of translation as a communication process, i.e., "those which depend on the relation of the linguistic message

to other messages conveyed by non-linguistic features" (1988: 356) provides a limited-scoped translation. Therefore, translators have to be aware of this complex relationship between the different elements in an audiovisual text and include and adapt the target features of the text to the extra-linguistic elements of the source text so there is a synchronicity among text, image and sound in the target product too.

The concept of "constrained translation" and how audiovisual texts require a special approach to translation will be dealt with in more detail in the following chapter, which will be devoted to the audiovisual and multi-media translation types. A brief tour around the different methods and their characteristics is provided with some examples. Voice-over translation will have its own chapter just after, in which we will see in detail the multiple characteristics of this translation mode from an academic and AVT industry perspective. In this work, the focus of attention will be an essentially linguistic element, i.e. intonation, and the way it can affect, enhance, or even spoil a translation. But before that, let us now delve into the audiovisual and multi-media translation world.

1.2 Audiovisual and multimedia translation

This chapter will provide an overview of the different audiovisual and multi-media translation modes in order to place the voice-over method in the translation context. But we will first analyse the differences, if there are any, between "multi-media" and "audiovisual" translation; a debate that has been puzzling academics and professionals alike.

Audiovisual translation and multimedia translation are sometimes understood as the same field, or type of activity. However, opinions differ as to whether these terms refer to the same activities. Chiaro (2009: 141) considers that *audiovisual* translation, together with *media* translation, *multimedia* translation, *multimodal* translation, and *screen* translation, are overlapping umbrella terms, which cover "the interlingual transfer of verbal language when it is transmitted and accessed both visually and acoustically, but not necessarily, through some kind of electronic device." However, this scholar considers that screen translation (translation of films and other products for the cinema, TV, video, etc.) is both audiovisual and multimedia *in nature* (2009: 141). This is true considering that nowadays most films and tv series, whether

dubbed, subtitled, voiced-over, etc., can be played in different (multiple) media. In line with this, Remael (2001: 13) understands that audiovisual translation for film or television (dubbing, subtitling, voice-over etc.) could also be seen to involve multimedia activities, although the term *multimedia* could also make reference to the wide range of hi-tech products of our time, like interactive TV, virtual reality, digital video, etc. Both Remael and Chiaro refer to the fact that there is a technological component that applies to multi-media products; and consistent with this view is Mayoral's (2001: 34) for he considers multimedia products as audiovisual and computerised at the same time. The current technological development with smartphones, tablets, smart watches, and smart TVs allows users to interact, or communicate, with the device. This interactivity between user and machine is what from our perspective, makes a product *multi-media* rather than just *audiovisual*. Those new technologies, allow users to, for example, choose the language in which they want to watch a programme (movie, documentary, series, etc.), put subtitles in its original language or a translated version, or resort even to audio description. This applies not only to audiovisual products like movies and TV series, but also to video games (which are audiovisual too but belong to the realm of localisation), and audio books and audio guides[4], which can be played in different languages and devices, thus also allowing interactivity, and can also be audiovisual products - although sometimes they only include the audio component without the visual element. For Cattrysse (2001: 01-02), multimedia translation has been understood (like media translation) as the verbal translation of the linguistic part of multimedia messages. But this scholar considers this problematic since it isolates the linguistic part from the rest of the multimedia message, so he prefers to define multimedia translation as "the translation of a message into an AV (audiovisual) or MM (multimedia) message" (2001: 02). This definition is regarded as more advantageous because it allows the integration of linguistic translational aspects into the global AV or MM communication.

Cattrysse (2001: 01) refers to the interactive process inherent in *multimedia*, which this author defines in two ways: in a *usual and narrow sense* when there is a process and a presentation of text, graphics and pictures, if not animation and motion video, which generally requires *interactivity*; and in a *wide sense,* when the process and presentation of information is done in two or more media simultaneously.

[4] This is the case of audio guide apps like MyWoWo (https://mywowo.net/en) and BMW Museum (https://itunes.apple.com/de/app/bmw-museum/id1079980613?mt=8)

In our view, multimedia products are those audiovisual products which are in digital format (.mov, .wmv, .mp4, etc.), not analogue, and allow some degree of interactivity between the viewer and the device on which it is being reproduced, like the possibility to choose the language, subtitles, audio description, the chance to watch extra footage, commentaries by the film director, cast, etc. This is the case of most movies, series, documentaries, etc. available on VOD platforms, such as Netflix and Amazon (Prime), DVDs, video games, etc. Multimedia products are then audiovisual ones with an interactive digital component. However, audiovisual products are not multimedia necessarily, since they could not have the interactive component, or the digital format proper of multimedia ones. This can be the case of the videotape VHS films before the arrival of the DVD, or even digital videos that do not allow any kind of interactivity, like subtitling or language options.

There are various modes of audiovisual translation, some of them multimedia, being the most common dubbing and subtitling, followed by voice-over, interpreting, surtitling, audio description, fansubbing, and fundubbing. A brief description these modalities of translation together with some examples will be given hereunder.

1.2.1 **Dubbing**

This is certainly the most widely used method of AVT in Spain (Rica Peromingo 2016: 73 & Sánchez Mompeán 2017: 20) and probably the preferred one for the general public in other countries considering that most audiovisual products translated into Spanish and broadcast on TV, or VOD (video on demand) platforms such as Netflix, Amazon, and HBO are dubbed.

As defined by Chaume (2004: 32), dubbing consists in translating and adjusting an audiovisual script and the subsequent interpretation of this translation by actors under the direction of a dubbing director and the advice of a linguistic adviser. In other words, it consists in replacing the original voices in a film with others in the translated language, in such a way that there is "no difference" between them, that is, that the viewers of the translated version have the impression that they are watching an original and not a translated film. In order to achieve this level of translation accuracy, a good adjusting of the script is paramount, since lip-synchronisation should be as close as possible to the original film. However, the perfect lip-

sync is not possible, as there will always be words in each of the languages involved in the process that have different phonetic realisations; this means that for the target text and the dubbing performance to be credible, changing the target text is sometimes necessary.

The examples shown below belong to two scenes from two movies with extreme close-ups. These kinds of frame imply a challenge for translators and adjusters, who have to find a target text which closely matches the lip movements of the actors on screen.
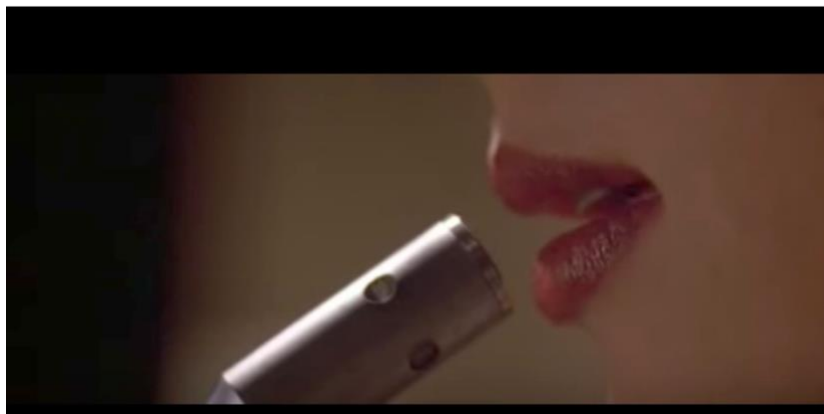
Image 1. Extreme close-up to Mia's (Uma Thurman) lips in Pulp Fiction (1994)

Example 1: *Pulp Fiction*

- Original: *Go make yourself a drink and I'll be down in two shakes of a lamb's tail*.
- Spanish Dubbing Translation: *Sírvete una copa y bajaré en lo que tardas en bebértela*.

Image 2. Extreme close-up to the Terminator's (Arnold Schwarzenegger) lips in *The Terminator* (1984)

Example 2: *The Terminator*

- Original: *I'll be back.*
- Spanish Translation: *Volveré.*

These examples show the relevance of finding words with similar phonetic realisations in both languages. In the case of *The Terminator*, it is clear that there is a match in the number of syllables (3) – *I'll-be-back | Vol-ve-ré* and the two bilabials in "be back" and "volveré" (in Spanish "v" is a voiced bilabial plosive, compared to English "v" which is a voiced labiodental fricative). In the example of *Pulp Fiction*, the lip-sync match is not so close as that in *the Terminator*'s but the fact of watching just Uma Thurman's mouth profile is convenient in cases like this.

There are other important technical features characteristic of dubbing apart from lip-sync, such as isochrony and kinetic synchrony. These features are also present in voice-over, so they will be described in detail further on in this work, in the section devoted to voice-over (section 1.3). Other synchronic features shared by dubbing and voice-over are: character synchrony and content synchrony. These two, according to Chaume (2012: 69), do not strictly fall within the definition of synchronization, but will also be described in section 1.3 since they will be relevant in our study.

The dubbing process is a very complex one, with several agents involved, though the number may vary according to the budget of the project, ranging from dubbing actors, translators, adjusters and dubbing directors, to sound engineers, linguistic advisors and voice extras for ADRs (Audio Dialogue Replacement), for example.

1.2.2 Subtitling

Subtitling consists in providing a written translation of film dialogues, together with the discursive elements that appear in the image (letters, inserts, graffiti, etc.) and the information that is contained on the soundtrack (songs, voices off) (Díaz Cintas and Remael 2007: 08), which are shown in the lower part of the screen, as the film is played. These *snippets of text* (Pérez-González 2014: 15) are usually displayed centred on the screen, in white colour, and

have a maximum of two lines and a limit of 37 characters each (Rica Peromingo 2016: 97). Although the figures may vary, according to Díaz Cintas and Remael (2007: 09) a subtitle has some 32 to 41 characters per line. The duration of the subtitles on the screen may vary: for monolinear subtitles the maximum duration is 4 seconds, while bilinear ones may last 6 seconds (Rica Peromingo 2016). They should all be displayed up to one-quarter of a second after the corresponding character starts talking so that the audience has time to identify who is speaking (Gambier 2006: 261). All subtitled programmes are composed of three major components: the spoken word, the image and the subtitles (Díaz Cintas and Remael (2007: 09), which together with the viewer's ability to read the images and the subtitles, and the size of the screen, determine the basic characteristics of this medium.

Together with dubbing, subtitling is one of the most widely studied and recognised translation modes, and it is now also widely accessible thanks to DTT (Digital Terrestrial Television), even in essentially dubbing countries like Spain. TV viewers can opt for subtitling just pressing a button in their remote controls, which not only contributes to the general growth of this AVT translation mode, since almost all TV programmes have subtitling possibilities, but also promotes the reception of TV content in its original version.

Like dubbing, the process of subtitling implies certain constraints, which also makes it a rather challenging translation method. The first issue relates to recreating in writing the oral elements inherent in the original text, like intonation, body language, etc. (Hurtado 2001: 80). The second involves time and space restrictions, which force translators to "frequently condense, if not reduce, the original information in the dialogue" (Sepielak 2013: 11). This will also be the case in voice-over. The impossibility of transcribing all the textual elements of the script, together with the *re-condensation* and selection of the material to be subtitled, result in what has been defined as *mistranslation* (Rabadán 1991: 104). On the other hand, some scholars, like Rica Peromingo (2016: 94), consider that, in order to provide good subtitling, it is not really necessary to translate all the elements of the original text, since the oral register on the screen uses many linguistic elements which can be considered redundant in the written caption, and Rica Peromingo (2016) highlights that the quality of subtitles should not be measured in terms of how much text is reduced as compared to the original one, but in how the target audience perceive the product: the audience cannot have the impression that they are not being given all the information needed to understand the audiovisual product.

Subtitles can be classified in three categories according to their linguistic parameters, as shown in the table below.

Intralingual subtitles	For the dead and the hard-of-hearing (SDH)
	For language learning purposes
	For Karaoke effect
	For dialects of the same language
	For notices and announcements
Interlingual subtitles	For hearers
	For the deaf and hard-of-hearing (SDH)
Bilingual subtitles	

Table 1. Classification of subtitles according to their linguistic parameters.

The first category shown in the figure above (intralingual subtitles) carries some sort of reluctance by some scholars to call it translation since it involves a shift from oral to written within the same language (Díaz Cintas and Remael 2007: 14). Within this category, subtitling for the deaf and hard-of-hearing (SDH) will have its own chapter in this work since it has gained tremendous popularity in the last decade, and there is also a norm in Spain which regulates the normalization of this translation mode for accessibility purposes. The second group of intralingual subtitles refers to those specifically designed for teaching and learning a foreign language. The third group is the one used to sing karaoke. The fourth group is the use of subtitles in films (TV programmes, movies, documentaries, etc.) in which there are people whose accents or dialects are hard to understand for the audience, who share the same language. An example of this is the case of Noel Gallagher, former guitar player and composer of the pop-rock band Oasis, who was subtitled on the MTV[5] interviews due to his Mancunian accent. The last group can be seen on screen in underground stations and other public areas for advertising and broadcast the latest news (Díaz Cintas and Remael 2007: 17).

[5] https://www.rollingstone.com/music/music-features/the-last-word-noel-gallagher-on-oasis-rumors-manchester-pride-127547/

The second category (interlingual subtitles) implies the translation from a source language into a target one, and is also known as *diagonal subtitling* (Gottlieb 1994), as stated by Díaz Cintas and Remael (2007: 17) because it involves changing not only the language (source into target) but also the mode, from oral to written. The first type of interlingual subtitling, for hearers, is the most common and is present in almost every multimedia movie, or TV programme. Children's programmes do not necessarily need subtitling for obvious reasons, although it is now common to have subtitling options in these kinds of programmes, as shown in image 4 below, which shows a frame of a TV programme whose target audience are toddlers and preschoolers. SDH can also be found for interlingual purposes, although it is not very common, and is gaining visibility thanks to DVD and pressure groups in countries such as Germany, the UK and Italy, which have managed to get many foreign films marketed in their countries with two separate tracks, one for the hearing audience, and one for the deaf and hard-of hearing (Díaz Cintas and Remael 2007: 18).

The final, and third, category (bilingual subtitles) is used in countries and regions where two languages are spoken, as can be the case of Belgium with the Walloon and Flemish communities (Díaz Cintas and Remael 2007: 18), or Spain with the Catalan, Basque, and Galician communities.

The following example show a graphic illustration of subtitles (for hearers) in a movie.

Image 3. A Subtitled frame from the TV Series *Stranger Things* (2016)

- Original dialogue: *I've told you a million times that my teeth are coming in*!
- Spanish Translation: *Os he dicho un millón de veces que los dientes me están saliendo.*

Image 4. A subtitled frame from the TV programme *Little Baby Bum* (2011)

Example 4: *Little Baby Bum*

- Original audio: *square, circle, triangle, rectangle.*
- Spanish translation: *cuadrado, círculo, triángulo, rectángulo.*

Image 4 above shows the audio and subtitle options available to watch this programme on Netflix. The subtitling possibilities are Castilian Spanish, English (intralingual), German, and Brazilian Portuguese; and the audio options include English (original and audio description), German, Castilian Spanish, and Brazilian Portuguese. These language options allow multiple possibilities to the audiences. However, it is unlikely that the target audience of these programmes, i.e. toddlers and preschoolers, will take advantage of subtitles when they do not have developed reading skills.

1.2.3 **Subtitling for the deaf and hard of hearing (SDH)**

Unlike general (interlingual) subtitling (for hearers), subtitling for the deaf and hard of hearing is normally an intralinguistic transfer, although it is becoming to be used for interlingual purposes as stated in the previous section. This mode of audiovisual translation basically consists in subtitling with the addition of certain elements that describe and communicate not just the dialogues of the characters on screen, but also elements like sound effects, suprasegmental aspects (accents, screaming, stammering…) and contextual information, as is established in the Norm [6] UNE 153010 (see hyperlink or annexe). According to Rica Peromingo (2016: 121), the most relevant characteristics of subtitling for the deaf are:

- A maximum duration of 6 seconds for 1 line and 10 seconds for 2 lines.
- There must be a break of at least 100 milliseconds between one subtitle and the following one.
- Each character in a movie has a specific colour assigned, depending on their relevance in the plot: yellow for the protagonist, green for the second most important character, blue for the third most relevant character, pink for the fourth character, and finally white for the rest of the characters, as well as for contextual information (which usually appears in brackets) and for the narrator of the film.

The images below show two frames of two English movies which are available in English SDH:

Image 5. An SDH frame from the TV Series *Homeland* S1.E1 (2011)

---

[6] Hyperlink to the Norm UNE 153010: hyperlink

Image 6. An SHD frame from the film *Pulp fiction* (1994)

1.2.4 **Audio description (AD)**

Audio description, together with SHD, is a breakthrough for the people who cannot fully experience a movie or audiovisual product due TO disability reasons, in this case blindness. It consists in incorporating an audio track to the film in which a narrator describes the action, body language, facial expressions, scenery and costumes (Benecke 2004: 78). In Spain, as in the case of SDH, there is a norm which applies to this translation method: the UNE 153020 (see hyperlink and annexe), according to which, the audiovisual products with AD possibilities are:

1 TV productions: movies, series, documentaries, etc.
2 Any recorded production, such as movies, series, documentaries, etc.
3 Theatre films.
4 Live shows: theatre, musicals, etc.
5 Monuments subject to be visited by tourists: churches, palaces, etc.
6 Museums and exhibitions.
7 National parks, theme parks, etc.

According to Benecke (2004: 79), in Germany, the AD process carried out by a team of three people, four if the narration is done by someone else, one of them being blind. Two people watch the film; in this way, they can monitor and complement one another to write the description; the blind member will then indicate the parts of the film in which the description

is needed, how much and what kind of information is required. The final text will then be revised, recorded by a narrator, and finally mixed with the film soundtrack. The process of AD in Spain is similar, although the number of participants may vary: referring to the UNE 1530020, Solé (2005: 12), states that the norm recommends that the script (the description) should be checked not by the writer himself/herself, but by someone else who could take notes to be included in the final script to be narrated. Moreover, there is no mention of having a blind person helping produce the final text, and audiodescribers in Spain have been mostly volunteers with no formal training, and without either social recognition or legal protection (Solé 2005: 16).

1.2.5 **Surtitling**

Surtitling is probably one of the least researched audiovisual translation methods (Rica Peromingo 2016: 146) and consists in placing a screen usually above a theatre stage on which a translated text is displayed. Surtitles were invented by the Canadian Opera Company during the 1982-83 opera season, and the name SURTITLES™ is a registered trademark since then (Mateo 2002: 52). However, the word *surtitles* has become the standard term used to refer to this mode of translation, although there are others such as *supertitles* and *supratitles*, which are also used by academics (Díaz Cintas and Remael 2007: 25).

Surtitles are commonly used in live performances like operas, theatre plays, and musicals, and there are several agents involved in the surtitling process: the translator and/or surtitler, the stage director, and the musical director (Mateo 2002: 57). The fact that surtitling is performed live makes spotting one of the trickiest issues.

The number of lines shown in the surtitle are between two and three, and unlike subtitling the surtitler has more freedom in terms of timing, however he/she must be aware of when the scenes change to keep the synchronisation with the action on stage. There is no uniformity in terms of the format of surtitling, according to Mateo (2002: 64) in terms of the usage of capital letters, the font of the text, punctuation, or its presentation. The number of characters varies from 25 to 35 per line; the texts are sometimes aligned on the left side of the screen, and sometimes centred; the colour of the letters may change too, it can be either red, green, or yellow; punctuation is not consistent, and it can be quite annoying since some

programmes do not include accent, exclamation or question marks, which hinders the comprehension in Spanish.

Among the main characteristics of surtitling are: (Rica Peromingo 2016: 149)

1-   Space-time limitations: there is a limited time which has to be synched with the performance on stage, and a screen on which the lines are displayed.

2-   Simultaneous channels of communication: there is the visual channel (performance on stage and reading of the lines) and the acoustic channel (music and actors' voices).

3-   Technical specifications: two lines with a limit of forty characters per line, a simultaneous display with the performance, etc.

Mateo (2007: 146) identified 4 technical aspects of the surtitling process. The first concerns the software system used, and although theatres commonly use Power Point when they do the surtitling themselves, there are a few specific surtitling programmes. The second aspect deals with the projection of the surtitles, which is usually made on a LED screen, although it can be sometimes more rudimentary in the form of a simple wooden board and a black cloth. Thirdly, the person in charge of the projection is mostly a technician from the surtitling firm, although if the commission is made to a freelance translator, it is the translator him/herself who deals with the projection of the surtitles. The final and fourth technical aspect of surtitling deals with unexpected changes in the performance; that is, if it is the translator who is projecting the surtitles, he/she may realise that a change has been introduced live on stage and not project the original corresponding title. This, in Mateo's (2007: 146) view depends "on the ability of the translator to respond quickly and on the type of change required".

The image below shows a scene of the play Cyrano de Bergerac performed in Paris, with English surtitles.
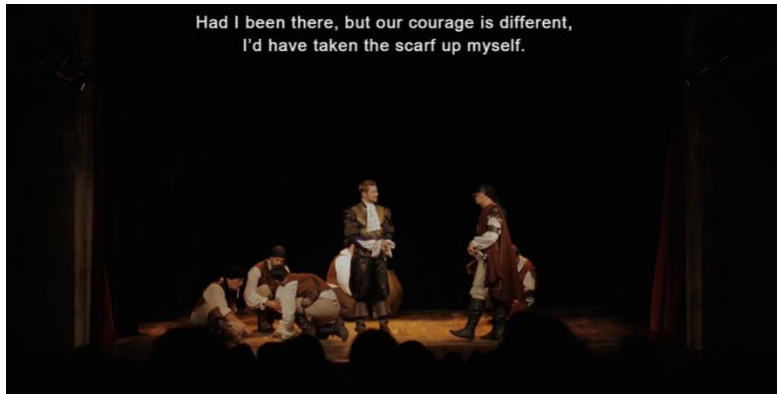
Image 7. A Surtitled scene from the play *Cyrano de Bergerac*. Source:
https://www.theatreinparis.com/blog/theatre-surtitles-past-present-and-future

According to Rica Peromingo (2016: 146), this audiovisual translation mode has increased exponentially in the last decade in Spain. However, this fact has not had surtilting obtain a relevant presence in the audiovisual translation undergraduate and postgraduate degrees in Spain.

1.2.6 **Interpreting: simultaneous and consecutive**

Interpreting is probably the most common translation method in live interviews, press conferences, meetings and court trials. Interpreting can be divided into two categories: consecutive and simultaneous. According to Hurtado (2001: 81), citing Jimenez (1999), the difference between the two lies in the fact that consecutive translation occurs when the oral reformulation (the translation) takes place after the oral speech in a source language; whereas simultaneous translation occurs when there are no interruptions for the translation in the original oral speech and the translation takes place simultaneously as the source text.

Within simultaneous interpretation, Hurtado (2001: 83) identifies three modalities depending on how it takes place. If there are electronic means it will be *simultaneous interpretation* proper. Interpreters frequently work from an isolated booth where they can listen to the speaker's words through headphones and translate them simultaneously, by speaking to a microphone which is connected to loudspeakers or earplugs, which the audience wear to receive the translation. This is, for instance, how political speeches are translated at the

European Commission meetings, (see hyperlink7 for more information on EU interpreting standards). If no electronic means are used, the modality is *whispering interpreting*. The interpreter would normally be close to the person needing the translation of the message in the source language. The third modality is called *relay interpreting*, that is, "interpreting from one language to another through a third language" (Shlesinger 2010: 276), and it happens when there is no interpreter for a linguistic combination and the solution is that one booth (interpreter) connects with another Jimenez (1999: 69): the translated text in the first booth will become the source text in the second booth, from which a translation is then provided. The image below, provided by Jimenez (1999: 70) shows graphically the flow of relay interpreting:
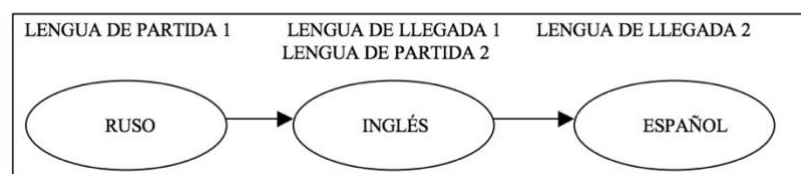
Image 8. The relay interpreting process (Jimenez 1999:70)

Unlike consecutive interpreting, simultaneous interpreting (in its three modalities) requires, as stated by Hurtado (2001: 83), the ability to reformulate an utterance while listening and understanding further utterances at the very same time, together with the need to develop strategies to compensate the gap between the original speech and its reformulation. On the other hand, in consecutive interpreting the speech to translate may be anything "between a minute and twenty minutes in length" (Gillies 2017: 05). This implies that in order to translate such long utterances, the translator/interpreter makes use of a combination of notes, memory and general knowledge to reformulate the original speech. According to Gillies (2017: 05), this form of interpreting is commonly known as l*ong consecutive*, to distinguish it from *short consecutive*, in which the source language speaker normally stops after each sentence, or couple of sentences, to allow the interpreter translate what has been said. In addition to this, Gillies (2017: 05) considers that consecutive interpreting is still today "an essential part of an interpreter's repertoire", and it is considered by many to be "the superior of the two skills" (simultaneous and consecutive). This is the reason why it is often better paid on the free market (Gillies 2017: 05).

---

7 Hyperlink: https://ec.europa.eu/info/departments/interpretation/standards-interpreting-facilities_en

1.2.7 **Fansubbing**

Fansubbing is an activity carried out by the fans of films, TV series, and cartoons (mainly *anime*), which consists in subtitling these programmes domestically before they are broadcast and translated in their countries. This practice has been defined as *domestic subtitling* (Chaume 2012: 04), and involves teamwork (Chiaro 2009: 151); fansubbers have to download the original videos from the Internet, process the timing, translate the dialogues, do the typesetting, edit and encode the subtitles, and finally distribute the programme (Díaz Cintas & Muñoz Sánchez 2006: 38-9). There are several differences between fansubbing and subtitling considering the former is an *amateurish* production. For Díaz Cintas & Muñoz Sánchez (2006), citing Ferrer Simó (2005), the main differences between these two audiovisual translation modes lie in the following characteristics observed in fansubbing (2006: 47):

- The use of different fonts throughout the same programme.
- The use of colours to identify different actors.
- The use of subtitles of more than two lines (up to four lines).
- The use of notes at the top of the screen.
- The use of glosses in the body of the subtitles.
- The position of subtitles varies on the screen (scene timing).
- Karaoke subtitling for opening and ending songs.
- The addition of information regarding fansubbers.
- Translation of opening and closing credits.

Fansubbing is considered technically illegal (Díaz Cintas & Muñoz Sánchez 2006: 45), considering, for example, the copyright laws which are infringed just by downloading the programmes from the Internet. However, some companies may see fansubbing as beneficial, since it helps to introduce episodes of new series and films in several countries so the owning companies can test their popularity among the target audience and thus conduct a market study or survey without spending a euro on it. Other companies may consider fansubbing as a "fan-activity" as long as it does not affect the sales of these products; finally, there are other companies that do not want to waste their time and money prosecuting foreign violations of their copyright (Díaz Cintas &Muñoz Sánchez 2006).

The images below illustrate two examples of fansubbed manga films.

Image 9. A fansubbed frame from an *anime* film. Image taken from Google Images. LINK

Image 10. A fansubbed frame from an *anime* film. Image taken from Wikipedia. Link

1.2.8 **Fundubbing**

Fundubbing consists in replacing a dubbed version of a film with a completely different target text normally for humoristic purposes (Chaume 2012: 04). Only short clips of films are usually fundubbed, sometimes the most famous or well-known scenes, and never the whole film. The reason for this is that fundubbing is usually done on comedy and entertainment TV programmes by professionals, and sometimes by amateurs on social media who enjoy dubbing for fun.

Fundubbing in Spain, at least into Spanish, became very popular thanks to the TV programme "El Informal" (The Informal one), which was broadcast in the late 1990s and early 2000s and included fundubbed clips in almost every programme. It could be said that El Informal started a fundubbing fashion that is used still today in several Spanish TV programmes like "El Hormiguero", "El Intermedio" or "90 Minuti". In the regional TV channel in Asturias (TPA), videos of this kind were very popular in the early 2000s thanks to "Terapia de Grupo" (Group Therapy), a programme conducted in Spanish but which included video-clips fundubbed in Asturian, the regional dialect.

But fundubbing is not only popular on TV: Social Media like YouTube and Twitter include hundreds, even thousands, of fundubbed videos. In addition, there are also smartphone apps which allow users to fundub short clips, as is the case with Dubmash[8]. This app is quite interesting because it allows users to fundub video clips but the other way around: instead of changing the text of the film, users change the video and the setting of the scene itself. By using the camera of the smartphone, users record videos with the original lines of the film and it is them who utter the original text.

The images and web links below show some fun examples of this audiovisual translation method, which is gaining popularity thanks to social media and new technologies:

Image 11. A fundubbed frame from the film "El burlador de Castilla" (1948), a classic for fundubbers.

Image 12. El Informal team was made up by: (standing) Miki Nadal, Florentino Fernández, (sitting) Javier Capitán, Patricia Conde, Felix Álvarez *Felisuco*

Web links:

- The best of "El Informal": https://www.youtube.com/watch?v=nrPuOkqU51I
- Terapia de Grupo:
  https://www.youtube.com/playlist?list=PL01D0B8BEAEE5DDA2
- Dubmash compilation: https://www.youtube.com/watch?v=4u2E51Xz_AA

1.3 Voice-over

This section is devoted to provide a thorough description of what *voice-over* is both from an academic and industry-related perspective. It starts describing what voice-over is from a Translation Studies perspective, which, as we will see, differs from the conception that industry experts and academics of other disciplines have of this term. Next, we will see the linguistic features that make it, according to some scholars, the *most faithful* translation mode, and finish with the technical features that characterise and shape this translation mode.

Before describing the main characteristics of voice-over as an audiovisual translation mode, it is convenient to note and be aware of the terminological controversy that surrounds the term itself. Since this issue will be discussed in detail in section 1.3.1 below, let us now simply mention that for Díaz Cintas & Orero (2006: 477) citing the OED (Brown, 1993: 3596) voice-over is "the narration spoken by an unseen narrator in a film or a television programme; the unseen person providing the voice". These authors agree in the fact that this definition does

not belong to the Audiovisual Translation field since there is no reference in it to the translating activity. In fact, this is more of a definition suitable for the Media and Film studies field, and, as will be discussed further on, it is the definition most voice actors and agencies would subscribe to.

From a Translation Studies perspective, and as stated by Díaz Cintas & Orero (2006), voice-over can be defined as:

> A technique in which a voice offering a translation in a given target language (TL) is heard simultaneously on top of the source language (SL) voice… [T]he soundtrack of the original programme is reduced to a low level than can still be heard in the background when the translation is being read. It is common practice to allow the viewer to hear the original speech in the foreign language for a few seconds at the onset of the speech. (2006:477)

The process of creating a voice-over translation, from a Translation Studies perspective, can be twofold (Franco, Matamala & Orero 2010): voice-over for production, and voice-over for post-production. The difference between the two consists in the fact that in voice-over for production translators are not provided with the source text, just with the video material, and they have to transcribe and then translate the text; on the other hand, in voice-over for post-production, translators are provided with the text, a character list and the video file to work with. Since this audiovisual translation method has been selected as the object of study of this doctoral thesis, these two modalities, which constitute the process of voice will be analysed in more detail in the following section (1.3.1). In addition to this, Matamala (2019: 67-68) includes two more categories within voice-over regardless of whether they are for production or post-production: first and third person voice-over, depending on whether it is a direct voiceover (first person) or reported (third person); and single and multiple voice-over, in which only one voice is heard by the audience, as in the case of translated TV programmes in Poland, or more voices are involved (both male and female), as in Spanish voiced-over programmes.

Voice-over is the preferred translation method for different TV programmes in Spain nowadays, and its presence has grown exponentially in recent years, probably due to the fact that it is considerably cheaper and faster to produce than dubbing. The vast majority of the programmes that resort to voice-over in order to be translated are the so-called *unscripted*

programmes; which are those in which the action, the plot and the dialogues develop spontaneously without the need to have a planned script. Some of these programmes might have a short script to provide a narration in order to join the different scenes, or parts, of the programme. The table below shows a list of examples of this kind of programme (to mention a few), together with the channel and VOD (video on demand) platform they are broadcast on, and their names in the English (original) version and in the Spanish (target) version.

Channel / Platform	ENGLISH TITLE	SPANISH TITLE
Netflix	Story of Diana Lock Up: disturbing the peace Big Family Cooking Showdown The Royal House of Windsor	Story of Diana Entre rejas: perturbar el orden La mejor familia en la cocina The Royal House of Windsor
DMAX	The Ultimate Survivor River Monsters	El último superviviente Monstruos de río
DIVINITY	Love it or list it Income property Auction Hunters Wheeler Dealers	Tu casa a juicio Tu casa lo vale Guerra de subastas Joyas sobre ruedas
CUATRO	The dog whisperer	El encantador de perros
MTV	Geordie Shore Teen Mom	Geordie Shore Teen mom

Table 2. Some voiced-over TV shows available in Spain

There are some technical requirements that apply to voice-over translation, as in the case of dubbing and subtitling. These requirements, or constraints, are spatial and temporal. Spatial since the translated text has to match the text length of the original text, or even be shorter; and temporal: because the voice providing the translation has to be in synchrony with the original voice. This synchrony is known as phrase-sync, which differs from dubbing's lip-sync in the fact that it is less technical, and therefore, cheaper to produce as less studio time is required. In addition to this, there is no need to search for words with similar phonetic realisations for close-ups and extreme close-ups, as is usually the case with dubbing.

Voice-over translation is regarded as a partially vulnerable translation mode (Orero, Franco, Matamala 2010 citing Díaz Cintas 2003), since the two languages (source and target) are available to the audience. This means that the audience can compare and judge the translation in those parts in which both languages are most clearly heard: at the beginning and at the end of the phrase. On the other hand, this vulnerability feature plays a role in providing "reality, truth and authenticity" (Orero, Franco, Matamala 2010: 25), which make voice-over the preferred translation mode for the types of programme mentioned above.

Therefore, voice-over can be described as vulnerable, constrained and cheap, and less relevant (in terms of research) than other translation methods. However, the fact of providing access to the two languages (source and target), its spontaneity, an undoubtedly real oral component, and the technical means available to produce it, make this translation method both exciting and challenging, considering the research possibilities it affords both from a linguistic and technical point of view. The following section will be devoted to analysing the specific characteristics of voice-over from a professional and an academic perspective in detail.

1.3.1 Industrial and academic conceptions of voice-over

The fast-growing market of online localisation agencies and home-recording studios has made *voice-over* a popular term, or buzzword, for the last ten years. If one searches the word *voice-over*, *voiceover*, or *voice over* (without a hyphen), on Google, one will get over sixty million results, the first fifty entries referring to mostly websites of online agencies, definitions and voice actors. If one does the same kind of search for the words *dubbing* and *subtitling*, the most prestigious and popular translation modes, the results are less than half those of *voice-over* (eleven million and barely half a million entries, respectively). This may look like a contradiction if one considers the popularity and relevance of each mode in the translation world, but the reason for this unusual numerical difference lies in the fact that *voice-over* is also used in computing to refer to VOIP (Voice Over Internet Protocol), which is the way the voice travels through the internet when using software like Skype, FaceTime, or Whatsapp, and it is technically defined by Varshney, Snow, McGivern, and Howard (2002: 89) as:

[…]sending voice transmissions as data packets using the Internet Protocol (IP), whereby the user's voice is converted into a digital signal, compressed, and broken down into a series of packets. The packets are then transported over private or public IP networks and reassembled and decoded on the receiving side.

However, leaving Google aside, voice-over does not seem to be as popular in the academic field as it is in the professional world, considering the large number of agencies and companies that include this term as part of their slogan or *motto*. In fact, the term itself is used in rather different, even confusing, ways depending on the field one is dealing with.

In the academical world, the term *voice-over* has generated some controversy among scholars, and the same concept has received a wide range of terms. This is due to the fact that *voice-over* was first used in the field of Film Studies and then borrowed by Audiovisual Translation Studies. Thus, according to Franco, Matamala and Orero (2010: 30-31), who conducted a questionnaire among voice-over practitioners and scholars, *voice-over* translation is defined as or compared to:

a) a category of *revoicing*, along with lip-synchronization dubbing, narration and free commentary (Luyken et al. 1991, Baker 1998, Chaves 1999, Russo 1995);
b) a type of *dubbing* (Fawcet 1983, Gavrilov Translation 2007), either "non-snchronized or non-lip sync dubbing" (Dries 1995); or (surprisingly) its opposite "doublage synchrone" (Kaufmann 1995), followed by *doublage voix off* (Kaufmann 2004); or just "half dubbing" / *demi-doublage / seidoppiata* (Gambier 2004, Mayoral 2001, Perego 2005);
c) *subtitling* (Espasa 2004) or *oral subtitling* (Gambier 1996);
d) a type of *interpreting* (Pönniö 1995, Gambier 1996, 2000, Chaves 1999, Díaz Cintas 2001b, Elias 2006, Espasa 2004, Gavrilov Translation 2007);
e) *traducción en sincronía* (Del Águila & Rodero Antón 2005);
f) *spoken translation* (BBC Broadcasting Research 1985);
g) *live translation* (Elias 2006);
h) *simultaneous translation* (Ávila 1997, Russo 1995, Del Águila & Rodero Antón 2005);
i) *wersja lektorska* or the *reader's version* (Garcarz 2006);
j) *a type of narration* (Karamitroglou 2000);
k) *l'adaptation en aval* (Kaufmann 1995);
l) *recorded commentary* (Mailhac 1998, Russo 1995).

This blurriness generated around the term is rather disappointing; if all these terms mentioned above refer to the process of voice-over, why then is not the word *voice-over* used as the umbrella term? An interesting example of this terminological mess can be drawn from the fact that Chaume (2012: 03) finally considered *narration*[9] as a type of *voice-over*, in contrast with an earlier publication (2004: 37), in which *voice-over* and *narration* were two different audiovisual translation modes. In addition to this, Sánchez Mompeán (2015: 272) adds one more term to the debate: voice acting. In her work on dubbing animations, she states that the terms *voice-over* and *voice acting* are used interchangeably to refer to the same thing, in this case putting someone's voice to an animated character on screen. And to make matters worse, she refers to *The Glossary of Voice-Over Terms* (2013) (see [hyperlink](#)[10]), in which *voice acting* would be a more accurate term than *voice-over* to refer to the art of using the voice to bring life to written words. Sánchez Mompean's views contrast with Darwish and Orero's (2014: 06), for whom voice-over is often confused with voice acting, and consider that, as far as AVT is concerned, "voice acting has in fact been used where voice-over should" (2014: 07) and highlight that voice acting may cause serious dissonance between original and translation in terms of quality of voice, pace, pitch, matchability and register. Moreover, and to add more fuel to the fire of this term mayhem, a new word has recently appeared in Italy to refer to voice-over translation when the delivery is acted as in a dubbed film: *simil sync* or *semi-sinc* (Sielo 2018: 260). Thus, the list of expressions to refer to voice-over is enlarged, since, in our view, simil sync is dramatised voice-over, as Sielo (2018: 266) states:

> Il semi-sinc, pur ancora poco analizzato dalla critica e dispregiato da molti dialoghisti, sembra un ottimo compromesso economic e tradutivo, in quanto permette di godere di un dialogo sufficentemente drammatizzato e verosimile, in cui la voce originale a tratti quasi svanisce, in cambio di tarife ridotte e un tempo di lavorazione anch'esso sensibilmente inferiore rispetto al dopiaggio. [Semi-sinc, although still poorly analyzed by critics and disregarded by many dialogists, seems an excellent economic and translational mode, as it allows to enjoy a sufficiently dramatized and faithful dialogue, in which the original voice, at times, almost fades away, in return for reduced fees and studio time which, in addition, are significantly lower than those of dubbing. (my translation)].

[9] Narration is simply a kind of voice-over, where the translation has been summarized. (Chaume: 2011: 03)
[10] Hyperlink: https://www.voices.com/blog/glossary-of-voice-over-terms/

For Eliana Franco (2001), this terminological conflict reflects the isolated attempts made by audiovisual translation scholars and practitioners to mix the translative notion of the term (a translated voice on top of an original one) with the traditional notion of *voice-over* (a technique of presentation/location of a voice). This can be exemplified in Chaume's 2012 publication, in which he even considers voice-over not to be necessarily a type of revoicing citing Wright and Lallo's (2009) definition of *voice-over*, which is closer to the Film Studies perspective; nevertheless, Chaume finally discards this view since it would imply a "broader definition of the term, related to the voice talent's industry" (2012: 03). Paradoxically, this seems to be in contradiction with his admitting *narration* as *voice-over*, which is closer to Film Studies than Translation Studies, as claimed by Franco, Matamala and Orero (2010: 39), for whom the association between these two fields of study, and the very fact of considering *narration/commentary* as instances of *voice-over*, lead to negative consequences. For them, these two, sometimes three, terms (commentary/narration and voice-over) cannot be the same thing because the audience is hardly aware that commentaries are translated since nothing indicates there has been an interlinguistic process. These scholars prefer to use the term *off-screen dubbing* (2010: 41) to refer to commentaries as there is a process of adaptation, an absence of the original audio track and on-screen speakers, and no need for lip-sync.

Moving on now to the professional world, including agencies and companies that do *voice-overs* worldwide (studios, localisation agencies, and freelance voice talents), their view on what *voice-over* means is probably different from AVT scholars', and is more in line with Film Studies' understanding of the term. Let us take Google's number one ranking agency: VoiceArchive, from Aarhus, Denmark, whose motto is "*We do voice-overs*" (see image 13 below), by which they do not mean with this that they only provide the audiovisual technique in which one voice is heard on top of another, the original voice being reduced in volume.

Image 13. A screenshot of VoiceArchive's home page
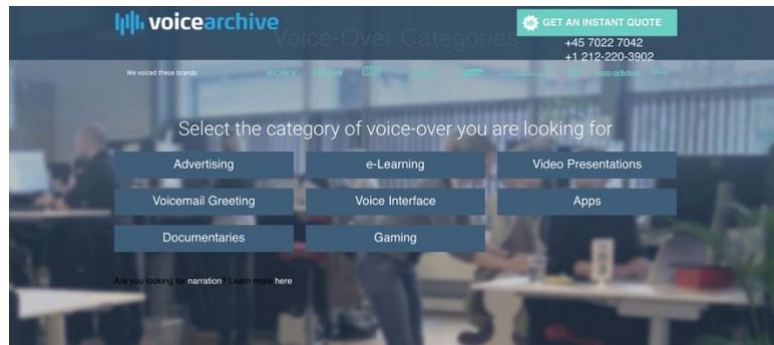
Image 14. A screenshot of VoiceArchive's voice-over category section.

What they are trying to convey through this motto is that they also provide narrations, commentaries, dubbing and voice-acting, as can be seen in image 14 above, which shows the categories of voice-over (notice the use of the plural – *voice-overs* - by this agency in the motto) provided by this company. Projects like e-learnings, and documentaries could be regarded as commentaries/narrations, and similarly, voicemail greetings (or interactive voice response, a.k.a. IVRs), advertising, apps, gaming and video presentations could be included in the voice-acting category, or even in dubbing. For professionals, *voice-over* is a term that blends the two definitions given by Film Studies and Translation Studies. Thus, the narration of audiobooks, audio guides, documentaries, e-learnings, video games, cartoons, even film dubbing and probably anything to do with a voice and a film, although the latter is not always needed, as is the case with audio guides, would be considered *voice-over.* In fact, VoiceArchive is not the only professional company around the world that refers to *voice-over* as a general term: there are others (see figure 6 below) which also use it this way, such as: GoLocalise (*London's leading voice-over, subtitling and translation agency since 2008*), and VSI (*The global service provider for dubbing, voice-over and subtitling services since 1989*), VoiceBooking (*the easiest way to book a voice-over*), and thousands of other agencies like Voices.com, VoiceBunny, Voice123, Hobsons International, Bodalgo, Tellary, etc. all offering *voice-over* services, from translating the script to mixing the audio. In addition to this, the people in charge of delivering the translation of an audiovisual product are known as voice talents or *voice-over* talent or artist, as stated by Matamala (2019: 65), showing that voice-over has a wider scope in the professional world than in the academic one.
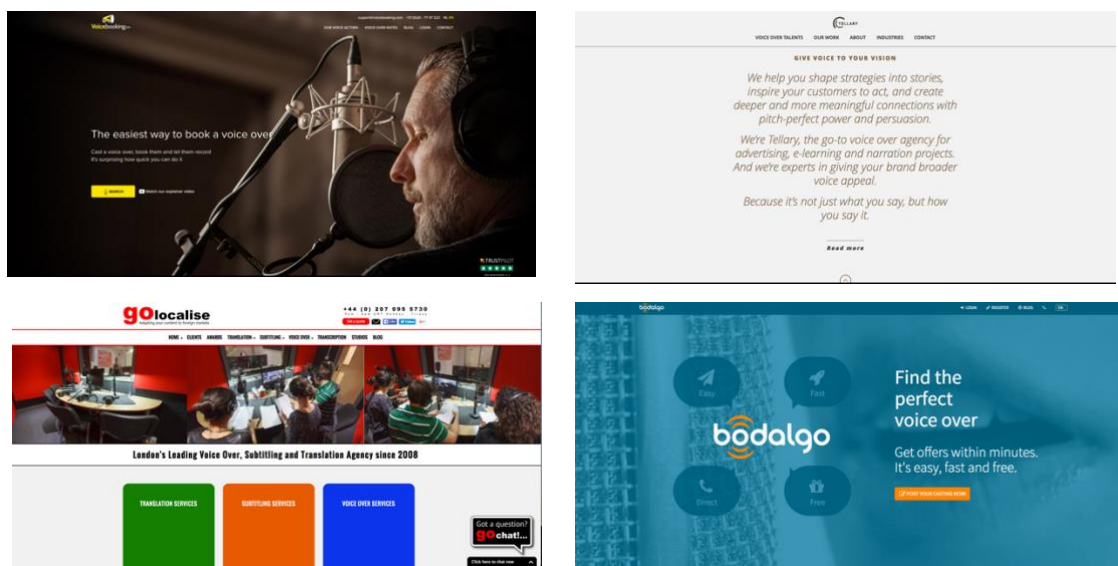
Figure 6. Home page of voice agencies VoiceBooking, Tellary, GoLocalise, and Bodalgo

In the light of all this, there seems to be no doubt that there is a clear gap between professionals and scholars, and most importantly among scholars themselves, who use rather different terms to refer to the very same activity. In Eliana Franco's view (2000: 34), this gap, or *minimal interaction,* between professionals and scholars of audiovisual translation is reflected in the terminological ambiguity around the term *voice-over*, and consequently leads to "discussions which lack the common sense that could contribute to advancement in the field." (2000: 34) For her, the conflict is not entirely terminological but conceptual, when she states:

> And most importantly, the prevailing discourse on voice-over reveals that the actual conflict is not so much about terminology, but about the object of study itself. That is, it seems that the status of voice-over as a mode of transfer *per se* has been neglected, mainly because it always appears as a technique or a sub-category associated with other "well-known" types of audiovisual translation." (2000: 34-5)

Eliana Franco is still right when she mentions that voice-over is considered a sub-category within the field of audiovisual translation. Despite the fact that there has been, professionally speaking, a voice-over boom in recent years, dubbing and subtitling are still the most common methods of audiovisual translation, and consequently the most favoured in terms of academic research and prestige. This is probably due to the fact that, as Orero herself states (2004: 76), "translation for voiceover has not been successful in attracting the attention of the

academic community, and more in particular the interest of those researching within the field of Audiovisual Translation (AVT)." How voice-over translation can succeed in attracting the academic community is certainly enigmatic; just as enigmatic as how the academic community has ignored voice-over for so long. Gambier & Suomela-Salmi (1994: 243), in line with this, already found this situation *surprising* given the "huge variety of audiovisual communication", and acknowledged the fact that dubbing and subtitling had been considerably more favoured in terms of research in the 20th century. This belief has been later shared by other scholars, like Orero (2004: 76) and (2005: 177), Franco (2001), Rica Peromingo (2016: 143) and Sepielak (2013: 23), and is graphically represented in the table below, elaborated by Franco, Orero and Matamala (2010: 23) in their monographic book, showing the number of publications on voice-over and other AVT modes until 2010.

Year										
	1	2	3	4	5	6	7	8	9	10
2008										
2007										
2006										
2005										
2004										
2003										
2002										
2001										
2000										
1999										
1998										
1997										
1996										
1995										
1994										
1992										
1991										
1985										
1983										

> ■ research on voice-over translation
> ▨ research on a different AVT mode
> 2008 indicates three works that are forthcoming and one work already made public
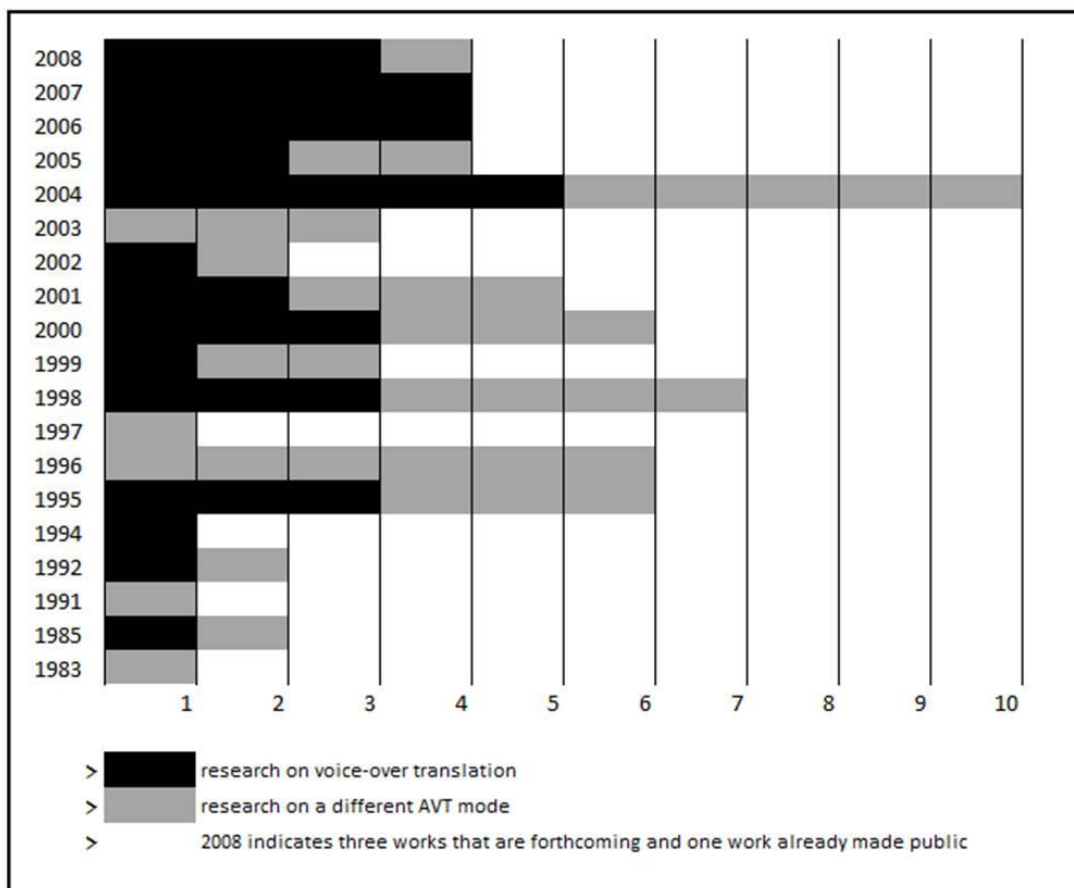
Table 3. Number of publications on voice-over and other AV modes according to Franco, Matamala, Orero 2010: 23)

Indeed, there is rather important aspect which may play against voice-over translation's popularity: money. Studio time is considerably pricey, even nowadays, when all (if not most) recordings are done using digital software. This *push-button technology* (Chiaro 2009: 146) is

notably easier, more reliable and less time consuming than old analogic recordings for which tapes were needed. In the audiovisual world, time is money, and voice-over is, as opposed to dubbing, less complex technically speaking, and less demanding (Díaz Cintas & Orero 2010: 478). There is no need for lip-sync, so the production of a voiced-over product will be considerably cheap compared to dubbing, given the short studio time needed to edit and synch the audio track to the film. In line with this, Górska (2015: 68) considers that voice-over is cheaper than dubbing not only in terms of studio time, but also in the number of professionals involved: a translator, a proof-reader (if any), a *reader* and a sound engineer. Cheap products normally tend to be considered poorer in quality than expensive ones, and this very fact rather than making it more popular, may actually go against voice-over.

Nevertheless, the fact that voice-over is cheaper than dubbing can be debatable nowadays, considering the large number of online professional agencies posting dubbing jobs (Voices.com; Voice123.com; Bodalgo.com; VoiceBunny.com, etc.) and voice talents who can synchronise audio to video as they record from their facilities, namely home studios, delivering synchronised audio ready to mix. With the emergence and spread of professional digital recording software in the late 1990s and 2000s, the number of people with recording capabilities for music, and voice, be they professionals, amateurs, or aficionados, has gradually increased, and the number keeps growing today. The accessibility to top-notch software like Logic Pro, Pro Tools, and Cubase, or even freeware (free software) like Audacity or Garage Band (if you own a Mac) has made anyone with a computer and a USB microphone, for example, a potential voice talent. All this has made traditional recording studios lose a substantial part of their studio time, and therefore a significant amount of money. Consequently, the possibility for companies of skipping the studio, plus a fierce voice talent competition, has made the rates plummet in the voice industry. However, and as is usually the case with any project and industry, the budget always depends on the length of the project and on whether a buy-out fee, or royalty-fee, applies, which is not always the case.

Moreover, another disadvantage, if being cheap is one, could be that voiced-over products lack the glamour and appeal of dubbed/subtitled products like Hollywood films and TV series. Who would be willing to study and watch hours of documentaries, interviews and reality shows, when one can do that with Oscar-awarded films? Not to mention the difficulty one faces when trying to get the original and translated versions of the videos and texts of these programmes. Films and series in DVD format already include the original version and its

translation, both in dubbing and subtitling, and film scripts, if needed, are not very hard to get online. Nonetheless, and regardless of whether voice-over is considered a more or less interesting, attractive, or glamorous topic, it is undoubtedly a fast-growing genre within audiovisual translation that deserves its place and research.

There are numerous subgenres, categories or programme types, that use voice-over as their preferred translation method. These include:

1   Interviews (news reports, talk shows, bonus features on DVDs, sports programmes both on radio and tv).

2   Documentaries.

3   Web videos.

4   TV programmes: reality shows, talent shows, biopics, docudramas, and lifestyle shows.

These are mainly the kind of audiovisual products that use voice-over as their main translation mode in Spain. However, in countries like Poland the Baltic States and some Commonwealth countries (Díaz Cintas 2003: 196) voice-over is the preferred mode to translate films and movies.

Although – as has been made clear in this section -, voice-over is not as popular as other audiovisual translation disciplines like dubbing, subtitling and even audio description, this does not mean that voice-over is a discipline not worth researching. In fact, there are certain features that make this translation mode unique and, in some cases, a more suitable one than others. But before delving into the characteristics and features that make voice-over unique, it should be remembered at this point that voice-over can be divided in two categories, depending on the translation process, as we saw in section 1.3: translation for production, and translation for post-production (Franco, Matamala & Orero 2010). In the case of voice-over for production, the translator has to work with "raw, unedited material, which will probably undergo several processes before being broadcast" (Díaz Cintas & Orero 2010: 442). On the other hand, voice-over for post-production implies translators working from a final source text and video, which will not be subject to any further editing (Díaz Cintas & Orero 2010). This means that in some cases (production), translators have to translate the text by ear, as if interpreting, or ask someone to transcribe it for them and then translate it, which would normally imply a waste of valuable time. The fact of getting the script directly from the client, or the agency, makes things

significantly easier. However, be that as it may, in the end there will always be a script to translate notwithstanding the way it was obtained (transcribed by the translators themselves, or by someone else from the agency).

The main features of voice-over translation, both for production and post-production, could be divided into two categories: translational and linguistic on the one hand, and technical on the other; they will be analysed in the next two sections.

1.3.2 **Translational and linguistic features**

From a translational and linguistic perspective, voice-over may be considered an excellent translation method since it implies certain characteristics that others lack. Franco, Matamala and Orero (2010: 25) refer to several publications, claiming that:

> As pointed out repeatedly by most studies, voice-over is generally the preferred mode of transfer for the non-fiction genre, along with subtitles, because its defining features contribute to the appeals of reality, truth and authenticity that factual programmes count on in order to prove that their arguments are right or believable.

The notions of *authenticity*, *faithfulness*, *loyalty*, *truthfulness* and *equivalence* in translation have been the hot topics of debate among scholars for decades. These characteristics are usually linked to voice-over translation, as will be seen below. At this point, however, it is convenient to remember what each of these terms involves from a translational perspective. Hurtado (2001: 202) considers that *faithfulness* is understood in terms of fidelity to the meaning; that is, fidelity to what the original sender wanted to convey in the original text; and she highlights three categories that characterise and condition *fidelity* or *faithfulness*: subjectivity, understood as the compulsory intervention of the translator subject; historicity, which implies the impact within the socio-historical context; and functionality, conceived in terms of the implications of the textual typology, the language and the target means, as well as the purpose of the translation. Therefore, there are several ways of being *faithful* according to Hurtado. In line with this, Chaume (2012: 17) understands *fidelity,* or *loyalty,* as fidelity to the content, form, function and source text effect. Both Hurtado and Chaume emphasise the

importance of functionality, and the transfer of meaning in order to achieve a *faithful* and *loyal* translation. This point of view is also shared by Larsen (2001: 03), with the idea of *trustworthy* translations, and also by Gutt (2000), who states that the assumptions the communicator tries to transfer can be conveyed in two ways, by means of *explicatures* or *implicatures*, which were defined as follows: (2000: 80)

> Explicatures are those analytic implications which the communicator intended to communicate. Implicatures are a subset of the contextual assumptions and contextual implications of an utterance or text (…) Both explicatures and implicatures are identified by the audience on the basis of consistency with the principle of relevance (2000: 80).

The principle of relevance refers to one of the aspects of the cooperative principle, by which the participants or interlocutors cooperate with each other by observing certain maxims (see Grice 1975), and was further developed by Sperber and Wilson (1986). For Gutt (1990), the notion of relevance can be defined as a cost-benefit relation, by which the cost stands for the amount of mental processing effort needed to interpret the stimulus, and the pay-off consists in the contextual effects that derive from it. For this scholar, "the less effort the processing of a stimulus requires and the more contextual effects it has, the more relevant it will be" (1990: 142). These stimuli can be found in terms of their linguistic properties, or the perceptible phonic substance (Hatim & Munday 2004: 64), as can be the case of a language such as English, for example. This is one of the reasons why, in this work, we will try to analyse the rendition of the intonational and prosodic clues in voice-over translation.

With this in mind, the demand to keep the information content of the original corresponds to the demand that the explicatures and implicatures of the translated version should be the same as those of the original (Gutt 2011: 80). As a consequence of this equality regarding the implicatures and explicatures in both source and target text, the concept of *equivalence* arises. Hatim & Munday (2004: 52), understand that achieving *equivalence* involves a complex decision-making process, defined by Levy (1967) as if playing a game of chess in which there are many choices to make from several alternatives. According to Koller (1989), there are at least five frames for equivalent relations in translation, analysed by Hatim & Munday (2004: 50-51) and Pym (1997: 02), which are the following: formal (the formal-aesthetic qualities of the source text), denotative (based on extra-linguistic factors), connotative

(based on the way the source text is verbalised), text-normative (textual and linguistic norms) and pragmatic or dynamic (with regards to the receiver of the target text).

It is this last frame of equivalence, pragmatic/dynamic, which coincides with Rabadán's (1991: 45) view of equivalence (*translemic equivalence),* seen as a characteristic established through communicative coordinates whose main aim is not to achieve a "righ*t"* version of a translation, but to get an acceptable equivalent in the target polysystem: "the goal of an equivalent text (a translation) is to reach the audience of the target polysystem (1991: 45 - my translation). And we cannot forget the fact that, in this case, we are dealing with, *equivalence* in audiovisual translation, in which, as Vercauteren highlights (2016: 45), the visual and aural non-verbal elements of both target and source texts have to be considered, since they play a key part in any form of audiovisual product.

All the aforementioned scholars coincide in the fact that in order to label a translation as *faithful, loyal, trustworthy* and *equivalent*, the communicative purpose, and intention, of the original text has to be met in the target text.

Amongst the most relevant definitions of voice-over translation, collected by Franco, Matamala and Orero (2010) are the following, all of which include the notions we have been discussing:

1- "The faithful translation of original speech, and approximately synchronous delivery […] Voice-over is used where a sense of authentic presentation is to be combined with an almost full translation of the original text". (Luyken et al 1991: 80)
2- "Le voice-over ou interpretation simultanée, se characterise par une traduction asez "fidèle" de l'original et émise en quase synchronie." (Voice-over or simultaneous interpretation is characterised by rather "faithful" translation of the original broadcast in almost synchrony) (my translation). (Gambier 1996: 8)
3- "[In] sum, voice-over translation has to be faithful, literal, authentic and complete version of the original audio. Such definitions give voice-over the status of a trustful transfer mode, because after all, the definitions suggest that original speech will be rendered word by word". (Franco, Orero & Matamala 2010: 26)

Chaume (2004: 35) agrees with these scholars in the way voice-over is characterised as truthful and considers that the fact of not erasing the original audio track, together with the *sound bites* (hearing the original audio for a few seconds at the start and finish of the utterance), contribute to increasing the notion of truthfulness, compared to dubbing, for example. These technical features will be analysed, along with others, in the next section of this work.

The sense of realism and authenticity that voice-over provides makes it more attractive than other audiovisual translation modes for viewers, and that sense of truthfulness should be paramount in translation priorities. However, whether this is fully achieved will be challenged and put into question in this work; it is possible, for instance, that in audiovisual translation, and most specifically in voice-over translation, relevant communicative aspects conveyed by means of prosody, for instance, by intonation in English are not rendered into a language like Spanish, thus compromising these notions of fidelity and trustworthiness that usually accompany the definitions of voice-over. But before that, let us move on to the technical features of this audiovisual translation modality.

1.3.3 **Technical features**

Voice-over is considered to be technically less complex than dubbing, and therefore a more economical translation mode. The most noticeable technical features that any viewer of voice over 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 lapse, or gap, at the beginning and the end of the utterances during which one can hear for a few seconds the original voice, known as *sound bites*. This means that the translated voice starts a few seconds after the original voice, and ends a few seconds before the original.

There is no given pattern as to how long this delay should last or as to whether this could change according to the genre in which voice-over is used, since there are examples that show different possibilities. With this in mind, according to the scholars quoted in the table below, sound bites could last as follows:

AUTHOR	SOUND BITE LENGTH
Orero (2004: 83)	"Some two seconds before and after".
Espasa (2004: 189)	"Two seconds before and after, or three or four words."
Mayoral (2001: 03)	"Three or four words."
Agost & Chaume (1999: 259)	"About three seconds at the beginning but both languages finish at the same time in the case of documentaries."
Darwish & Orero (2014: 143)	"A few words at the beginning, and usually original and translation finish at the same time in the case of voice-over in news broadcasts."
Chaume (2012: 03)	"The translation is heard a few seconds after the original voices."

Table 4. Duration of sound bites according to different authors.

Despite the evidence shown in these descriptions, the duration of this gap is not always existent; in fact, there are cases when it does not take place at all, although this is not very common. However, when it does take place, its length is more a matter of how long the translated script is and whether there is the chance to leave this gap at the beginning and the end. In any case, the fact that this gap exists makes the translation process more challenging, and given that there are around four seconds missing (2 at the beginning and end) can the result be as "faithful" and "loyal" as it is claimed? Can the viewer have such a feeling of authenticity knowing that there will be parts of the original that will be deleted, or are they not aware of this? More specifically, will basic and relevant intonational communicative elements be removed, or obviated? These are some of the questions that will be discussed in detail further on in this work.

But the time gap, or sound bite, is not the only technical feature of voice-over. Unlike dubbing, in which lip synchrony is paramount, the degree of synchrony in voice-over is less demanding and *just* requires to match the original audio in time length, what is known 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 reason why there has been a 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 only requirement is simply to produce a text translation which is similar to the original in terms of 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, but phrases 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, this poses a serious problem. This issue will be addressed further on in this work together with other factors related to the different means each language seems to prioritise in oral communication. Now let us examine the synchronic features of voice-over.

Indeed, the synchronic characteristics of voice-over translation are:

- Isochrony: the duration of the translated voice has to match in length the original voice. For Schwarz (2011: 399), "perhaps the most important synchrony is not qualitative (or lip synchrony) but quantitative. To an audience, the divergence of visually and acoustically perceived utterances is extremely distracting." This is really important when video editing is not possible: there may be cases in which the translation could be longer than the original script, 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 find a way to make their lines fit in the given time codes. This is one of the main challenges of voice-over isochrony for Orero; in these cases, the translated texts have to be reduced or condensed. This researcher identifies three "strategies" (as she calls them) (Orero 2006a: 260):

    1- In the case of 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 the case of interviews it is often the case that the post-production team can edit the video to fit the voice track.

3- In the case of 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.

- Literal synchrony: it consists of a literal 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. This contributes to the 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 a few seconds later (in Spanish in our case). If there is a content match the audience will probably be happy with, and trust, the translation; if there is no content 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 narrator is actually saying. Franco, Matamala and Orero (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), a literal translation will have to be given even in cases in which a well-constructed semantic translation may go.

- Kinetic synchrony: body language has to be "in sync" with the script being conveyed orally (voice). Schwarz (2011: 400) emphasises the relevance of this feature stating that "the interplay of speech and gesture is an entity, and confusion can ensue when the two do not harmonise". It is important that the voice talent's delivery be in sync with the gestures and attitude of the person on screen. In order to achieve this, and to avoid dissonance, the translator has to make the translation match the original in those parts that carry the most relevant information, both verbally and on camera. In Schwarz's view:

> An actor's delivery gives rhythm to syllables and emphasizes the important parts of speech. Nodding, raising of the eyebrows, or gestures coincide with the stress-bearing component of an utterance, which is also referred to as `nucleus´. The interplay of the nucleus with the body movement plays an important role in the perception of speech (2011: 400)

Kinetic synchrony is paramount to make the product believable to the audience. If a character on screen is, for example, shouting, waving the arms, or even dancing, and the voice providing the translation does not show any emotion compared to the original, the sense of truthfulness might get lost.

- Content synchrony: content synchrony is successful when "semantic content of the SL and TL script versions match each other closely" (Fodor 1976: 77 in Schwarz 2001: 400). Matching content in voice-over translation is considerably difficult considering that in some languages, as is the case of Spanish (see above), there is an approximate 30% increase in word count compared to English originals. This means that the target text will have to be reduced, with the obvious loss of content if isochrony is to be achieved. Therefore, a sort of *straitjacket* (Schwartz 2011: 400) is imposed by isochrony and lip-sync on content synchrony. Moreover, there is the different intonational behaviours of languages, like English, which can displace the tonic syllable, or nucleus of intonation, showing a flexibility which cannot be equalled by other languages, like Spanish, for example, and therefore make content synchrony in translation between them extremely difficult. This feature will be at the heart of this study and is substantially relevant to achieve the touch of authenticity, truthfulness and fidelity that voice-over translation claims. These linguistic features, plus the presence of sound bites, together with the general 30% increase in word count, and literal and content synchrony, make the translation of voice-over, a rather difficult task, at least in the English-Spanish direction, and the present study will examine whether translators actually have all these features in mind when translating voice-over.

- Action synchrony: there should be an agreement between the target voice and what is shown on the screen in terms of content and image. For example, in the case of interviews, Franco, Matamala and Orero (2010: 82) consider that if there is a reference to an element on screen, the translated version must keep this synchrony and refer to the same element at the same time it is shown on screen. Here the length of the translation is quite relevant; if the translated voice reaches the element somewhat after the corresponding image is shown on screen, then this kind of synchrony would be broken.

- Character synchrony: this could be understood as how similar the original and the translated voices are. The term was coined by Fodor (1976) and, in Schwarz's words, it is used "to describe the degree of correspondence between the dubbing voice e.g. timbre, tempo used, and the original actor's physique manners, and gestures" (2011: 401). Fodor sees it mainly as a psychological problem. While there is some freedom with timbre and tempo between SL and TL voice, the age of the speakers should harmonize (Schwarz 2011: 401). In line with this, one would suggest that not only the age of the speakers should harmonize but also their gender. Would a documentary on Arnold Schwarzenegger be at all believable if his original voice were voiced-over by a woman? Or would the reality show *America's Next Top Model* have Tyra Bank's voice voiced-over by a man in the target version? As Orero (2006b: 175) puts it "if the person speaking is a middle-aged man, the voice-actor usually matches the genre (sic.) and the age." Orero (2006b: 175) discusses the case of IMS (Independent Media Support Group plc.), an agency known for trying to match voices according to the expectation of the target audience. The role and relevance of voice quality within voice-over will be discussed further on in this work, and as will be seen, the synchrony presumed in this case is not always there.

Of these six synchrony requirements of voice-over, Orero (2006a and 2010) only considers four (kinetic, action, isochrony and literal synchrony) when characterising voice-over, leaving out content and character synchrony. In the present study, however, the quality of the voice that delivers the translation (character synchrony) together with how close the content of the source and target texts will play an important role and will be considered paramount. In our analysis, in addition, one more category will be introduced: prosodic-illocutionary synchrony.

After presenting the main features of the audiovisual text, along with some of the most relevant issues concerning audiovisual and multimedia translation, including its different modalities, and paying special attention to voice-over, chapter 2 will turn to the relevance of the phonetic (suprasegemental) component of language in audiovisual translation and how it is a substantial element for pragmatic and communicative purposes, and for effective translation. We will thus have the necessary theoretical background (both from the discipline of AVT and

from the study of [English] phonetics) to proceed to the examination of our specific object of study, namely, the role of intonational features in English-Spanish voice-over.

2   TRANSLATING INTONATION

2.1   The relevance of phonetics and phonology in audiovisual translation

Translation has traditionally dealt with words and the written text, the major problem being to find words in a target language that render the meanings stated in the source language (Mayoral, Kelly & Gallardo 1988). This implies that the decisions of the translator were essentially linguistic and textual/verbal, without paying attention to any non-verbal element. This problem might be due to the lack of awareness of the existence of other text-types (Zabalbeascoa 2008: 23). Nonetheless, all this changed with the arrival of audiovisual texts, which combine different elements to the same degree of importance (verbal, nonverbal, audio and visual) which are to be regarded as inseparable for a "fully satisfactory communication event" (Zabalbeascoa 2008: 25).

There is one fundamental part of the English language (and languages in general) that has been largely neglected, and not very much researched on (Mateo 2014) in relation to translation: its phonetic component.

The role played by phonetics in any language is of crucial importance within communication. And it does not only refer to how words are pronounced or the different segmental realisations that can occur in a given language considering its regional and social variations, in terms of what is sometimes considered as the standard of pronunciation; the suprasegmental and prosodic elements of a language are probably more substantial for communicative matters (at least in English) than the phonetic/segmental realization of individual words. Thus, the suprasegmental richness of English, characterised by the different intonation patterns and their functions, the roles of rhythm and stress, and processes like the accenting and de-accenting of new and given information, among others, are worth researching on in audiovisual translation since these features usually play a crucial role in the communicative value of an English oral text and do not always have a similar realization, or function, in the target language. In addition, and focusing now on the translation mode which is the object of study here, voice-over, if we consider the oral component and the time

limits/constraints that characterise it, the treatment that these phonetic and phonological features receive in the process becomes highly interesting since it is more than possible that they are often ignored. This is one of the core issues of this work.

There are two main aspects within phonetics that are worth considering in audiovisual translation. As was mentioned above, the first aspect has to do with the segmental elements (the phonemes), which have mostly been studied in the field of dubbing for lip-sync matters. Fodor (1976) was a pioneer in this respect and coined the term *phonetic synchrony*, which is nowadays commonly known as lip-synchrony, or lip-sync, and consists of adapting the translated script to match the movement of the actors' lips when they are on screen, especially in close-ups. The translation, thus, replaces the source text while somehow reflecting the original realisation of some of the source text sounds. This way, when watching a film, viewers will not be distressed because the original actor's lips match the kind of lip movement that is produced by the dubbing actor in the target language. Lip-sync works mainly with words that have a similar segmental realization, like *no*, and *yes* and *no* and *sí* in Spanish (see table below), but it does not for other kinds of utterances that may be relevant in the film like *don't go* and *no vayas*, in which there are two close round vowels in English and one close round vowel but two mid-open half-spread (unround) vowels, and one more syllable, in Spanish.
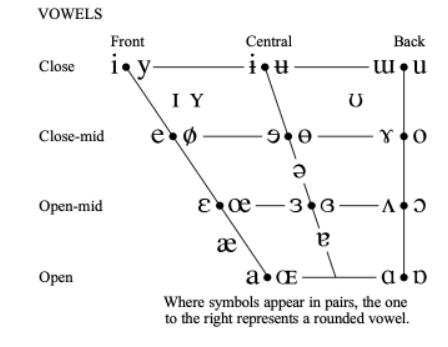
ENGLISH *YES*	/e/ mid front	
SPANISH *Sí*	/i/ close front	
ENGLISH *No*	Br En /əʊ/ centring diphthong Gen AM /o/ half-close back	
SPANISH *No*	mid back [o] half-open central /ɐ/	

Figure 7.11 Comparison of Spanish and English vowels at segmental level

The second aspect concerns the suprasegmental, or prosodic, elements; which have barely attracted any attention either by AVT researchers or practitioners, although more and more studies are now being published on this issue. Prosody encodes both linguistic and

---

11 The phonetic symbols used belong to Quilis (2012) for Spanish and Wells (1982) for English, and the vowel chart was retrieved from the IPA website.

paralinguistic meaning (Wharton 2012: 102), which is crucial in spoken discourse. It seems of vital, and indisputable, importance and indisputable, to incorporate such a relevant aspect of human communication in the field of translation, and even more in the field audiovisual translation, in which oral texts occupy such a relevant place. Prosody, and intonation, allow us, for instance, to *mean* the opposite of what we are saying; "human linguistic communication has the ability to understand the behaviour of others in terms of the intentions behind it – sometimes known as the `mindreading´ ability" (Wharton 2012: 104), and suprasegmental features are often crucial cues to those intentions. Intonation is therefore a paramount issue which needs to be properly dealt with in translation, if the communicative value of the source text is to be transferred adequately. Errors, or oversights, in translating intonation might have fatal consequences in the target version with loss of meaning or of relevant information.

2.2   Analysis of English intonation: form and functions

2.2.1   The form of English intonation

Spoken discourse has no punctuation marks, unlike written discourse, to show when a sentence starts and finishes, to show questioning, or to show surprise, which we indicate with an exclamation mark, for instance. All these are conveyed through the use of melodic patterns and other phonetic elements which can be considered within what we call "intonation".

There are several phonological models for the study of English intonation, being the most relevant the British School, also known as the School of London (Granato, 2014: 197), and the American School. Other scholarly traditions for the modelling of intonation include [12] the Dutch School, the Pierrehumbert School and Fujisaki's Model, among others. The main differences between the British School and the American School are the way of describing tones and how to interpret intonation phrases (Estebas-Vilaplana 2014b). Sánchez Mompeán (2017) adds one more difference: the way they represent pitch movements.

The American School proposes analysing intonation by means of *tone levels*, or *tonal targets*, which can be categorised into 4 types, following Pike's (1945) model: 1 (extra high),

---

[12] For a detailed description of phonetic modelling see Taylor 1992, "A Phonetic Modelling of English Intonation".

2 (high), 3 (mid), and 4 (low). Some scholars, such as Roulon S. Wells (1945), use the complete opposite categorisation, being 1, low, and 4, extra high, as stated in Ladd (2015: 119). The American School does not include nuclear tones, unlike the British School, to describe attitude; what they use are *boundary tones,* or *terminal junctures*, which occur at the final edge of an intonation phrase. Trager and Smith (1951) identify three types of terminal junctures: a single bar juncture [/] for sustention or levelling, a double bar juncture [//] for rising, and [#] for falling. For these scholars, any utterance made in English ends in one of the terminal junctures, and minimal complete utterances, also called *phonemic clauses*, have no other terminal junctures within them (1951: 49-50).

American School symbology				
Pitches or tone levels	1 Extra High	2 High	3 Mid	4 Low
Terminal junctures	/ Levelling	// Rising	# Fall	

Table 5. The symbols of the American School to describe intonation

Since the present study will follow the British School in the analysis of intonation within the chapter devoted to corpus analysis, let us have a look at the main tenets of this school in more detail. The British School proposes an intonational modelling based on two configurations: the nuclear configuration and the pre-nuclear configuration (Estebas-Vilaplana 2014b), and suggests dividing each intonation phrase, or IP, into the following parts: *pre-head*, *head*, *nucleus*, and *tail*. IPs can be made up of one single word *yes!*, or by several words *Can you help me?*. The only rule applying to all IPs is that they must have one intonation nucleus only, which is the most prominent syllable, "typically has a marked change in pitch, and is somewhat longer and louder than the rest" (Collins & Mees 2013: 142). The British School includes seven types of nuclear tones, compared to the four of the American School, which are: *high fall*, *low fall*, *rise-fall*, *mid level*, *high rise*, *low rise*, *fall-rise*.

The nucleus is used to mark the information that is relevant in each IP, normally that which is new. The pitch pattern carried by the nucleus is the *nuclear tone*; this begins on this syllable and will finish there if that syllable is the last syllable in the IP. Otherwise, the tone will be completed throughout the syllables following the nucleus, which are known as *the tail*

of the IP. The elements occurring before the nucleus within an IP are called the *head* and *prehead*: the former comprises the elements within an IP from the first accented syllable up to and including the syllable preceding the nucleus; the *prehead* contains any unaccented syllables before the *head*.

It was a ˈrea lly ˈ**bo** ring ˈgame
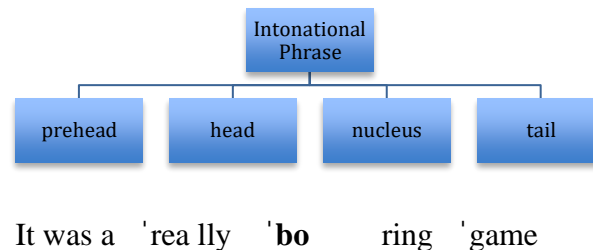
Figure 8. Example of the structure of an IP according to the British School.

Let us analyse the structure of IPs in detail starting with the first possible component: the *prehead*. There are two types of preheads: low and high. Low preheads are the most common type, usually uttered in a comfortable low pitch, and are known as *unmarked preheads* (meaning that they do not stand out in any way, or convey any special connotations); they are also unmarked when transcribing intonation (Ashby 2013), and can be of any length. High preheads, on the contrary, are usually uttered in a relative high pitch, approximately as high as the beginning of a high-fall nuclear tone (Ashby 2013), making them quite noticeable to listeners. They are known as *marked preheads* for this reason, and are less frequent than low preheads. High preheads are transcribed using the only tonetic symbol that is not also a stress mark: [ˉ] (Ashby 2013); it is placed above and before the first syllable of high preheads. Let us compare high and low preheads in the following example, from Ashby (2013: 44):
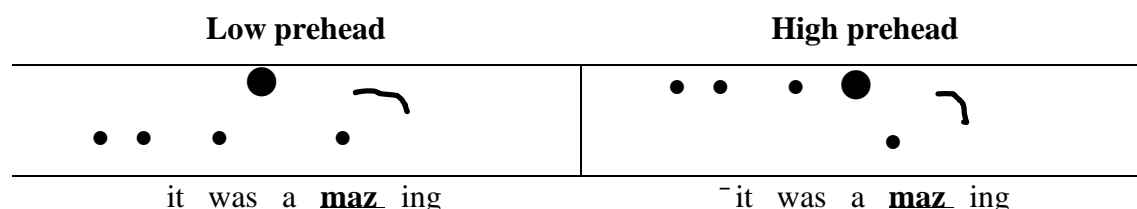
Figure 9. Examples of high and low preheads

Moving now to the *heads*, there are different ways of categorising them. For Wells (2006) we can distinguish between two types of heads: simple and complex, for which there are four types respectively; and according to Ashby (2013), heads can be basic and emphatic, with four and three types respectively. Despite the fact that these types established by Wells and Ashby have different names, their meaning and purpose is practically the same.

Simple or basic heads are those which include only one accented syllable, although there can be more stressed syllables13 within them. Ashby's basic head structure is based on O'Connor & Arnold's (1973) model. Both simple and basic heads can be divided into four categories, as follows:

Simple & Basic heads	
Wells 2006	**Ashby 2013**
High level	High head
High falling	Falling head
Low rising	Rising head
Low level	Low head

Table 6. Types of heads according to Wells (2006) and Ashby (2013)

High level or high heads are uttered in a high pitch, though not extremely high. For Ashby the high head is probably the most commonly heard type of head in English and combines with all nuclear tones except the fall-rise. The examples below show how Ashby and Wells represent the intonation of these types of head. Note how the first syllable in them is indicated by a tonetic stress mark ['], which is the same as the symbol used to mark sentence stress. In Ashby's and Wells' representation, any subsequent stresses are marked using the small raised zero [º] symbol14.

13 Stressed syllables are strong syllables (those not including a weak vowel /ə, u, i/) made prominent in terms of loudness, as in *water* /'**wa** ter/ and *computer* /com **'pu** ter/. Stressed syllables can be made even more prominent by means of pitch, thus making them accented. An accented syllable is a stressed syllable made prominent by pitch.

14 Rhythmic stresses, which are irrelevant for intonation, are marked with the zero symbol [º], or not marked at all.

Example 5: Wells' representation of a high level head (2006: 209):

Well "make up your `mind.
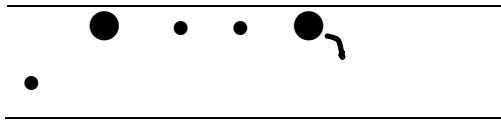
Example 6: Ashby's representation of a high head (2013: 44):

I'll "try to °ring you °back this `eve ning.
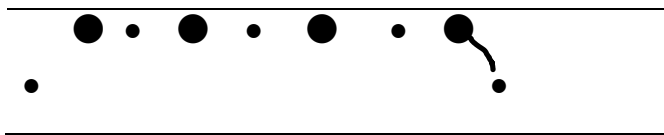
High falling or falling heads start with a fairly high pitch, which falls gradually, syllable by syllable, to a mid pitch. These types of heads are usually followed by a fall-rise nuclear tone. The first syllable of these heads is indicated by a back-slash mark [\]. In Ashby's representation, any subsequent stresses are marked using the small raised zero [°] symbol.

Example 7: Wells' representation of a high falling head (2006: 209)

Well \make up your ˇmind.
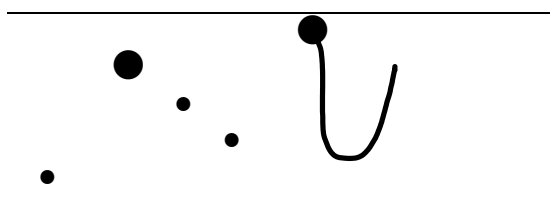
Example 8: Ashby's representation of a falling head (2013:45)

I'll \ try to ºring you ºback this ˇeve ning
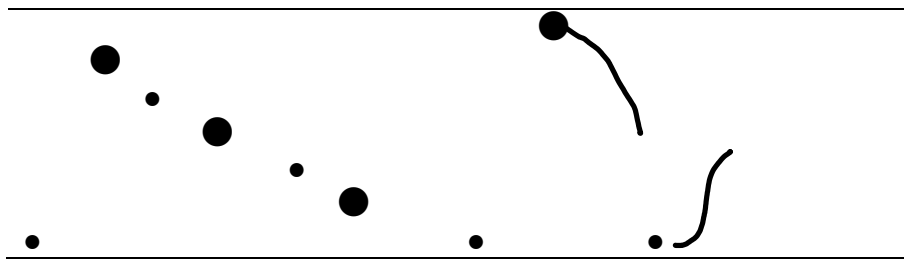
Low rising or rising heads start with a fully low pitch which rises gradually, syllable by syllable, to just below the starting pitch of the high fall nuclear tone. These heads are indicated with the forward slash [/] tonetic symbol just before the first stressed syllable. In Ashby's representation, any subsequent stresses are marked using the small raised zero [º] symbol showing that there is no new pitch information (the pitch is following the instruction issued by the initial [/]).

Example 9: Wells' representation of a low rising head (2006: 209):

Well /make up your `mind.
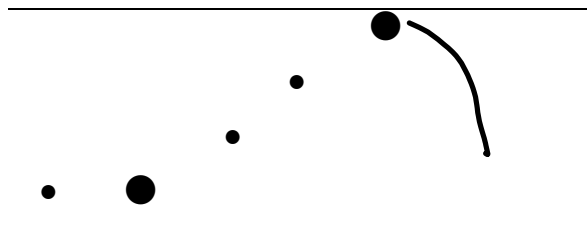
Example 10: Ashby's representation of a rising head (2013: 45):

I'll /try to ºring you ºback this `eve ning.
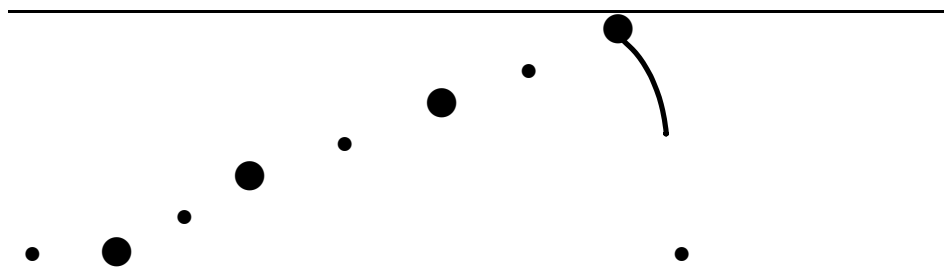
Low level and low heads are uttered with a fully low pitch, and are commonly followed by the low-rise nuclear tone. The first syllable of these types of heads are indicated with the tonetic symbol [ˌ]. In Ashby's representation, any subsequent stresses are marked using the zero [0] symbol at the bottom line showing that there is no new pitch information (the pitch is following the instruction issued by the initial [ˌ]).

Example 11: Wells' representation of a low level head (2006: 209):

Well ˌmake up your /mind.
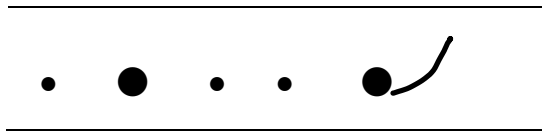
Example 12: Ashby's representation of a low head (2013: 45):

I'll ˌtry to ring you back this /eve ning.
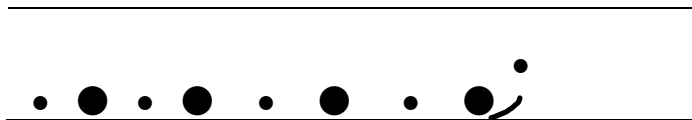
The communicative value of the different types of head will be seen later on in this work since it depends on how the heads can be combined the nuclear tones. Complex heads are, according to Wells, those which include more than one accent, unlike simple heads. Emphatic heads are similar to complex heads, and differ from basic heads in the fact that while simple heads could have more than one stressed syllable, by means of rhythmic stress [º], in emphatic heads these rhythmic stresses get pitch prominence and become accented [15][']. There are three main types of complex/emphatic heads: *stepping heads*, *climbing heads*, and *sliding heads*. Wells includes one more category, *mix complex heads*, which are a combination of different heads within the same IP, and whose description will not have a place in this work.

Stepping heads, *complex high level heads,* according to Wells 2006, are the emphatic variant of high heads or high level heads. The head starts with the same high pitch as a high head/ high level head but at each subsequent stress, the pitch goes down a step creating a pitch

[15] In O'Connor & Arnolds (1973:19-20) the tonetic symbol [º] is used to mark accent and not rhythmic stress, unlike Ashby and Wells. In p.35 the use of high and low zero symbol is similar to Ashby's and Wells'.

change thus accenting the syllables. Each stress is marked with the tonetic symbol ['] to show a new level of pitch on each stressed syllable.

Example 13: Example of a stepping head (Ashby 2013: 46):

I'll 'try to 'ring you 'back this \eve ning
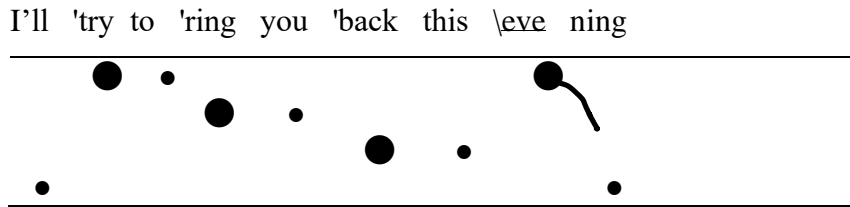
Climbing heads, *complex low rising heads* according to Wells (2006), are the emphatic version of the rising head. Each stressed syllable is identified with the tonetic symbol [/] in order to show that on each new stress, the rising accent starts again. In these types of heads, each stressed syllable is uttered in a slightly lower pitch than the preceding unstressed syllable. Both stepping and climbing heads are commonly used before a high fall nuclear tone.

Example 14: Example of a climbing head according to Ashby (2013: 47):

I /ne ver i /mag ined she'd be /so up \set
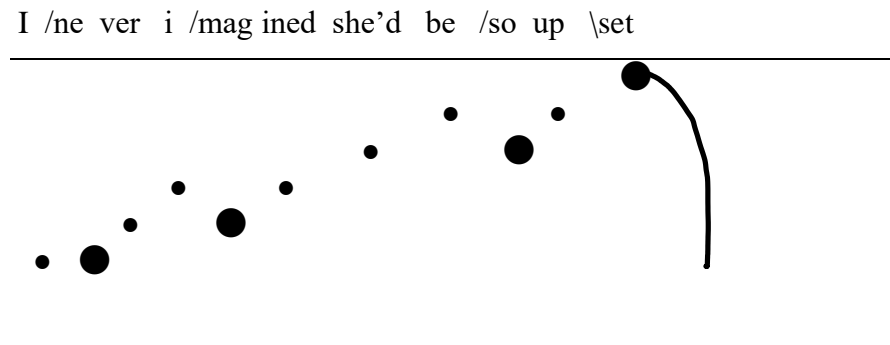
Sliding heads, or *complex falling heads* according to Wells (2006), are the emphatic version of falling heads/high falling heads, and typically precede a fall-rise nuclear tone. Each stress in these types of heads is marked with the tonetic stress [\] and each stressed syllable is higher in pitch than the preceding unstressed syllable.

Example 15: Example of a sliding head according to Ashby (2013: 47):

I \ne ver i \mag ined she'd be \so up ˇset
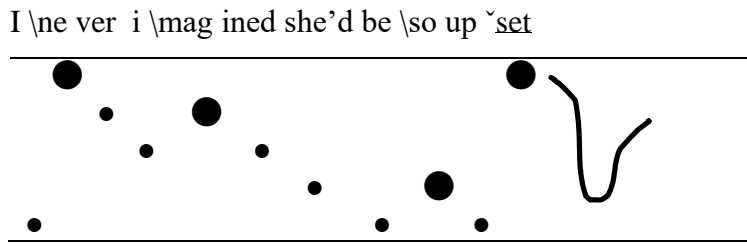
The most relevant part within an IP is the *nucleus*, which is the last accented syllable[16] within the IP, and carries an identifiable pitch movement, or tune, known as nuclear tone. This melody, or tune, continues throughout the following syllables, if any, which are known as the *tail*. There are seven nuclear tones in English which are identified by their own tonetic symbol: high fall [`], low fall [°], rise-fall [ì], mid level [–], high rise [æ], low rise [,], and fall-rise [<]. Let us delve into the different nuclear tones possible in English.
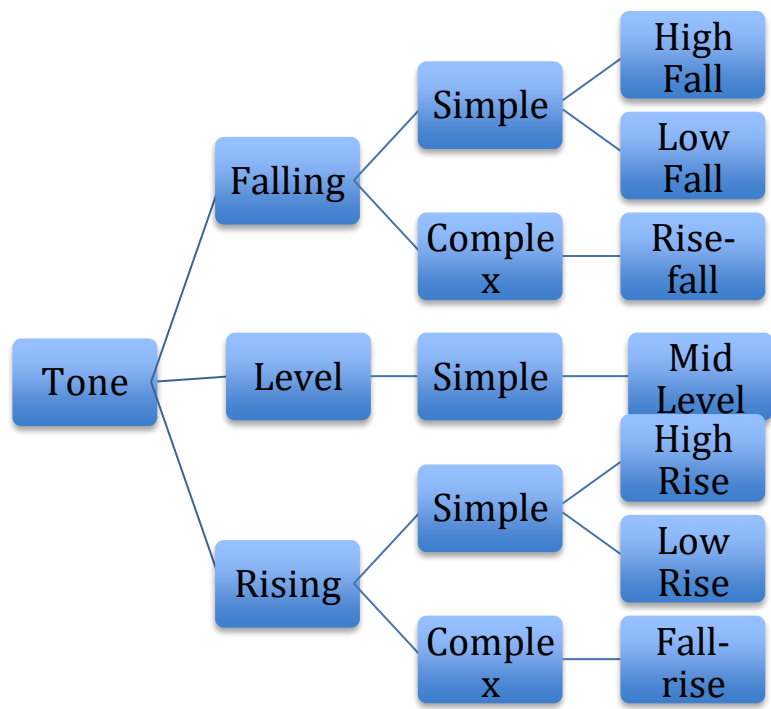
Figure 10. Types of nuclear tones according to the British School

The meanings of the nuclear tones in English depend mainly on the context in which they are used, since all intonation is context-sensitive (Ashby 2013: 38). However, there are some generalisations on the different functions they serve as can be seen in the chart[17] below.

---

16 The phonetician Geoff Lindsey (1981: 18) states that tone is a property of the whole sentence rather than a syllable, and thus his placement of tone symbols before the tonic Word in transcription is arbitrary, with two invented unitary nuclear symbols within his intonation system.
17 Chart taken from UCL's SCEP (Summer Course in English Phonetics) 2013 handbook p.38-39

The identification of the meaning of the tones is paramount to identify and interpret the tones of the examples analysed in our study. With this information we will be able to discuss how accurately the meaning and communicative value of the original tones in English are rendered in the Spanish voice-over translation.

Falling tones	
Typical semantic associations:	All normally signal definiteness and completeness.
Typical grammatical associations:	Are all typically used with declarative utterances/statements
Typical attitudinal associations:	These vary from tone to tone (see below).
Tone-specific associations:	**High fall** is also typically used with *wh-*questions; signals a lively and engaged attitude on the part of the speaker. Probably the most used of all nuclear tones.
	Low fall also typically used with *wh-*questions; often signals a more subdued attitude on the part of the speaker (low spirits, lack of interest, boredom, etc.) Maybe used rather less frequently than the high fall.
	Rise-fall is also sometimes used with questions but not specifically *wh-*questions like the simple falling tones. It often conveys an attitude of assertiveness or authoritativeness – has a challenging ring to it, the speaker disclaiming responsibility; can convey (wry) humour, can sound argumentative. Note that this is probably the least frequently used of all nuclear tones.
Level tone	
Typical semantic associations:	Incompleteness, marking non-finality or (maybe) intentional vagueness on the part of the speaker.

Typical grammatical associations:	**Mid level** is typically used with statements (usually incomplete fragments), interjections, filled pauses (hesitations phenomena).
Typical attitudinal associations:	None.
Rising tones	
Typical semantic associations:	All can signal either completeness/finality or (perhaps more often) incompleteness/non-finality.
Typical grammatical associations:	All associate with declarative utterances, statements. The simple rising tones associate with questions, especially inversion questions.
Typical attitudinal associations:	These vary from tone to tone (see below)
Tone-specific associations:	**High rise** also typically used in echo-questions and in statement questions. Can sound rather casual and sometimes indicates a tentative attitude on the part of the speaker.
	Low rise often carries a "sub-text" – in questions it can be disbelieving and in both statements and questions can sound (very) critical. However, this tone can also carry soothing, reassuring and encouraging connotations.
	Fall-rise, one of the most frequently used nuclear tones, can signal an unsure/tentative attitude. It is very often used for offering *polite* contradictions and corrections. It can be used with echo-questions.

Table 7. tone types and their communicative effects (SCEP: 2013)

To conclude with the different components of the intonation phrase, tails are optional units in IPs and can include stresses, but never accents since they simply follow the movement started by the pitch change in the nuclear tone, and therefore, cannot include any new change in pitch. They can be of two types: *rising*, or *low* (nuclear tones with a mid level pitch have mid level tails). In addition, some IPs may have no tails, i.e. when the nuclear accent is the very

last syllable in the intonation unit. The chart below shows the possible combination of heads, nuclear tones and tails, permitted in the O'Connor & Arnold's descriptive framework.

Head/N Tone	HF	LF	RF	LR	HR	HF+LR	FR	ML
High	✔	✔	✔	✔	✔	✔		✔
Falling							✔	
Rising	✔							
Low				✔				
Head/Tail	Low Level			Rising				Level

Table 8. Summary of the head + nuclear tone + tail combination

Native speakers of English play with IPs at their own discretion when they speak to convey, understand, and chunk the length of their messages, highlight the relevant information and show their attitude. In other words, they use all these formal elements of intonation to communicate different values. The following section will therefore describe the strategic functions implicit in every IP, which are transmitted by means of intonational features.

2.2.2 The functions of intonation

Intonation in English serves different purposes, or functions, which can be summarised in four according to most scholars (Collins & Mees 2013, Roach 2009, Cruttenden 2008, Wells 2006, Tench 1996, Couper-Kuhlen 1986): attitudinal, grammatical, focusing, and discourse. Couper-Kulhen, Tench, and Wells, however, include more functions: indexical and illocutionary for Couper-Kuhlen, and discourse, indexical and psychological for Wells; while Tench also introduces the *sociolinguistic* function, which relates to "the user rather than the use of language" (1996: 29), by means of describing, comparing, and contrasting one accent (or dialect) with another. We will here focus on the main four functions of intonation, although the indexical and illocutionary will be dealt with in a different chapter since these prosodic features are especially relevant to our study.

The British School distinguishes three main elements of intonation which are crucial in oral communication, commonly referred to as the 3Ts (Wells 2006: 06): *tonality*, which deals with the chunking or division of spoken material into IPs (Intonational Phrases), *tonicity*, in charge of highlighting the important information by placing the nucleus in each IP (normally the accented syllable of the last content word), and finally, *tone*, that refers to the tune or melody carried by the nucleus in each IP. These "3Ts" are paramount in our study because they are used to analyse spoken material, such as what information is relevant, what melody is used and what meaning is carried within that tune, and how spoken material is divided. In voice-over translation we can perceive the two languages at the same time, and it can be possible that these intonational features be not adequately transferred, or that they are transferred by means of a different linguistic resource, such as syntax, for example. Chapter 2.5 deals entirely with the translation into Spanish of the "3Ts", but before that, let us analyse the functions of intonation in English and their relevance for translation purposes.

The Functions of English Intonation					
Wells (2006)	Grammatical	Focusing	Attitudinal	Discourse	Psychological	Indexical
Cruttenden (1997)	Intonation groups (Grammatical)	Nucleus placement (Focusing)	English nuclear tones (Attitudinal)			
Collins & Mees (2013)	Grammatical	Focusing	Attitudinal	Discourse		
Roach (2009)	Grammatical	Focusing (accentual)	Attitudinal	Discourse		
Couper-Kuhlen (1986)	Grammatical	Informational	Attitudinal	Textual/Discourse	Illocutionary	Indexical
O'Connor & Arnold (1973)	Grammatical		Attitudinal			
Tench (1996)	Syntactic structure (Grammatical)	Organization of information (Focusing)	The expression of attitude (Attitudinal)	Realization of communicative functions (Discourse)	Textual structure (Discourse)	Identification of speech styles (Indexical)

Table 9. The functions of intonation according to different scholars of the British School.

2.2.2.1 **Grammatical function**

Intonation functions to help distinguish different grammatical structures and constituents by means of tonality (the way the boundary between utterances is indicated), and also tone 18 (syntactic function). Tonality also affects the way we interpret syntactic structures, and for scholars such as Mott (2011: 222), Tench (1996: 21), and Wells (2006: 11), there are several different uses of tonality in this respect. However, with regards the grammatical functions of intonation, in this study we will just mention the ones we consider are relevant for communication in English and for audiovisual translation, for the sake of brevity, and which will be under study. The most relevant aspects of the grammatical function of English intonation which should be considered in voice-over translation are:

Subject and predicate division: In English, it is sometimes desirable to highlight the subject of a sentence by separating it from the predicate so that it gets the maximum attention, and according to Mott (2011: 223) in order to do this, two intonation phrases (IPs) are preferable. (But Anne | continued to be Anne) (The idea of not having an heir | was unthinkable)

Example 16

a) The idea of not having an ˇheir | was `unthinkable
RATHER THAN:
b) The idea of not having an heir was `unthinkable.

In example 16a we can see that there is a boundary that separates the subject (the idea of not having an heir) from the predicate (was unthinkable), thus creating two different IPs with two nuclei and two tones. This way, the hearer can notice that the subject, with its own IP and a high-fall nuclear tone in "heir" is important regarding what comes next. In example 16b this effect is not achieved, and the relevance the subject can have within the context of the conversation is not as important as in a)

---

18 Another component of intonation that can be said to have grammatical significance is the choice of tone on the tonic syllable (Roach: 2009: 155).

Distinguishing between defining and non-defining relative clauses: defining relative clauses normally have the same intonation pattern as the ones shown in the example above (1a) with two IPs, although they can have just one IP as well. However, non-defining relative clauses have three IPs since it is necessary to separate the non-defining clause from the main clause. It could be possible that in the translation process of oral texts as in the case of voice-over this might be overlooked, and the meaning of the original text be rendered with a different one.

Example 17

a) Defining relative clause:

The boys who were at the end of the line failed the test.
The boys who were at the end of the ˇline | failed the `test |

b) Non-defining relative clause:
The boys, who were at the end of the line, failed the test.
The ˇboys | who were at the end of the ˇline | failed the `test |

Example 17a show two IPs and implies that only the boys that were standing at the end of the line failed the text, the rest did not. However, example 17b implies that all the boys failed the test and it happened that there were also standing at the end of the line.

Differentiating questions from *exclamations and statements*: exclamative utterances are said using a falling tone (an exclamatory fall for Wells 2006: 59), while interrogative ones are said (sometimes) using a rising tone.

Example 18

a) More wine? | More ´wine |
b) More wine! | More `wine |

In addition, a statement can be turned into a question by just changing the pitch direction of the tone from falling to rising.

Example 19

a) More wine (question) |More ´wine |
b) What d'you say?   More wine. |More `wine |

These examples above show how we can change the pragmatic meaning of a sentence by just changing the tone. This is quite relevant considering the fact that most of the products dedicated to voice-over translation are unscripted programmes, and as such, they do not have a script like movies, for example, and therefore, transcribers and translators have to be aware of these tunes to identify the right purpose of the utterances before translating them.

Topics, cleft sentences and pseudo-cleft sentences: the first element in a sentence is the *topic*, or *theme*[19], coinciding most of the times with the grammatical subject, while the succeeding is the *comment*, or *rheme*[20]. In English it is possible to topicalise other clause elements apart from the grammatical subject in order to give some sort of *psychological prominence* (Leech and Startvik 1994: 200), and this is usually signalled with an intonation break, according to Wells (2006: 199). Tench (1996: 37) refers to this specific usage of intonation as "marked theme".

Example 20

(Topicalised rheme) *The survival of the dynasty* is what is at stake.

Cleft sentences include a topicalised constituent introduced by *it is* (*it was*, etc.) which carries focused material and consequently an intonation nucleus. If there is further focused material to follow, then there must be an intonation break. (Wells 2006: 199)

Example 21

It's Henry's order, that she's taken out and beheaded.
It's Henry's `order | that she's taken out and `beheaded |

[19] Wang (2007: 166) following Halliday (1985: 38) considers the *theme* is the given information of a message, what is known.
[20] Wang (2007: 166) following Halliday (1985: 38) considers the *rheme* is the new information of a message.

Example 22

It's precisely the magnitude of the charges, that makes them convincing.

It's precisely the magnitude of the `charges | that makes them `convincing.

Pseudo-cleft sentences involve *what* or any other *wh* question word (*who, where*, etc.), and there is usually an intonation break between the two clauses. (ibid). In pseudo-clefts the spotlighted information or rheme is not introduced by the *wh* clause but by the clause with the verb *to be*. They usually start with the *wh* clause so that the new information is placed at the end (Parrot 2010: 321). Example 5 above shows a pseudo-clef whose rheme has been topicalised or fronted. The examples below show a commoner usage in English of pseudo-clefts.

Example 23

What makes Anne interesting and fascinating *is precisely that she's not like Catherine of Aragon*.

What makes Anne interesting and fascinating | is precisely that she's not like Catherine of Aragon |

Example 24

What he sees is conclusive evidence that his wife is not a good woman.

What he sees | is conclusive evidence that his wife is not a good woman |

These types of constructions are quite relevant in English communication considering the emphasis and focusing of information achieved through their structure. From a translating studies perspective, these constructions are quite interesting since there are several ways of translating them into Spanish, as will be seen in chapter 2.4, and in the case of voice-over translation their translation could entail synchrony issues (literal synchrony at the beginning and end).

2.2.2.2 **Focusing function**

Another strategic function of English intonation is, the one related to focusing; native speakers do this by tonicity (see page 18) and the placement of other accents. When

people speak, they produce IPs which may contain one or several words, and not all the words have the same relevance within the IPs. Within each IP native speakers select one word as particularly important for the meaning, normally carrying new information, and on that very word they place the nucleus, or main, most prominent accent, which will bear the nuclear tone or last pitch movement (Wells 2006). In English, the nucleus is usually located on the last content[21] word of the IP, and according to Wells (2006: 95), "by definition, the nuclear accent is the last accent in the IP". Let us analyse the following example:

Example 25

"Meet me in "front of the "pub at \seven.

The natural place for the tonic syllable (the nucleus) will be the first syllable of the word "seven", which is the last content word in the IP. In this example the nucleus indicates the end of the new information. The speaker is not just focusing on "seven" but on the whole plan of meeting in front of the pub at seven. In this case, the focus is *broad*, and embraces the whole clause. When the nucleus or tonic syllable falls on the last lexical item of the IP, the tonicity, or position of the nucleus within the IP, is considered to be neutral, i.e. *neutral tonicity*. (Tench 1996: 56).

There are cases in which only part of the information in an utterance is brought into focus. This is known as *narrow* focus, and old, or given information, is usually left out of focus. Let us see how narrow focus works in the example below taken from Wells (2006: 118):

Example 26

Who's bringing the food?　　　"Mary.

　　　　　　　　　　　　　　"Mary is.

　　　　　　　　　　　　　　"Mary's bringing it.

　　　　　　　　　　　　　　"Mary's bringing the food.

　　　　　　　　　　　　　　It's "Mary that's bringing it.

[21] Content words (nouns, verbs, adjectives, and adverbs) carry the nucleus in IPs, unlike function words (pronouns, prepositions, auxiliary verbs, and articles), which are only tonic in special cases like narrow focus.

There may be some cases in which the context of the conversation makes the speaker change the tonic syllable, and the focus, of the utterance. Let us consider the following conversation, using the sentence shown in example 27 below.

Example 27

"Meet me in "front of the "pub at \seven

Example 28

Ok, I'll see you in front of the theatre at seven.

Example 29

No |"Meet me in "front of the \pub at seven |

Example 30

Ok, I'll see you in the pub at seven.

Example 31

No |"Meet me in `front of the pub at seven |

In the examples above the focus does not fall on the last content word, as in broad focus, but on specific words the speaker wants to highlight and bring into focus. This is called *contrastive* focus and is a kind of "narrow focus", since the utterance contains both new and old information.

These are cases of *marked* tonicity, and it normally occurs in situations in which the nucleus, i.e. the tonic syllable, does not fall on the last lexical item of the IP, its very last word. However, there are cases of narrow focus with neutral tonicity, in which the nucleus does fall on the very last word of the IP. These are the cases in which all the information preceding the nucleus is old or given information. The examples below exemplify how this works.

Example 32

a) I think I'll go and have a cup of \tea.
b) (Well) why don't you come and have a spot of \lunch?

The example 32b above (Tench 1996: 59) shows narrow focus with neutral tonicity. This can be explained from the context since the only thing which is new information is "a spot of lunch"; "you come and have" refers to the "I'll go and have" from 32a in the previous IP. Therefore, given that the nucleus should indicate the end of the new information, "lunch" should carry the nucleus in this case.

Example 33

a) Look at that man up `there.
b) And look at `that man.

In this case, example 33b (Tench 1996: 60) shows narrow focus with marked tonicity since the nucleus does not fall on the last word of the IP. The explanation for this is that "man" and "look" in b) are anaphoric; they are given or old information from the previous IP, and as such they are deaccented. The case of deaccenting anaphoric elements in English oral communication and its relevance for translaton will be dealt with in chapter 2.5.2

As it can be seen from the examples above, given the flexibility that the English language shows when it comes to emphasise the relevant information in oral communication, it is worth considering how this is reflected in Spanish voice-over translation for different reasons. One reason can be found in the fact that Spanish intonation resources to mark focus are not as flexible as they are in English. There are other resources characteristic of Spanish that apply in order to achieve the same communicative effect as in English. However, these could be obviated in some cases if the translator is not aware of the way intonation functions in English.

Tonicity, that is, the placement of the nucleus in an IP, is normally influenced by whether the words in the utterance bring new information, or not. In English, old information, or information that is already given, or repeated, is *destressed* and therefore *deaccented*. Thus, the placement of the nucleus marks "the end of the new information in an intonation phrase" (Wells 2006: 107). This rule of thumb even applies to any repeated words, or near synonyms, in a conversation, since that means that the information is already known, considered as given, and therefore it is not accented.

Example 34

a) And that would be like a\dultery, like \gross adultery.
b) The king and queen would sit in the royal Pew above the \Chapel, in the \body of
the Chapel.

As can be seen in the two examples above, both "adultery" and "the Chapel" are deaccented in the second IP because they are old information, i.e anaphoric[22], and do not provide anything new to the context of the conversation.

However, this deaccenting rule does not work every time a word is repeated, or already given in a conversation. There are cases in which old information is brought into focus with marked tonicity. One of these cases are *echoes*. Echoing consists in repeating the same word someone has said, and it is usually brought into focus when the second speaker wishes to show disbelief, surprise, or just comment on that. The example below shows how echoes are given marked tonicity:

Example 35

a) In your view, which are the positive aspects of being in the EU?
b) I don't know a lot about the \positives of being in the EU. All I really hear about is
the negatives.

The example above shows how 35b reaccentuates "positives" to highlight it as the focus and then make a comment about it.

Another case in which given information is reaccented are *insists*. English speakers may reaccent old information in order to highlight and emphasise that information again, and therefore insist that what they said was important.

Example 36

a)   But \Anne, continued to be \Anne.
b)   The final driver of anything under \Henry, is \Henry.

---

[22] The destressing of old information is also called Anaphora Rule by Mott (2011)

In example 36a the speaker is talking about how Anne Boleyn's behaviour could have been the cause of her own death. A behaviour that she did not change when the rumours of her punishment started. For that reason, the speaker emphasises and accentuates "Anne" in both IPs. The second example (36b) shows how "Henry" is reaccented in the second IP, insisting on the fact that Henry is responsible for everything happening within his court.

Once again English intonation shows how resourceful it is when it comes to emphasise and de-emphasise information in fluent oral speech. These resources might not be present in Spanish and the way these are translated can be relevant if, for example, what is brought into focus in the target text is the information that the speaker has consciously and intentionally deaccented in the original text. The problems that these intonational features can carry in audiovisual translation will be dealt with and exemplified in detail in chapter 2.4.

2.2.2.3 **Attitudinal function**

This is probably the best-known function of intonation, and for some[23] it is the most important. It is not what we say, but how we say it, that matters. Emotion can be linguistically signalled using words and syntax (*Could you please give me the money? /Hand over the bloody cash!*), or it can also be signalled by means of using different tones and pitch[24] levels so as to sound bored, interested, happy, sad, excited, anxious, etc. Our attitude may not always be the same when we speak and that affects the way the hearer interprets our message. Wharton (2012) considers that the natural aspects of prosody "encode information that conveys information about emotional states about emotional states and attitudes, or creates impressions" (2012: 110). Assigning meaning to English tones is a troublesome business for several reasons and scholars such as Wharton (2012), Roach (2009) and Wells (2006) agree with this. For example, in addition to the tone we use, there are psychological and physical factors that affect the

---

[23] Kate Scott (2013: 64) considers the attitudinal function as the most important function of intonation.
[24] Pitch refers to the frequency with which our vocal folds vibrate (Rose 2002: 292). The higher the frequency, the higher the tone, as for example a girl's voice. Contrarily, the lower the frequency, the lower the tone, as in the voice of baritone and bass opera singers.

way we speak; everything sounds different when we smile25, that's why smiling is a must do for people working in call centres; if we are agitated, we may speak fast, and the opposite if we are tired. All these are factors that affect our attitude and the meaning of what we say in addition to the tone we choose for each IP.

The impossibility to demonstrate 100% that there is one default tone for any sentence type is claimed by Wells (2006: 91), for whom the burst of uptalk26 into everyday English has changed the intonational landscape for many native and non-native speakers of English, which implies that some of the generalisations about the correct usage of tones can be debatable nowadays.

Nonetheless, despite the possible contradictions of assigning specific meanings to tones, there is a general agreement within the literature of phonetic science and EFL (English as a foreign language) that tones have a "given" meaning depending on the context. Some of the generalisations about the meaning of tones, which include attitudinal meaning, have already been mentioned in section 2.2.2 (see Table 9) above. The charts below show some more generalisations focusing on the attitudinal function of English nuclear tones according to Roach (2009) and Collins & Mees (2013), Wells (2006), and Couper-Kuhlen (1986).

ROACH	
Tone type	**Meaning**
Fall	Finality, definiteness
Rise	General questions, listing, more to follow, encouraging
Fall-rise	Uncertainty, doubt, requesting
Rise-fall	Surprise, being impressed

Table 10. Tone types and their meaning in Roach (2009: 117-148)

---

25 This online article shows how important it is to smile on the phone for people working at call centres: https://blogs.20minutos.es/comunicacion-no-verbal-lo-que-no-nos-cuentan/tag/sonrisa-telefonica/

26 Uptalk refers to the usage of a rising tone in statements in which a fall would be expected.

COLLINS & MEES	
Tone types	**Meaning**
High fall & Low rise	Neutral
Low fall & high rise	Reinforce, emphasise and exaggerate the speaker's basic attitude
Fall-rise	Doubt, correction, reservation, appealing the listener to reconsider.

Table 11. Tone types and their meanings according to Collins & Mees (2013: 146)

The two tables above are quite basic and only show the tone types and their generalised meanings. Wells' table, shown below, is more complete as it includes the sentence type associated to each tone and meaning.

Wells 2006		
Tone type	**Meaning**	**Sentence type**
Fall	Definitive (includes exclamatory)	Statement Exclamation Wh question Answer Command Interjection
	Insistent	Yes-no question (includes tag question and elliptical question)
	Reinforcing	Adverbial
Non-fall (=fall-rise or rise)	Dependent (includes open limiting)	Clause or smaller element

Fall-rise	Implicational (includes contrastive, reservation, tentative, polite, correction, partial statement, negative, warning)	Statement Command
Rise	Encouraging (includes soothing)	Statement Wh question Command
	Non-supportive (includes truculent, perfunctory)	Statement
	Yes-no (includes pardon question, uptalk)	Yes-no question Tag question Indep. Ellip. Question Declarative question Statement (uptalk) Pardon question interjection

Table 12. checklist of tone meanings according to Wells (2006: 91)

Couper-Kuhlen's (1986) table below is probably the most complete in terms of representing the relationship between the melody of the tones, the attitude and the physical and psychological state of the speaker. However, it does not include the sentence type as in Well's.

Fig. X/2 Reported prosodic correlates of selected emotions and attitudes

Prosodic feature / Attitude or emotion	Pitch level (average)	Pitch range		Tone	Loudness	Tempo	Other
		Intervals between tones	Width of glides				
Anger	high (D,W/S) low (C,H)	greater than neutral (W/S) / steps down (C)		simple falling (C)	loud (C)	fast (C)	strong stress (C) high unstressed syllables (C)
Amusement				rise–falls (C)			
Boredom	low (D,H)			level (C)	soft (H)	slow (H)	
Complaint				semi-tone rises (F/M)			smooth melody rhythmically interrupted (F/M)
Confidence	high (S/L/W)				loud (H,S/L)	fast (H,S/L)	
Coquetry					suppressed (F/M)		soft tertial up-glide on last stressed syllable (F/M)
Excitement		high step-ups (C)	wide (C)	falling and rise–fall type (C)	loud (C)	fast (C)	
Fear	lower than for anger (W/S) mid (F/M)	extremely narrow (F/M) occasional high peaks (W/S)				duration longer than for anger (W/S)	rigid melodic line (F/M)
Happiness (= joy)	high (D,H, F/M)	increased (F/M)		frequently ascending at irregular intervals (F/M)	loud (D,H)	fast (D) lively (F/M)	irregular stress distribution (F/M)
Longing		narrowed (F/M)		slightly rising, descending, gently ascending finally		restrained (F/M)	
Pleasure	high (H)						
Puzzlement	high (C)	high step-ups (C)	wide (C)	rising type (C)	piano (C)	lento (C)	
Sadness (= sorrow, grief)	low (D,H) lower than neutral (W/S)	narrow (W/S)			soft (H)	slow (W/S)	
Surprise		increased (F/M)		sudden fall (F/M)			
Tenderness	high (F/M)	narrow (F/M)			restrained (F/M)		audible off-glide in long stressed syllables (F/M)
Timidity					soft (H)	slow (H)	

C = Crystal 1969, D = Davitz 1964, F/M = Fónagy/Magdics 1963, H = Huttar 1968b, S/L/W = Scherer/London/Wolf (1973), W/S = Williams/Stevens 1972

Table 13. Check list of tones and attitudes in Couper-Kuhlen (1986: 181)

How can concepts such as attitude and emotion be translated in audiovisual translation and voice-over? Unlike in literary translation, in which punctuation shows clause types and certain comments can be included in brackets to show psychological states (happiness, sadness, excitement, etc.) as in *I'm so sorry. (he said bursting into tears),* in voice-over translation that is not possible. In fact, that is one of the advantages of voice-over as a translation mode, because the audience is provided with a twofold translation. On the one hand there is the translation of the original oral text, which is done in writing; and on the other hand, there is the oral delivery of that written text, in which the voice talent can add the attitudinal and emotional touch. However, there are situations in which the pragmatic value condensed in a high-fall nuclear tone, for example, together with some physical conditioning factors such as tempo or loudness, may not be adequately translated in the target text. For example, if someone says "it's wrong" with a high-fall nuclear tone on "wrong", fairly loud and at a more slowly pace than normal, a suitable translation into Spanish would be "está muy mal" rather than just "está mal", which is the literal translation. It is to this intonational nuances that translators should pay attention to provide the target audience with a faithful and reliable translation. The translation of the pragmatic, and illocutionary, force implicit in English intonation will be dealt with in Chapter 2.4.

2.2.2.4 Discourse function

English intonation is said to have a discourse function, also known as *cohesive* (Wells 2006), and *textual structure* (Tench 1996), which is in charge of giving coherence, comprehensibility, and a structure to spoken discourse, very much like punctuation in writing. In a conversation, the discourse function also signals turn taking or whether someone has finished speaking; and it allows the hearer(s) to identify which information is relevant, overlapping with what has been termed "the accentual function"

Both tonicity (or the position of the nuclear accent) and tones are used in English to indicate which information is relevant in an utterance, as seen in Section 2.2.2.2. The rule of *destressing* given information, or anaphora rule (Mott 2011: 205) signals which

part of the utterance shall be taken as new/relevant and which as old, as in the following example:

Example 37

No woman had ever made that step from royal mistress to the throne, getting the \Queen, a \real Queen, out of the way.

As regards to tones, rising tones indicate non-finality and imply that information is sought, or anticipated, rather than unloaded, whereas falling tones suggest the opposite, that is, finality and allow the hearer to understand that no more new information is coming. (Collins & Mees 2013)

Example 38

We don't know to what extent she /loved him, if she ever /did, or if she operated on a basis of cold am\bition.

In the example, the rising tones allow the hearer realise that more information is coming next, until the falling tone is heard, showing, therefore, the end of the message.

Another way of focusing information by means of intonation is done in English by modulating the pitch and the speed of the voice. By doing this, some IPs (the important ones) are given more prominence, and therefore more relevance, than others in a conversation. This intonational resource is known as *intonational subordination* (Roach 2009: 158). Wells (2006: 243) calls this resource *parenthetical intonation* because it is as if putting the IP in brackets as in writing, and the IPs which are subordinated are referred by Mott (2011: 229) "throw-away IPs" since they do not provide any relevant information to the message. Sometimes, these subordinated IPs are uttered using a lower volume and a faster pace than the relevant ones.

Example 39

a) Anne is \not, ↓ as she had \hoped ↓, Henry's front line political ad\visor.

b) Cromwell has said to him, ↓ and it's written in French, the original ↓ , that he's set to conspire and think up the said affair.

In the example 39a above the relevant information is "Anne is not Henry's front line political advisor", and is uttered in a higher pitch than the IP in the middle "as she had hoped", which is uttered in a lower pitch, shown with the arrows facing down. In example 39b the subordinated IP is "and it's written in French, the original" which is lower in pitch and faster in pace than the other relevant IPs.

Subordinated IPs can be worth considering in voice-over translation given the fact that they do not provide relevant information to the message, and could be deleted in the target text to comply with the time restrictions imposed by the original text.

Another way pitch can help structure a spoken discourse is by indicating when a new sentence starts. In English, the first element of a sentence, and paragraph, normally start on a fairly high pitch (Tench 1996), and from there, the pitch falls gradually in each sentence belonging to the same paragraph (i.e. idea). This is referred to by some as *phonological paragraphing* (Tench 1996), and *declination* (Mott 2011). The usage of pitch in this respect is quite relevant in situations in which the speaker is quoting someone else's words. Quotations normally come after reporting verbs, such as *claim, argue, say*, which are uttered in a fairly high pitch followed by a pause to mark the beginning of the quotation. The standard tone for quotations is, for Hewings (2007: 126) a level tone at the beginning and a falling tone at the end to mark the end of the quotation, as in the examples below:

Example 40

a) And he's ↑ saying, (pause) "Perhaps my marriage was always null and void for lack of proper consent."
b) He's ↑ thinking, (pause) "Can I get out of my marriage to Anne?"

However, not all quotations always start after a reporting verb. Sometimes the mere use of pitch shows it, and this can be tricky in the translation process if the translator does not have the video file to follow the text to translate, and also if he/she

is not aware of the way pitch functions to signal quotation in English. The examples below, with the quoted texts in inverted commas, illustrates this:

Example 41

He thinks, that as King, he has a direct relationship with God. "So why? Am I still not on the right side of God?"

Example 42

As relations between Anne, Henry and Cromwell all become increasingly fraught, the signal must come from Henry. "I am fed up. We want this woman out of the way".

Another case in which intonation serves a fundamental role regarding the interpretation of a message is that of tag questions. Tag questions are a very common resource in spontaneous oral speech (Leech & Svartvik 1994: 14), and can be of two types, for which there are different intonation patterns. The first group are *checking tags,* also known as *reverse polarity tags*, that is, the tag is negative if the main clause is positive, and vice versa. Checking tags have their own IP and can have different communicative meanings depending on the tone they have (rising or falling). A rising tone normally implies asking, it is a *genuine question* (Vince & Emmerson 2003: 182) so the receiver should answer the tag. If, on the other hand, the tag is not a real question, and one is sure of the answer, then the tone is a fall, and the receiver should not answer the question. Falling tags are tricky because they can have different meanings. They are not real questions and rather than asking for information, they may appeal for agreement (Wells 2006: 49) and an answer may be given. However, they can also be used to express an opinion (Leech & Svartvik 1994: 151), in which case, no answer is expected.

Example 43

a) The match was a disaster, | /wasn't it? (rising = asking - answer expected)
b) The match was a disaster, | \wasn't it? (falling = confirmation - answer not expected)

The other group are *copy tags*, which are commonly used to show surprise or disbelief, have the same polarity as the main clause (normally positive-positive, are always accompanied by a rising tone and for Tench (1996) they do not necessarily need their own separate IP.

Example 44

John: *So you're happy with the promotion, are you?* (He thought Sam would not like it because the workload is much heavier.)
Sam: *Sure, the new post is truly challenging. I love it!*

Questions tags are a very common resource in spontaneous oral speech and are distinctive markers of orality, and as such, they should be considered in audiovisual translation in general, and voice-over translation in particular, for two reasons: first, speakers do not expect an answer when the tone of the tag is falling, and therefore this type of tag is not used as "standard" questions (Wh-questions and yes-no questions) in English; "the sentence is more like a statement than a question" (Leech & Svartvik 1994: 127). And second, they are always at the end of sentences, which means that they could be heard during the sound bite between the translation and the original, which is relevant in terms of synchrony.

2.2.2.5 **Indexical function**

Intonation functions as a marker of identity for each person. One can easily identify if a speaker is young or old, a man or a woman, foreign or native, etc. by means of the features of each individual's vocal mechanism (Rose 2002: 297). In English - probably in all English-speaking countries, according to Wells (1982: 13) and Collins & Mees (2013: 02) - there is a strong connection between language and social class, so that speech stratification correlates with social stratification. Language variation, (pronunciation of a language –i.e. "accent" – together with grammar and vocabulary – i.e. "dialect" -) can be divided into two categories: regional variation, and social variation (ibid.). Regional variation is accepted by everyone without question; people from Kent are not expected to speak in the same way as the people from Hull, Glasgow,

or Cornwall. However, social variation is a more controversial issue "especially where the link with social class is concerned" (2013: 02).

The sociolinguistic pyramid shown below (Collins & Mees 2013: 03) represents the accent continuum which could be applied to The United Kingdom, and is made up of *basilects27* (traditionally associated with the working-class and the less privileged in terms of education and other social factors), *acrolects* (the most prestigious form of speech, linked to the most advantaged people in terms of wealth, education and other social factors), and *mesolects28* (which stands in between the other two).
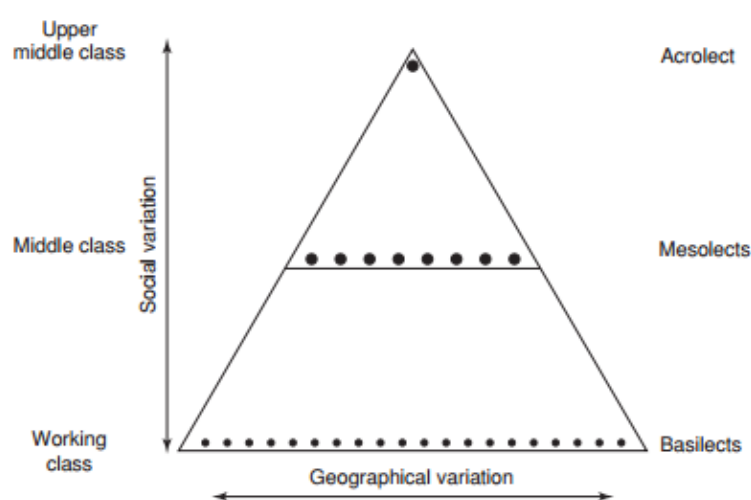
Figure 11.The sociolinguistic pyramid (Collins & Mees 2013: 03)

An example of this speech-social association is clearly shown in the TV series *Downton Abbey*, which portrays the everyday life of a family of British aristocrats in the early 20th Century. In this TV production it is noticeable how the aristocrats speak in perfect *RP English29*, an acrolect traditionally linked to the educated sector of British society, while the people working at the manor house (house maids, chauffeurs, cooks, butler, etc.) use a different type of accent, closer to the lowest part of the sociolinguistic pyramid but also showing variation depending on their position in the house: for

27 The terms "basilect" and "acrolect" were coined by William Stewart (1964).
28 The term "mesolect" was coined by Bickerton (1973) to refer to the varieties standing between acrolects and basilects.
29 RP – short for Received Pronunciation can also be referred to as "Oxford English", "BBC English", and "The Queen's English" (Collins & Mees 2013: 03)

example, the butler (Mr. Carson) uses an accent which is higher in the pyramid, and could be considered a mesolect, than the one used by Miss Daisy, one of the cooks.

The human voice is a clear marker of identity (Verhoeven 2002: 179) in every language and, as it could be seen in the previous paragraph, the case of the English language is a clear example of how one's voice can denote not only geographical origin, but also economic and educative; features that are relevant in communication and which should be explored in translation.

The indexical information provided by the human voice can be divided into two categories: *extrinsic features*, those under the speaker's control; and *intrinsic features*, those which are not (ibid). Thus, a regional or a social *accent*, in this case meaning the pronunciation pattern used by an individual, is an extrinsic indexical vocal feature. The accent forms part of one's idiolect, characterising the person as belonging to some geographical region and/or social class. An individual's pronunciation can also be, for scholars such as Wells (1982), Couper-Kuhlen (1986) typical of the speaker's sex, age group, and education level. Nonetheless, according to Rose (2002: 46), modern sociolinguistics believes that individuals do not have a "single invariant linguistic system, but use several variants in different sociologically defined circumstances". Therefore, the use of sociolinguistic variables, such as using postvocalic /r/ in British English in cases in which it would not be used, as in "I've bought a car" /aɪv ˈbɔːt ə ˈkɑːr /, a more "refined" pronunciation like RP English in a job interview, or the use of a Cockney accent when watching a football game at the pub, is a matter of speaker's choice, and as such it is extrinsic. The intrinsic features, which identify a voice, are the following: age, sex, physique, and health, including the psychological state of the speaker. According to forensic phonetics (Rose 2002: 297) the voice can carry a substantial amount of information on the speaker's organic state. Thus, one can know whether the speaker is in good health, pre-menstrual, deaf, has a cold, is drunk, or even suicidal.

Rose's analysis of voice indexical information is based on Laver and Trudgill's (1991) typology of markers of identity, which is, in turn, a revision of Abercrombie's (1967), and identifies three categories to characterise personal vocal features (1991: 03):

a) *Group markers*: those that mark social characteristics, such as regional affiliation, social status, educational status, occupation and social role.

b) *Individuating markers*: those that mark physical characteristics, such as age, sex, physique and state of health.

c) *Affective markers*: those that mark psychological characteristics of personality and affective state.

In light of this, one can notice that *group markers* fall into Rose's (2002) *extrinsic* indexical features category, while *individuating* and *affective markers* within the *intrinsic* category of indexical features.

Rose (ibid) proposes a model (see Image 15 below) to analyse the different components of a voice, which consists of four main parts: two inputs and two mechanisms. The two mechanisms are *linguistic* and *vocal,* and the two inputs are labelled *communicative intent*, which maps onto the *linguistic mechanism*, and *intrinsic indexical factors* (age, sex, physique, health), which maps onto the *vocal mechanism*. The speech wave coming from the vocal mechanism represents the final physical acoustic wave, which carries all the functions and sociolinguistic nuances distinctive in a voice.
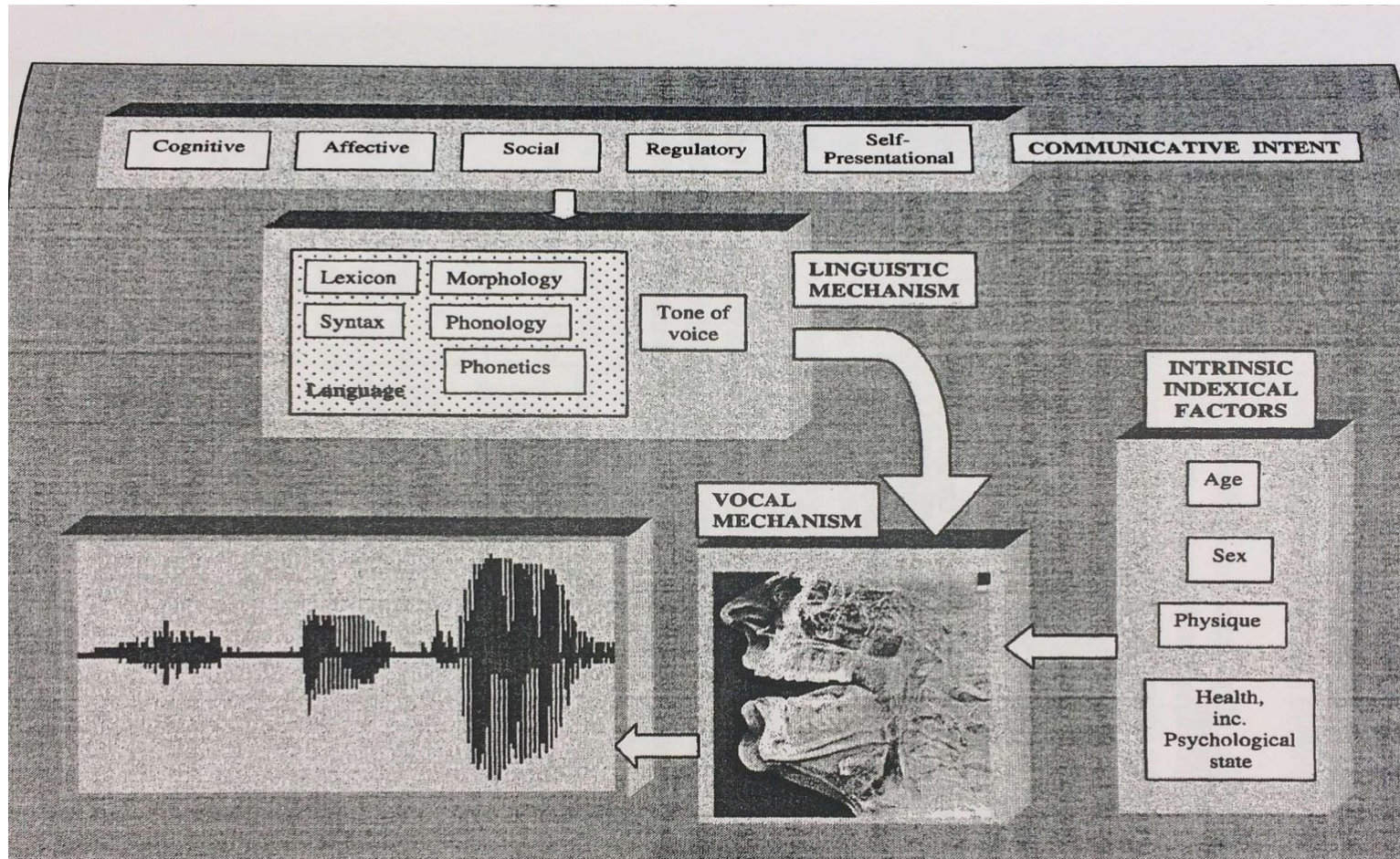
Image 15. Rose's voice model (2002: 285)

Rose's (2002) model is certainly quite interesting, relevant and useful in order to identify the communicative value of the messages in an oral source text prior to its translation.

As has been seen, intonation plays a major role in English communication not only by signalling new and old information, or structuring spoken discourse into paragraphs, or providing attitude and emotion to what we say. As happens in most languages, it also acts as an inherent marker and identifier of speakers as human beings with unique non-transferable vocal features that differentiate each individual and make them unique. In addition, given the paramount importance that intonation performs in oral communication in English, and the fact that sometimes the specific intonational features do not coincide between languages, it (intonation) should occupy a relevant place in the analysis of a source text within the translation process, regardless of whether the target text is to be written down or revoiced. This matter will be analysed in more detail in the following chapters.

2.3   Translating orality

We have seen in the previous section that intonation serves several relevant and crucial communicative purposes in oral communication, constituting the most important part of English prosody (Wells 2006: 05); and despite the fact that some intonational meaning can also be shown in writing, by means of punctuation, most is not. Thus, "spoken English, as spoken by native speakers, is richer in information content than written English" (2006: 05).

If this is so, the crucial role of English intonation should be borne in mind in audiovisual translation (somehow). Voice-over scripts should be carefully scrutinized in search for the intonational and prosodic values in them to be finally translated in the best possible way into the target language, both in the form of a script and a voice performance. Given the oral component inherent in audiovisual texts, and particularly in voice-over, these prosodic and suprasegmental features (intonation, accenting, deaccenting, etc.) are paramount to this mode of translation.

Most of the studies concerning the issue of orality in audiovisual translation have to do with *discourse markers*: how these linguistic items are translated, and also created. Discourse Markers or fillers can be categorised in three different groups, according to Leech & Svartvik (1994: 13): purely interactive, such *as, ah, aha, mhm, oh, yes, yeah, yup,* etc.; mainly interactive, like, *I see, I mean, you know, OK, that's OK, well, sure, right,* etc.; and, finally, also interactive, for example, *anyway, in fact, maybe, perhaps, probably, absolutely*, etc. The first two categories of markers (purely interactive and mainly interactive) are, above all, characteristic of conversation, while the last one (also interactive) is more grammatical and frequently used in public speaking and writing. The use of these markers in an oral conversation help signal for turn taking and also reinforce the cooperation principle, since although they may add little information, "they tell us something of the speaker's attitude to their audience and to what they are saying" (Leech & Svartvik 1994: 14).

The idea of a prefabricated orality in audiovisual translation (Baños & Chaume 2009) implies that oral features characteristic of human communication, such as discourse markers, hesitations, etc. which will be translated into the target language, are created (fabricated) ex profeso in the original script, as a result of the scriptwriter's imagination and creativity. In line with this, Chaume (2001: 79) states that "el discurso oral de los personajes de pantalla no es más que el recitado de un discurso escrito anterior. Sin embargo, tal recitado ha de parecer oral." For Braga & Rica Peromingo (2015: 104), this notion of prefabricated orality, which they study for the translation of orality in literary texts, is also referred to as "oralidad fingida", "oralidad construida", or "oralidad ficticia".

Matamala (2008) considers that translators for voice-over should be trained to omit certain oral features without referential content, as well as hesitations and fluffs. This is in line with Barzdevics' (2012: 67) recommendations, referring to discourse markers, hesitations etc.:

> En estos casos, el traductor ha de recordar que al espectador no le interesa la destreza lingüística del personaje. Le interesa lo que dice. Por ello eliminaremos todos estos elementos que sólo sirven para molestar y le daremos más fluidez. (2012: 67)

Another scholar who shares this opinion concerning discourse markers and spontaneous oral features is Chaume (2001: 81), for whom the most noticeable aspects of spontaneous oral language should be "obviated", i.e. deleted, in order to achieve the perfect balance between credibility on one side, and the compliance of linguistic and stylistic norms, on the other. However, this scholar assumes that some discourse markers are "indispensable to the logical composition of ordinary conversation or written discourse, and without which conversation or written discourse would fall apart." (2004: 844). Chaume alludes to the redundancy of audiovisual texts (the same information coming from two channels) as an alibi to dispose of discourse markers in the process of translating, since this does not produce any semantic alteration in the target text, although "it does in terms of interpersonal meaning" (2004: 854). Nonetheless, some scholars like Crystal (1988), Aijmer (2002) and Muller (2005), consider that discourse markers provide coherence to the text and contribute to the pragmatic meaning of sentences, performing a fundamental role in the pragmatic competence of speakers. This is why for Rica Peromingo (2014: 180) discourse markers have a paramount function is spoken language.

The translation of orality and the oral traits of human communication, leaving discourse markers aside, has been the topic of research, and debate, for several scholars within the discipline of audiovisual translation in recent years. In this respect, Franco (2001) considers orality as a central issue of voice-over translation in documentaries, not merely because the discourse is purely oral, but also because orality is "directly linked with the idea that there are individual speakers in the documentary" (2001: 80). This is fact brings back the indexical function of intonation, which characterizes every speaker, as discussed in the previous section. In her analysis of 22 documentaries about Brazil produced in western Europe and voiced-over from Brazilian Portuguese into French and German, she found that the transcriptions left many parts of the original out by means of omissions, and other parts were reproduced incorrectly due to inaccuracies. This led her to conclude that "'facts' are already filtered or selected before they are translated" (2000: 81).

This filtering and selection in the transcription process compromises the notion of *faithfulness* in voice-over translation (*"the notion of faithful reproduction is at*

stake") (2001: 77). In her analysis, she found there was a deletion of those utterances that showed the interviewees were aware of the camera, which in turn would, for this scholar, uncover the fact that spontaneous testimonies were actually directed.

In addition to this, Franco (2001) understands that the written reproductions of sociolects (Brazilian sociolects in her study) can hinder the translation process, since they may misrepresent the original spoken speech.

> If similarity in pronunciation represents the key to the transcriber's misunderstandings, then it also turns into the translator's trap from which other misunderstandings may be generated. This means that the notion of faithfulness is at stake also due to the transcriber's background and his/her ability to recognize and convey non-standard language. (2001: 77)

Franco (2000) is directly pointing here at the role of the transcriber as a filtering device, which can have a direct and relevant impact on the translation process, affecting the desired faithful representation of the original. The role of the transcriber as an active participant in the translation process of voiced-over products will be analysed in detail in chapter 3.

Another view regarding the relevance that oral features have in voice-over translation is Gorska (2015). This scholar emphasises the importance of oral traits in voice-over given the fact that the transfer channel (the original voice) is not deleted and reaches the audience. She acknowledges that some elements of the original script (oral script) are deleted in voice-over, for the purpose of synchrony, as is the case with names, foreign words with the same pronunciation in both languages, and repetitions of expressive nature. Nonetheless, for her, the audiovisual translator has in voice-over the chance to show the oral imprints present in the original text, which is not possible in disciplines like subtitling due to the rigid nature of written language. This is one of the challenges that the voice-over translator has to face, as stated by Matamala (2008), in her article on the teaching of voice-over translation:

> Finally, translators must be able to render the translation of an oral product in an adequately reformulated written form, bearing in mind that viewers will receive it as an audiovisual product […] this is one of the main difficulties students encounter when

being trained in AVT, because they are generally used to written translations. (2008: 118)

It could be that the habit and tradition of translating written formats makes no room for spontaneous orality in audiovisual translation. However, it all depends on the product and who the final audience will be. In some cases, depending on the genre that is being translated, oral markers will be more or less condensed. Translating voice-over for unscripted programmes such as a documentary, for which a sort of sobriety is expected, is rather different from rendering an episode of a reality show, such as "*Love it or list it*30", or "*Teen Mom*31" which have a friendlier and more casual approach.

As can be seen in the previous paragraphs, orality has been predominantly analysed by researchers within the field of AVT in terms of whether discourse markers should be omitted or not, and the fake nature of dialogues which pretend to sound natural and real in films. However, there are other oral features inherent in speech and intonation such as accenting and deaccenting, rhythm and stress, and voice quality that are utterly pertinent in the original oral text and which need to be rendered accordingly in the target oral text to comply with the assumed generalisation that voice-over is the most faithful translation mode. In fact, "translation does not end with `text´ but with delivery" (Gambier & Gotlieb 2001: xix), which brings the idea that in audiovisual translation, the translation process is not just a matter of the translator alone (Mayoral 2001: 05) but also of the director, adapter, and finally the actors whose voices will reach the audience, and which are one of the main tools of expressiveness (Gorska 2015: 67).

The oral features inherent to an individual's voice and speech - be those cultural, social, economic, etc. - are therefore relevant too. For Wells (1982: 08) "one of the most obvious things we notice about a person's speech is that it tells us about where he comes from. Accents are powerful indicators of geographical identity", as we saw in the section devoted to the indexical function of English intonation. Hatim & Mason (1990) consider that translators have to be aware of the geographical variations of speech and

---

30 Information on the TV programme "Love it or list it" can be found here:
https://www.imdb.com/title/tt1773182/
31 Information on the TV programme "Teen Mom" can be found here:
https://www.imdb.com/title/tt1566154/?ref_=fn_al_tt_1

of the ideological and political implications that may accompany each particular variation. For them, "accent is one of the most recognizable features of geographical variation and is often a source of problems" (1990: 40).

Translating accents, and the oral features proper of certain characters, is not an easy task and can sometimes lead to "awkward" translations, mostly from the audience perspective. There are some techniques that might be useful when it comes to translating dialects, which have been gathered by Braga & Rica Peromingo (2015: 134-147) in an attempt to show all the available options that translators have in order to deal with these oral features in literary translation. They highlight the techniques suggested by Perteghella (2002) and Tello (2011), which are described in the following below.

PERTEGHELLA (2002)	TELLO (2011)
Dialect Compilation	**Compensation**
This strategy consists of converting the original dialect or slang into a mixture of dialects or slangs characteristic of the target language.	This strategy allows some marked structures of the original dialect to be translated as the standard of language in the target text, while others which are not marked are marked in the target text.
Pseudo-dialectal Translation	**The Creation of a Dialect (norm infringement)**
It means that the translator can create a non-existent dialect using a mixture of non-standard language and some idiomatic traits characteristic of some of the dialects available in the target language.	This implies making up a dialect for the target text, including wrong grammatical constructions, for example.
Parallel dialect Translation	**Dialectal Translation**
This technique allows translating a dialect as an existent/real dialect in the target language; this target dialect normally resembles the characteristics of the original one in the source language.	It allows choosing a dialect in the target language as an equivalent of the one from the source text. Normally these two dialects share the same characteristics in both languages (social, political, etc.)
Dialectal Localization	**Colloquial Translation**
This technique is very similar to domestication and acculturation since the dialect is translated an existent dialect of the target language and all character names and cultural references are also translates into the target language.	It means replacing the original dialect by a colloquial and informal language in the target text. This implies using phonological, grammatical, lexical and syntactic elements characteristic of the target language.
Standardization	**Neutralization**
The dialect or slang is translated as the standard of the language in the target text. This technique is also known as "levelling".	The dialect or slang is translated as the standard of the language in the target text. This technique is also known as "levelling".

Table 14. Perteguella's (2002) and Tello's (2011) techniques to render foreign accents

Although these techniques described by Braga & Rica Peromingo (2015) are intended to literary translation, they could be applied to audiovisual translation too. Some examples can be found in Renée Zellweger's dubbed voice in Spanish in the film *Cold Mountain*32 (2003), or Brad Pitt's in *Snatch*33 (2000). In these films the actress and actor characterize an American southerner (Zellweger) and an Irish boxer (Pitt), but their accents and voice properties in the translated versions are invented for the occasion, using colloquialisms, mispronunciations and a poor vocalization which have nothing to do with the specific and characteristic nuances of their original accents. We can notice that, for these two movies, the translators have opted to translate the original dialects using Perteghella's (2002) pseudo-dialectal technique, or Tello's (2011) colloquial technique.

These oral features - which were discussed in the Indexical Function section, and illustrated with the series *Downton Abbey*, are normally neutralized in voice-over translation since it is almost impossible to convey them in translation, as in the case of regional accents, since these are characteristic and unique to each language. Other oral features, such as voice quality, age, etc. can be adequately rendered in the target version, as happens in dubbing (Sánchez Mompeán 2012: 285), when the same actors do the original and translated versions of certain movies. Examples of this can be found in, Antonio Banderas and Salma Hayek in *Puss in Boots* (2011), and Paz Vega in *Madagascar* (2012). Antonio Banderas is an actor who has dubbed himself into Spanish in most of his English-speaking films, unlike most of the Spanish actors who worked in foreign productions.

In Landers' words: "no dialect travels well in translation" (2001: 117), and the same applies to accents, as one of the oldest and most recurrent problems not only for dubbing, but also for dialogues in novels and plays (Zabalbeascoa 2001: 52). In line with this, Romero Ramos (2005: 246), in her study of the translation of social and geographical dialects in subtitling, distinguishes between three branches: linguistic, extralinguistic and pragmatic. For this scholar, the omission of equivalent geographical

32 Cold Mountain film information available on :http://www.imdb.com/title/tt0159365/
33 Snatch film information available on: http://www.imdb.com/title/tt0208092/?ref_=nv_sr_1

nuances in the translation implies the deletion of the social nuances inherent in the use of the original dialect (2005: 251), which she sees as inevitable:

> En cuanto a la ausencia de marcas geográficas equivalentes, consideramos que, probablemente, ésta sea la única solución viable, ya que el hecho de hacer hablar a un personaje en una variedad geográfica del español cuando la acción narrativa se desarrolla en Italia, rompería con el principio de verosimilitud y transgrediría el umbral de permisividad del espectador. No obstante, la ausencia de rasgos geográficos equivalentes implica también la desaparición de las connotaciones sociales implícitas en el uso del dialecto geográfico en el original.

Petite (2005: 62) is aware of this loss of meaning concerning style and register in the source text (and performance) and declares that, unlike subtitling, dubbing (and in our case voice-over too) allows for compensation by means of intonation, stress and volume with the performance of the dubbing actor. Nonetheless:

> It is not possible to convey the multiple meanings of each and every utterance but the spirit of the entire audiovisual text, composed not only of language but also by other auditory effects (…) (2005: 62)

In the case of voice-over, the situation described in the dubbed films *Cold Mountain* and *Snatch* is not common, although in the UK it is quite common to use voices with the accent stereotypically associated with the nationality of the person speaking in the original version; a controversial procedure already stated by scholars such as Fawcet (1983), Franco, Matamala and Orero (2010: 79) and Barzdevics (2012: 68). This means that if in the original version there is a Spanish person speaking Spanish, the translated voice-over version into English will always have a Spanish voice talent speaking English with a Spanish accent. Although the purpose of this strategy is probably to provide authenticity to the translation, in fact, it can also be interpreted as reinforcing stereotypes, patronising, even racist. In line with this, Díaz Cintas and Orero (2010: 442) consider that this approach could be "far from deferential to the Other as it seems to perpetuate negative stereotypes by highlighting the inability of foreign people to speak correct English." Franco, Orero and Matamala (2010) agree with this view concerning the "reproduction" of original accents in voice-over, which

they regard as tricky since, on the one hand, it could be positively assessed as more authentic but, on the other, it can be condemned as a form of "caricature or fake" (2010:30).

Hatim & Mason (1990: 40) exemplify the controversy of endowing characters with accents (foreign or national) with the case of a TV dramatization of a foreign play in which Scottish accents were used to represent the speech of Russian peasants. This strategy - which implies replacing an original accent with another accent belonging to the target language - caused controversy in Scotland since the Scottish accent was used to represent low status. For Hatim & Mason (1990: 40) "like producers and directors, translators have to be constantly alert to the social and political implications of their decisions".

As a consumer of voiced-over products and as a professional voice-over talent and translator, I think that regional accents, the markers of social class status, and the cultural baggage implied in speech are extremely difficult to translate successfully into the target language in AVT. Nobody would expect that in a documentary about the steel and coal industry of Birmingham and Durham, places with very recognizable accents (Brummie and Geordie), the characters' speech was rendered in Spanish with the accents of Mieres or Avilés, in Asturias, simply because these are places with metallurgic industry and coal. The extralinguistic nuances of the accents of different languages (e.g. Geordie and Mieres) have cultural connotations that are not equivalent, and the same applies to the pragmatic value intended in their use.

The process of "reducing" the oral features of non-standard or regional accents (in the sense of neutralizing them by assimilating them to the standard in the target language) in audiovisual translation and voice-over, and is known as "leveling", which inevitably implies that the socio-economic and cultural richness implicit in every English, or other language, accent will be lost in translation (Barzdevics 2012: 67).

However, as we have just seen, there is no easy solution to this loss, since substituting accents from the target language is probably worse than the loss produced by the levelling of the source text accent. Chaume (2001) believes that the translation of certain phonetic features should be manipulated, or modified, so as to create a

"fabricated, or new, orality" based on the original spontaneous discourse, which satisfies the general public. These modifications are mostly segmental but there are also prosodic features to be considered. He recommends a tense and clear articulation, which contrasts with the relaxed articulation of colloquial speech, and avoid the reduction or suppression of consonants (proper of colloquial oral discourse), the loss of intervocalic *d* in Spanish, prosodic ambiguities and cacophonies, which could appear spontaneously in colloquial oral speech, and emphasise intonation 34. (2001: 82) Moreover, the translator must make an effort so that phonetic phenomena such as these do not happen in the translation. Applying Chaume's recommendations, the resulting translated discourse would not be as oral, nor as colloquial; and this applies to any audiovisual genre; even for those which try to achieve a higher status in terms of faithfulness through the usage of oral and colloquial registers (2001: 82), as is the case of voice-over.

2.4 Translating the illocutionary force

The pragmatic dimension conveyed in every *speech act*35 consists, for J. L. Austin (1962) (Hatim & Mason 1990: 59), of three different actions: locutionary act, illocutionary act, and perlocutionary act.

1. Locutionary act: the action performed by uttering a well-formed, meaningful sentence.
2. Illocutionary act: the communicative force which accompanies the utterance, e.g. promising, warning, conceding, denying, etc.
3. Perlocutionary act: the effect of the utterance on the hearer/reader; i.e. the extent to which the receiver's state of mind/knowledge/attitude is altered by the utterance in question.

For Mateo (2014) the illocutionary force implied in every speech act is the "real driving force of communication", and consequently, is "vital for translation" (2014:

---

34 The characteristic intonation of Spanish modern dubbed films is so recognisable due to its emphasis that some professionals call this way of speaking as "doblajitis" ("dubbitis" in English) as if it was a sickness. This can be seen in the video link: https://www.youtube.com/watch?v=9_26Qcxk62k

35 Actions performed via utterances are generally called *speech acts* according to Yule (1996: 47)

125). In line with this, Bell (1991), believes the translator may ask himself/herself certain questions regarding the rendition of the illocutionary force:

How do we make utterances count as particular speech acts?; how do we recognize what kind of a speech act a particular utterance is? And how are we to cope with the fact that differences in realization of the same speech act from language to language? (1991: 174)

All these questions relate to intonation, since this component of language "guides the receiver to the interpretation intended by the speaker of an oral text" (Mateo 2014: 125). The illocutionary force which is present in every utterance is already considered as a function of English intonation by Couper-Kuhlen (1986) (see Table 13 in chapter 2.2.2.3), for whom "intonation signals the intentional force of an utterance in a given context" (1986: 112). Gómez González & Sánchez Roura (2016) agree that intonation (by means of tone) is vital to encode the illocutionary force of an utterance, be that a command, a question, or a statement; and this is so because of the inexistence of a one-to-one relationship between grammatical categories (declaratives, interrogatives, imperatives and verbless constructions) and illocutionary categories (statements, questions, commands and exclamations (2016: 301).

Also referring to Searle's (1969) notion of illocutionary force and its pragmatic purpose, Hervey (1998: 18) highlights the importance of the illocutionary functions of the source language in a process of translation. In every language, he explains, there is a set of sentential units whose value is illocutionary, which are: illocutionary particles, intonation and sequential focus. These units do not have the same relevance or presence in every language, as "one or other of these categories may blatantly predominate in terms of frequency and expressive potential" in each different tongue (1998: 16). In addition to this, the illocutionary meanings carried in one source language might not find their equivalent through the use of the same units or elements in a target language: "the illocutionary meanings conveyed by intonation in a ST are not necessarily expressible through intonation in a TT" (1998: 18). Accordingly, Solé (1989) believes that when translating from e.g. English into Spanish, the differences in the intonation structure of these languages have to be born in mind, since "it cannot be taken for granted that there is a one-to-one relationship between the two languages" (1989: 181).

Hervey (1998: 17) categorises languages into three types: particle-oriented (German), in which a word or particle can have two distinct grammatical invariants under the same formal guise, as in "Er ist schon gekommen" (He has already come) and "Er wird schon kommen" (He will indeed come)[36]; intonation-oriented (English); and sequential-focus-oriented (Hungarian), in which syntactic functions are signalled by its morphology rather than word order, thus leaving word order free to convey illocutionary nuances.

Morphology		Word Order		Intonation
Rich	$\longrightarrow$	Free	$\longrightarrow$	Rich
Poor	$\longrightarrow$	Rigid	$\longrightarrow$	Very Rich

Table 15. Language categorisation in terms of morphology syntax and intonation

Accordingly, translating from and into these kinds of languages implies certain strategies that have to be considered, and which give a certain amount of "freedom" or flexibility in the translation of illocutionary functions. As Hervey suggests (1998: 18),

1- When translating from German (particle-oriented) be sure to pay special attention to illocutionary particles in the ST.
2- When translating from English (intonation-oriented), be sure to pay special attention to the illocutionary function of intonation in the ST.
3- When translating from Hungarian (sequential-focus-oriented) be sure to pay special attention to the illocutionary function of sequential focus in the ST (Hervey 1998: 18).

Therefore, if intonation in English is not equalled in Spanish in the conveyance of illocutionary force, the latter must be rendered by means of those other strategies or resources. Strategy number 2 above is the one studied in this work, and which brings up the question again of how the illocutionary function of intonation in the ST (English) can be rendered through the usual resources in Spanish. According to Hervey, "the

---

[36] This is a literal word by word translation to show that the emphatic particle goes before the verb in the German example (schon – indeed).

sensitive handling of the illocutionary function of sentences (as speech acts) is an essential aspect of skill in translating and in interpreting." It is both sensitive and essential, and in addition, it is to some extent, also unknown in voice-over translation.

Considering Hervey's classification of languages, we can identify the behaviour of languages regarding the conveyance of pragmatic meaning. In this sense, languages which are morphologically rich like Hungarian or Spanish have a free word order to convey pragmatic and illocutionary meaning, unlike languages with a poor morphology, such as English, which, as a consequence, have a rigid word order (Hickey 1993: 580) and need to resort to intonation to convey such illocutionary nuances. The figure below illustrates the pragmatic resources available in English and Spanish, which as can be seen, are almost the opposite.

English	Spanish
Rich Intonation	Limited Intonation
Rigid Word Order	Free Word Order
Poor Morphology	Rich Morphology
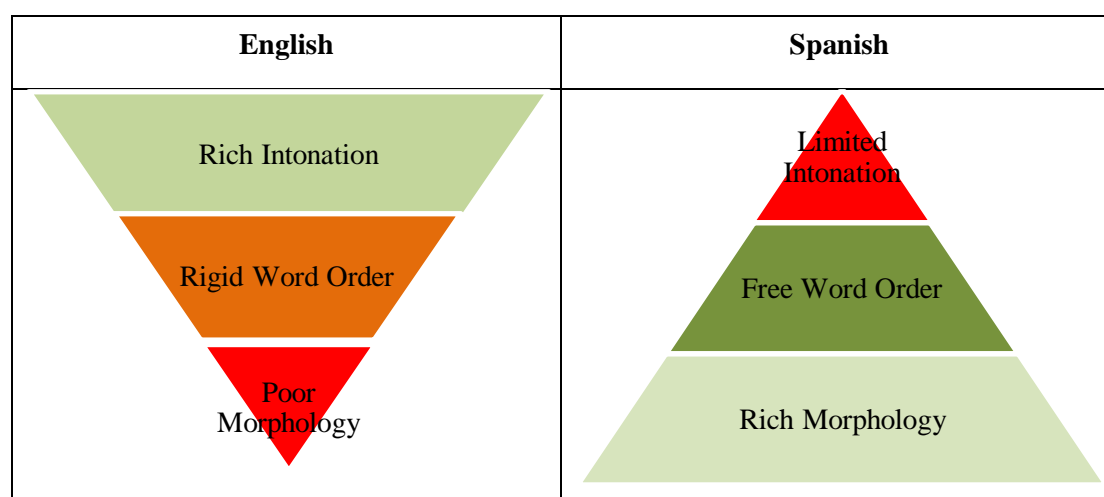
Figure 12. Pragmatic transfer of languages according to morphology, syntax and intonation

The transfer of the illocutionary force between English and Spanish has already been studied by scholars such as Mateo Martínez (1996: 78), who promotes the debate of how relevant the notion of illocution, and therefore intonation, is for successful translation. Understanding that there is an unavoidable loss of meaning in any translation, which can be more or less evident, altering the final result, he suggests that by comparing both the source and target texts that the translator will realize that something does not make sense, being left an unsatisfactory feel.

For instance, in the case of voice-over translation - in which target texts should be shorter in word count and time length than source texts due to the time restrictions already mentioned for synchronization matters - the translator has to consider how to translate the illocutionary function of the source text, bearing in mind that reduction, in order to convey the communicative purpose of the original. Translators ought to remember the pragmatic element of the script, and together with the voice-over talent and the director (if it is the case), reflect the illocutionary elements of the original script, or discourse, creating a faithful translation.

This author considers that there are also some cultural and social elements that have to be adapted to the target text in the transfer of illocutionary force between English and Spanish:

> La fuerza ilocucionaria no solamente refleja, como hemos dicho, aspectos de la intención real del hablante al emitir sus enunciados sino también cuestiones relacionadas con la configuración social y cultural de la lengua origen que deben adaptarse a los modelos análogos de la lengua meta (1996: 83).

These cultural and social elements, characteristic of every speaker, are somewhat hard to convey in voice-over translation since most of the translated products are translated and voiced into what is considered the neutral and standard version of the language, in our case, standard Spanish, or Castilian Spanish, as we saw in section 2.3. As a result, all the cultural and social elements inherent in the English spoken by a native English-speaking person in the original will be lost and standardized as a perfect Spanish discourse.

There is a special form of language use that is significant in terms of illocutionary force and prosodic value, which is not very much regarded in translation: quotations (Birbili 2000: 02). As we have seen, quotes are usually signalled by means of intonation (pitch variation) and the use of reporting verbs. Usually, when translators come across a quotation in a text, they decide to render the quote by means of reported speech or indirect speech, as for example:

Example 45 (from *The Last days of Anne Boleyn*)

Source text	Target text
He says, "I had sex with the Queen on three occasions."	Asegura haber mantenido relaciones sexuales con la Reina en tres ocasiones.

For Clark & Gerrig (1990) and Depraetere & Salike (2017) quotations are demonstrations; while on the other hand, reporting are descriptions, and both are "fundamentally different methods of communication" (1990: 764). In addition, these scholars consider that most quotations depict illocutionary acts, including the proposition expressed and, as we have seen, illocutionary acts consist of expressing a proposition plus specifying and illocutionary force (1990: 778). In the light of this information, we will consider quotes as demonstrations carrying relevant illocutionary weight, and as such they will have their own section in our analysis section.

The importance of knowing the function of the illocutionary force in translation is paramount, since it constitutes, as Mateo Martínez (1996: 82) states, the real communicative mechanism of speech acts, which includes the true intentions of the speaker. These are expressed by means phonological, suprasegmental, or stylistic (intonation, emphasis) mechanisms, word order, diacritics, etc. In fact, the illocutionary force contributes and determines the perlocutionary effects on the receiver.

Perlocutionary equivalence (Hickey 1998), which is in fact the key to successful translation, is achieved by using an equivalent illocutionary force in the target text, which will allow the translation to produce an equivalent effect on the target receiver, to that produced originally on the source receiver. For Mateo Martínez, phonology, intonation and suprasegmental features of languages are utterly important here. In line with this, Mateo (2014: 125) understands that we, translators, "must count this suprasegmental feature of the language", referring to intonation, as it "guides the receiver to the interpretation intended by the speaker of an oral text". In addition to this, Garcia-Lecumberri (1995: 91), considers that in English, intonation is one of the most important focusing devices, even more frequent than clefting or fronting. On the other hand, Spanish behaves quite differently in this regard so "intonational focus is a more common mechanism in English than in Spanish" (1995: 334). In fact, as Mateo points

out, to express marked focus, Spanish usually resorts to adverbial particles, syntactic movement, emphatic pronouns, idiomatic phrasing, etc., while the nuclear accent normally falling in final position in all cases (2014: 126). Therefore, Spanish would be, using Hervey's classification, both particle-oriented and sequential focus oriented (see above).

Braga & Rica Peromingo (2015: 114) focus their study of the illocutionary force in the translation of emphasis, one type of illocutionary force, to achieve the desired authenticity in the translating process. According to them, there are two types of emphasis which are relevant in translation: structural emphasis (*énfasis estructural*), which implies a word order shift to give prominence to a given element in the sentence, and emotional emphasis (*énfasis emotivo*), which is achieved by a certain lexis and the semantic burden of a word, phrase or expression.

As regards structural emphasis, Braga & Rica Peromingo state that in English it is very common to use cleft sentences and the passive voice, since placing certain information at the beginning of a sentence gives it psychological prominence. However, in Spanish this sort of emphasis is not as marked given the flexibility of the Spanish language to juggle with sentence constituents. Despite this, they consider cleft sentences really appealing for the translator given the different possibilities for their translation into Spanish. These scholars distinguish three types of cleft-sentences in Spanish based on Gutiérrez Ordóñez's (2000: 37-39) taxonomy, as exemplified in the following paragraphs.

The first type are *equational structures* and are formed by three elements: what is intended to highlight; the rest of the structure, which is introduced by a relative pronoun; and the verb "to be" in the required tense. Gutiérrez Ordóñez (2000: 37) proposes a syntagmatic scheme to exemplify this type of structure with the sentence: "el decano es quien ha convocado junta para el lunes". Thus:

Example 46

A is B: El decano es quien ha convocado junta para el lunes.
B is A: Quien ha convocado junta para el lunes es el decano.
It is A B: Es el decano quien ha convocado una junta para el lunes.

From these alternatives, although the three are valid from a grammatical and syntactical point of view, it is the first one (A is B) the one which provides the most expressive load (Gutiérrez Ordóñez's (2000:37).

Braga & Rica Peromingo (2015: 115), in turn, exemplify the rendition of emphasis in Spanish with the following examples, offered as possible translations of an English cleft-sentence:

Example 47

It is horror movies that I love
Las películas de miedo son las que me encantan / Son las películas de miedo las que me encantan.

Example 48

It was yesterday that we went to the movies
Ayer fue cuando fuimos al cine / Fue ayer cuando fuimos al cine.

The best translation for each of the examples above, according to Braga and Rica Peromingo (2015), and following Gutiérrez Ordóñez (2000) is the first option, and it would not be unusual that they carried some "insistence stress". Nonetheless, the second option would not be incorrect grammatically speaking, since flexibility that Spanish offers allows to juggle with the different constituents without changing the semantic meaning.

A second type of cleft sentences in Spanish, would be a variant of equational structures37, which for Braga & Rica Peromingo (2015) implies placing forward the relative pronoun or adverb at the beginning, and placing the emphatic object at the end of the sentence so there is an end-focus, as in the examples shown below from Braga & Rica Peromingo (2015: 115-116):

37 This category does not have a name, it is called "otras estructuras" (other structures) in Braga & Rica Peromingo 2015 since it is a variant of equational structures.

Example 49

It is horror movies that I love: Lo que me encantan son las películas de miedo.

Example 50

It was yesterday that we went to the movies: Cuando fuimos al cine fue ayer.

The third type are *conditional periphrasis structures*[38], which differ from equational structures in the fact that instead of a relative phrase, there is always aconditional in the first clause (Braga & Rica Peromingo 2015: 116).

Example 51

It is horror movies that I love: Si algo me encanta son las películas de miedo.
It was yesterday that we went to the movies: Si en algún momento fuimos al cine, fue ayer.

Moving now to emotional emphasis (*énfasis emotivo*), Braga & Rica Peromingo (2015) highlight that it is normally expressed by means of certain elements instead of playing with the distribution of the syntactic elements[39]. Those elements include: In this case, there should not be any problem to render the emotive emphasis from English into Spanish, considering the parallelisms shown by both languages (2015: 116-117). Furthermore, Spanish has its own lexical resources, such as terms like "menudo", "lindo", "bonito", or the exclamatory particle "qué", that contribute to marking focus; and it can sometimes use the pronoun "lo" as a fronting focusing strategy, as in the examples below:

Example 52

What a girl!: "Qué tía / ¡Menuda chica!"
Wasn't it a marvellous concert!: "¡Menudo concierto! / ¡Qué concierto!"
What a beautiful day we've had: ¡"Lo bien que lo hemos pasado!"
Hasn't she grown!: "¡Lo que ha crecido!"

[Braga & Rica Peromingo 2015: 117]

---

[38] This category is labelled as "Estructuras ecuandicionales" in Braga & Rica Peromingo 2015
[39] What Martínez Caro (1995) calls "orden de constituyentes"

Braga & Rica apply their study of emphasis to literary translation, in which there are no time restrictions and word count is not too relevant. However, in the realm of audiovisual translation, be that voice-over, dubbing or subtitling, some of those translations in which the target text is longer than the source text, or which include the repetition of words (and syllables) to make some meaning clear (*es un fallo, fallo*) can hinder the synchronisation of the target audio and the source video.

Nevertheless, Braga & Rica (2015) do acknowledge that English is special when it comes to expressing emotional emphasis, since it has a resource which is very unusual in Spanish: phonic emphasis, a resource which, they agree, is sometimes obviated in translation. "Este tipo de énfasis sónico pasa en ocasiones desapercibido para los traductores" (2015: 118).

Besides the conveyance of emphasis, another relevant difference between English and Spanish worth mentioning at this point, is the fact that English shows a preference for juxtaposition and coordination, whereas Spanish tends to subordinate, as is pointed out by López Guix & Wilkinson (2001: 89):

Example 53

Eventually came the day we had been longing for.
- Al final llegó el día que habíamos estado esperando.

[López Guix and Wilkinson 2001: 90]

Nonetheless, Chaume (2001) seems to encourage the opposite of this tendency observed by López Guix & Wilkinson concerning the preference of subordination in Spanish, when he suggests "procurar usar de más a menos enunciados yuxtapuestos, coordinados y subordinados (en este orden)" (2001: 84).

The reason why Chaume is suggesting here something which apparently goes against the normal tendency of Spanish in terms of emphasis, is probably that, by using the same structures as in English (juxtaposition and coordination) in the Spanish translation, the synchronisation between the Spanish audio and the original English video will be more natural and harmonious. There will probably be fewer words in the

target text with this strategy than in the case of subordination, something which is fundamental for suitable synchrony.

The present study will examine how voice-over translation is traditionally done in Spanish, and analyse how these common translation strategies affect the illocutionary purpose of the original text, whether they seem to have taken into account the role of English intonation in this regard, and if there is a loss of meaning in the final version. The rendition of intonation and its illocutionary force should also be considered in the delivery (performance) of the translated audiovisual script by the voice-over talent, since it will contribute to the desired perlocutionary effect in the audience/viewers of the audiovisual product. As has already been stated, the translation in AVT does not end with the writing of the translated script; in fact, the most relevant part comes just after.

2.5   Translating English intonation

Intonation is not random. There is a structure, and there are rules (Tench 2013: 06). Intonation in English operates at three different levels (tonality, tonicity, and tone), known as *the three Ts*, as was discussed in section 2.2.2. The structures and rules conveying these different components of English intonation will often work differently in Spanish particularly since, as we have seen, Spanish seems to prefer linguistic resources other than intonation for the conveyance of some illocutionary functions. Since intonation plays such an important role in English oral communication, and the source texts in English-Spanish voice-over (our object of study) will be oral, let us examine the translation of English intonation by considering the 3 Ts.

2.5.1   Translating the 3 Ts

2.5.1.1   Translating tonality

According to Solé (1989), whose paper on the translation of intonation was pioneering, and an inspiration, when it comes to translation purposes from English into Spanish, the grammatical structure conveyed by intonation - mostly what the English school has termed "tonality" - has to be translated correctly, and "it cannot be taken for granted that there is a one-to-one correspondence between the two languages" (1989: 181). Moreover, the IPs (or tone groups) as she calls them, will frequently be translated by a different grammatical structure or vocabulary item in another language, in this case, Spanish. Solé provides some examples to show the grammatical and lexical changes normally involved in the translation of tonality from English into Spanish (1989: 182). Let us just see one illustrative example which could have two different interpretations depending on the tonality used in oral discourse:

Example 54

She dressed and fed the baby.

Option 1: She dressed | and fed the baby. (she dressed herself and then fed the baby)
Option 2: She dressed and fed the baby. (The baby was both dressed and fed).

Each of these options would correspond to a different translation in Spanish both in written and oral format.

Option 1: Se vistió | y dio de comer al niño.
Option 2: Vistió al niño y le dio de comer.
Alternative option (not included in Solé's paper): Vistió y dio de comer al niño.

[Solé 1989: 182]

It is evident that there are changes to be made when translating the grammatical structure of IPs from English into Spanish, but the way in which these changes are relevant in ATV, particularly in voice-over, needs further research. These changes may sometimes imply the re-writing of the whole sentence, which may, in turn, affect the length of the script, and therefore, also the timing or pace for the voice talents to voice their lines. All this, combined with the time restrictions (around 2 to 4 seconds before and after the translated voice starts), might end up blurring a translation, by rushing the oral discourse or deleting some relevant parts of the message.

For example, in option 2 *she dressed and fed the baby* the nucleus would normally fall on *baby*, whereas in the Spanish translation provided, the nucleus would fall on *comer*. Both of Solé's translations of the examples are perfectly valid. However, the alternative option suggested (*Vistió y dio de comer al niño.*) could probably be more suitable, in this case, since it has one less syllable than option 2, and the last words coincide in both texts (*baby/niño*).

The example shown below shows how the rewriting of the translated text can affect timing restrictions and hinder synchronic and tonality features.

	Original English Text	**Original Spanish Translation**	**Alternative Spanish Translation**
5	Sitting in London Brexit stares us in the face	Desde nuestra sede en Londres, el brexit nos mira directamente a los ojos.	Desde Londres el Brexit nos mira a los ojos.
6	Will it be hard? Soft?	¿Se optará por una versión dura o blanda?	¿Será este duro? ¿O blando?
7	Orderly? Disorderly?	¿Se llevará a cabo de una forma ordenada o desordenada?	¿Ordenado? o ¿Desordenado?

Table 16. Example of long translation taken from TN17-005[40]

The excess of text in the original translation is evident and exemplifies how an unreasonable and excessive word count can frustrate a translation for voice-over. The question, again, is how aware are translators of this grammatical and lexical juggling when translating voice-over? Do they use any special techniques to avoid such inconveniences? Or do they play by ear in this case?

---

[40] Full text and video available in the annexe section

2.5.1.2 **Translating tonicity**

Another strategic function of intonation in English is the one related to focusing. Native speakers do this mostly by tonicity, or the placing of the nuclear accent as was discussed in section 2.2.2.2 above, and the placement of other accents.

Martínez Caro (1995: 367) studied the way tonicity worked in English and Spanish from a pragmatic, accentual and word order perspective and came to a conclusion that coincides with what was observed above in relation to Hervey's (1998: 15) classification of languages according to their preferences to convey the illocutionary force:

> La dimensión de la focalidad es expresada mayoritariamente en inglés por medio de mecanismos de tipo prosódico, que se refieren a la movilidad del AN (acento nuclear) en el grupo fónico (IP), y en español por medio de construcciones que conllevan a una variación del OC (Orden de Constituyentes). Se compensa así una estructura sintáctica relativamente rígida en inglés con la flexibilidad de la posición de su AN y la mayor rigidez de los patrones prosódicos en español con su flexibilidad sintáctica (1995: 367).

In agreement with this, Estebas-Vilaplana also shows that English and Spanish resort to different mechanisms to highlight the most informative part in an utterance. While in Spanish there is a syntactic reorganization of the sentence with the respective displacement of the element carrying the new information to the end of the utterance, in English, conversely, the syntactic structure remains unchanged and what is displaced is the intonational prominence (2014c: 235):

Example 55

Spanish	English
¿Quién canta la canción?	Who's singing?
La canta María.	Mary's singing.

[Estebas-Vilaplana 2014c: 236]

Regarding this syntactic flexibility, which characterises Spanish, López Guix & Wilkinson (2001: 67) state that despite the fact that both languages (Spanish and English) have the same word order pattern (SVO), Spanish enjoys a greater freedom compared to English, and even allows placing the subject after the verb for emphatic and contrastive purposes (*Te lo digo yo!*). Furthermore, this frequent subject-verb inversion in Spanish is, for López Guix & Wilkinson, probably a result of the demands of the trochaic rhythm of the language (óo), different from the basic iambic rhythm characterising English (oó), which tends to accent the last syllable of the sentence (unlike Spanish) (2001: 68).

Example 56

Your father gave it to me. (subject-verb)
  a) Me lo dio tu padre. (verb-subject)
  b) Tu padre me lo dio. (subject-verb)

[López Guix and Wilkinson 2001: 68]

The example above shows how Spanish tends to use verb-subject inversion to keep the trochaic rhythmic pattern, although it is still possible to change the sentence construction (subject-verb) to use a iambic rhythm, just like English, thus placing the verb at the end. Nonetheless, Spanish constructions usually avoid placing the verb at the end.

On the other hand, the placement of the nucleus within an IP is a more flexible matter in English than in Spanish. This is evident in the focusing or accentual function of intonation by tonicity. Caldiz (2012) observes significant differences between the intonational patterns of English and Spanish, in particular between accentual patterns within the IP. In line with this, Mott (2011) and Solé (1989) state that there is no such focusing /emphatic mechanism in Spanish intonation as there is in English, as the following examples illustrate (Mott 2011: 227):

Example 57

The idea was *mine*.   It was *my* idea.

La idea fue *mía*.   Fue idea *mía*. (*Not Fue *mi* idea)

As Mott explains (2011: 227), while in Spanish we have to resort to "mía" instead of "mi" to achieve the desired emphasis, English has the chance of moving the nucleus from "mine" to "my" thanks to the flexibility of its intonation. In addition to this, Solé (1989) states that this focusing device of tonicity in English has no regular counterpart in Spanish: for, "no such device of displacing the tonic (nucleus) to highlight the word in focus is possible in Spanish" because this language has a "different rhythmic pattern which does not allow shifting the tonic about". She gives some examples to show the awkwardness this would produce in Spanish (1989: 187):

Example 58

– Do you take your tea with *sugar*?

- With*out* sugar, please.

        vs

- ¿Toma el café con *azúcar*? *Sin* azúcar, gracias.

  Or

- Are you going to read *Platoon*?

- I'm going to *see* Platoon.   *vs*   Voy a *ver* Platoon.

These Spanish versions would not sound natural in Spanish. Solé describes three ways for translating marked tonicity from English into Spanish (1989: 187-188-189):

a) The Spanish construction does not have a displaced tonic and, consequently, the repeated/emphatic/contrastive word (or syllable) is not in focus or highlighted by intonation. The tonic appears in final position in the tone group.

   "¿Toma el café con *azúcar*? Sin *azúcar*, gracias."

   "Estaba *enfermo*, muy *enfermo.*"

b) Spanish changes the grammatical construction so that the contrasted or emphasized word appears at the end and the tonic does not have to be displaced.

   "Fue idea *mía*."

"Ha salido el sol."

c) Spanish uses an emphatic word or grammatical construction, which might attract the tonic.

"Sí que fue idea *mía*."

"Sí que voy a ver Platoon".

"De entre todos los presentes, sólo41 Pedro conocía la respuesta."

The flexibility of English tonicity when it comes to moving the focus within an IP therefore has its consequences in translation into Spanish, since the Spanish version might have to change the accentual pattern and the grammar of the sentence/IP, and consequently, the conveyance of attitude and meaning implicit in the English version might be problematic (Solé 1989: 188).

Another relevant aspect of the way in which tonicity can be translated is that raised by Schwartz (2011: 400), who introduces the notion of nucleus placement in voice-over translation stating that word order in different languages can create a shift of emphasis, which in audiovisual translation may result in a mismatch between the audio and the visual emphasis that spectators can see on screen. For this scholar:

> Verbs in German are often pushed to the end of a clause, while in English, French, or Italian the verb follows the subject more closely. In the translation, the nucleus will be in another part of the sentence and will therefore no longer coincide with the accompanying movement. (Schwartz 2011: 400)

This brings back the idea of how relevant synchrony is in audiovisual translation and how carefully translated scripts should be in this regard. Script length and the right choice when identifying the nuclear accent, i.e. the focus, together with the right delivery, are paramount for an adequate translation.

2.5.1.3 **Translating tone**

---

41 According to Tovar (2014: 07) The Spanish particle "sólo" always introduces new information and has a focusing role.

In written translation, the translation of tone has usually been done by means of punctuation marks (e.g.?, !, "") in order to show attitude; sometimes the mood of the speaker has been conveyed through lexical specification. According to Solé (1989), tone direction is normally expressed in written discourse through adverbs and verbs, which can also show other characteristics of phonation, such as pitch range and voice quality. She provides the following examples (Solé 1989: 192):

Example 59

- "She's nice", he said cheerfully.

      he said passionately.

      he said groomily.

      he grudged.

For Solé (1989: 192*)* Spanish and English do not use exactly the same tunes, or tone directions, to convey the attitude indicated by the lexical item in those examples. For example, a fall-rise tone is not usual in Spanish, which poses a "hard task for the interpreter" (1989: 192) and it can be a hard task for the voice-over translator too. Since there is not a one to one correspondence between English and Spanish as regards the tones used to show attitude and mood, the translator must strive to find a suitable tone in order to convey the intrinsic value of the original tone in the source text.

There is a type of construction in English in which the tone used has a significant pragmatic and semantic role, or *pivotal role* (Reese & Asher 2007: 448) that is substantial in terms of discourse function, and therefore, translation: question tags. As we saw in section 2.2.2.4, depending on whether the tag rises or falls, the meaning can be that of a question, a statement, and an opinion. Question tags are usually translated into Spanish by means of *ratified*[42] questions, which include a statement, followed by *¿no?* or *¿verdad?* as in the example below.

Example 60

Te gustan las lentejas, ¿no?

---

[42] Cid Uribe (1996: 63) calls the Spanish version of English question tags "preguntas ratificadas", also present in Álvarez del Valle (2015: 168) PhD thesis.

Example 61

Sabes nadar, ¿verdad?

However, according to Cid Uribe (1996: 63), while the tag in English fulfills a semantic function depending on the type of tone (rise or fall), Spanish only uses the rise in the ratification particles (¿no?, ¿verdad?), which, does not reflect any semantic value as tags do in English. Therefore, translators should be careful when dealing with questions tags and using ratified questions in Spanish because there could be other available options to translate the tag.

Example 62

	"You want to marry me when my husband's dead, \don't you?"
Translation	"Te gustaría casarte conmigo cuando el rey muera, ¿verdad?"
Alternative	"Sé que querrás casarte conmigo cuando el rey muera"

The example shows a tag question with falling intonation which, in this case, works to state an opinion. The speaker, Anne Boleyn, seeking an answer but confirming her thoughts.

Apart from ratified questions, tag questions with a falling intonation can be translated as *non-neutral interrogative statements* or *absolute exclamatory questions*[43], which, according to the RAE (2011: 469) have a rich pragmatic value, since they can imply commands, surprise, even hypothesis confirmation. These types of questions have a more ample rising boundary and are higher in pitch at the end compared to standard questions in Spanish.

Example 63

	"You can't get rid of them now, can you?"
Translation	"¿Cree que puede deshacerse de ellos?"

---

[43] These constructions are called "enunciados interrogativos no neutros" o "preguntas absolutas exclamativas" by the RAE (2011: 469)

As we can see, tone contributes to meaning in many different ways, and considering that intonation plays a major pragmatic role in English, the translated script should show those characteristics if what it is to be achieved is a faithful translation.

2.5.2 Other English prosodic elements to be considered in AVT

(…) it would be erroneous to think that the written representation of paralinguistic information, such as stress and intonation, is just a matter of voice-over translation performance and not of translation proper.

<div align="right">Franco (2000: 95)</div>

After reviewing the relevance that intonation has in English communication, and how different its usage is compared to Spanish, - and consequently for translation (fundamentally in oral texts) - the following sections will be devoted to reviewing those elements of English intonation which should not be obviated, or skipped, by the English-Spanish voice-over translator.

2.5.2.1 Anaphora rule - Deaccenting

One of the prosodic phenomena characteristic of English concerns the different prosodic behaviour of new information, which is always accented, and given, or known, information, which is deaccented, as stated in section 2.2.2.2 Translators need to be aware of this relevant prosodic feature to provide a reliable and faithful translation by exploiting the resources available in Spanish which have a similar communicative effect. In Spanish, unlike English, we tend to give prominence to even words which are repeated, or which have already been selected previously if they appear in final position, which is typical of Spanish nuclear accent. As a consequence, Spanish:

(…) evidencia una tendencia a mantener el acento nuclear al final del grupo tonal, incluso sobre palabras que no agregan información nueva por haber recibido prominencia prosódica con anterioridad (…) Esta característica en nuestra lengua incide significativamente sobre la reacentuación de información dada. (Caldiz 2012: 01)

Therefore, in Spanish there is no such thing as deaccenting, but the very opposite one, a reaccentuation of given/old information.

García Lecumberri (1995: 197) mentions that "there are two strong accentual tendencies in Spanish: nuclei tend to fall on the last lexical item, and given information is usually accented"; while Mateo, also discussing how each of these languages expresses contrast and the fact that the deaccenting of given information is not a common phonetic process in Spanish, observes that this language "tends to prefer elision of repeated items to make a contrast clearer" (2014: 127), as in the examples taken from Mott (1991: 236):

Example 64

- Do you want a room *with* a bath, or *without* a bath?
(Translation) ¿Quiere una habitación con *baño* o sin *baño*? OR: ¿Quiere una habitación con o sin *baño*?

As can be seen above, Spanish reaccentuates the word *baño* or eliminates it to show the contrast. The usage of elision could be a reasonable mechanism in voice-over to avoid word repetition, thus reducing script length.

For Mateo, deaccenting is also an important presupposition trigger, "since it is applied by speakers on the judgements about what they share with the receivers" (2014: 128). One cannot take for granted that everything in a source script should be considered equally relevant and brought into focus by the distribution of prominence in the utterance, something which the translator should bear in mind.

Mateo exemplifies the relevance of deaccenting in translation and how meaning could be altered if this prosodic feature is not carefully considered, with her translation

of the following example taken from Mott (1991) (deaccented words are reduced in size) (Mateo 2014: 120):

Example 65

- Doesn't he *know*?   - I've told the fool *twice*.
(¿No lo sabe?) - ¡Al idiota ese/A ese idiota ya se lo he dicho dos *veces*!

As Mateo explains, it is thanks to the usage of the deixis of the demonstrative (*ese*) that it is clear for the Spanish recipient that the speaker is referring to the same subject as the question. In addition, the placement of the deaccented information in theme position (the fool - al idiota ese/a ese idiota) helps keep the focus in final position in the translation; and:

> What a translator into Spanish should be aware of is that not reflecting the value of the deaccented source-text item by means of some lexical or syntactic device in the target text may produce a different pragmatic value: "Ya se lo he dicho al idiota dos veces" would keep the critical tone but it would now not be directed to the same person referred to in the question but to somebody else who probably knows him, so the illocutionary act of the utterance will have changed (Mateo 2014: 120).

The observation Mateo makes here about how ignoring the deaccenting function of English can hinder a good translation shows that prosodic features, often neglected in translation, are extremely important. For voiced-over products, in which one might find two speakers, or more, talking about the same thing (or object), deaccenting will probably be a common phenomenon, so translators should pay particular attention to the role deaccenting plays and how they could transfer that resulting interpretation into Spanish. Nevertheless, one should bear in mind that voiced-over scripts are probably often translated without the translator watching the video (i.e. from transcribed scripts), so that the deaccenting of given information is not regarded, thus leading to possible misinterpretations and therefore mistranslations.

In the translations of the example above, Mateo also shows that syntactic reorganization in Spanish helps the recipient identify given information. Let us consider the three translations of the example *I've told the fool twice*.

A – Al idiota ese/A ese idiota ya se lo he dicho dos veces.

B – Ya se lo he dicho al idiota dos veces.

As the author clearly explains (2014: 120) the less suitable translation is B, not only because there is no usage of an anaphoric pronoun (*ese*) to refer to the same subject as in the question, but also because the object has not been placed in theme position in the sentence, which is a common resource in Spanish to draw attention away from a given element. According to the RAE (2011), syntactic variation in Spanish influences the focus structure and the thematisation process works in order to distribute new and given information:

> Las lenguas románicas suelen emplazar el foco informativo al final del enunciado, y lo marcan prosódicamente como el elemento más prominente. La DISLOCACIÓN de la información temática es un recurso comúnmente empleado para situar la información conocida en una zona periférica del enunciado, antepuesta o pospuesta a la oración principal, con el objetivo de conseguir mantener el foco al final del enunciado principal (2011: 483).

Some examples of this "dislocation" of information are:

<div align="center">Example 66</div>

De fútbol, no pienso hablar contigo.
El cordero, aquí lo hacen muy bien.

<div align="right">[RAE 2011: 483]</div>

In the examples above *de fútbol* and *el cordero* would be the given information, or the least relevant information, and thus they are fronted to place the focus on the significant part of the two messages: *contigo*, and *muy bien.*

The opposite movement of information, that is placing the theme to the right of the rheme (*posposición temática*) is a less usual way of focusing in Spanish according to the RAE, as in the sentences in the following example:

Example 67

Ya no me habla, *mi antiguo profesor*.

Yo ya se lo dije, *que no encontraría casa*.

El traje de lentejuelas *se puso la muchacha*.

[RAE 2011: 483]

For Mott (2009), fronted nuclei are not as abundant in Spanish as in English, but they play a significant role in Spanish intonation. Some of the factors that cause this situation are (2009: 74):

- Contexts where strong contrast or anaphora influences the speaker's choice of placement of nucleus.
- The degree of emotional arousal elicited by the context (amazement, incredulity, impatience, exasperation, etc.)
- The use of interjections (like por favor `please´, gracias `thank you´) and other words of a parenthetical nature in final position.
- The use of nominal preforms after the verb whose identity is taken for granted and which are therefore unstressed.
- The semantic force of particular words in the utterance.

This moving of the nucleus and focus to the left of the sentence and the IP shows in Spanish, according to Caldiz (2012: 02) and Caldiz (2009), marked contrastive forms, with a high polyphonic touch. This is, in fact, also a resource of Spanish to show contrastive focus, and it is precisely what Chaume (2001) suggests audiovisual translators should use, together with juxtaposition and coordination: "procurar la topicalización de los elementos más relevantes informativa o expresivamente" (2001: 84). According to this, then, and following Chaume's suggestions, another possible translation of the example Mateo used to show deaccenting in English and Spanish would be: "Dos veces se lo he dicho a ese idiota/al idiota ese", as shown below:

Example 68

Original: Doesn't he know?   I've told the fool twice.

Mateo: ¿No lo sabe?	¡Al idiota ese/A ese idiota ya se lo he dicho dos *veces*!

Using Chaume's recommendations:

¿No lo sabe?	Dos veces se lo he dicho ya a ese idiota/al idiota ese.

Now, if one considers the advantages and disadvantages of topicalisation of given information and the rhematisation of the most relevant part in audiovisual translation, and particularly in voice-over, I would suggest that the best way to translate such a sentence would be that proposed by Mateo, the reason being that the two sentences (original and Spanish translation) coincide in the last words (*twice* and *dos veces*). It is important to remember that, according to most experts' recommendations for voice-over translation, for authenticity and faithfulness reasons, translated versions should try to start and finish with a literal translation of the corresponding bits in the original script so the viewer perceives (if they can understand the original language) that what they are being told is true. In addition, it is now worth recalling Schwartz's comment about the risk of moving the nucleus around in audiovisual translation, since the relevant word may no longer coincide with the accompanying movements of the person speaking on screen (2011: 400). Therefore, if that sentence was to be said in a voice-over video, the last translation suggested would show a mismatch between the English and Spanish nucleus:

Example 69

I've told the fool *twice*.
Dos veces se lo he dicho *al idiota ese/ese idiota*.

This would be noticeable for listeners/viewers if the speakers accompanying movement was some kind of gesture showing two fingers to reinforce that the number of times he told the fool was two.

2.5.2.2 **Stress and rhythm**

There are other very relevant features of English pronunciation which can affect English-Spanish translation. Speech is highly rhythmical (Solé 1991). It may be affected by the way we breathe, our heartbeat rate, the speed at which we talk, or even who we are talking to, and it tends to have a regular beat. That regular beat is rhythm, which does not work in the same way in every language, thus making recipients perceive different rhythmical sensations from speakers of other languages (RAE 2011: 445).

As is well known, English and Spanish show different rhythmical patterns, which may influence audiovisual translations. English may be considered a stress-timed language, as opposed to Spanish, which can be classified in the group of syllable-timed languages. In the former type, stressed syllables tend to occur at equal intervals of time, regardless of the number of unstressed syllables between them (Gómez & Sánchez 2016). This is commonly known as rhythmic isochrony, which means that a language, in this case English, can extend or reduce the duration of syllables between two stresses so the distance between them is the same (RAE 2011: 446); Solé (1991: 154). On the other hand, Spanish, as a syllable-timed language (Martínez-Celdrán et al. 2003: 255) shows syllabic isochrony (RAE 2011: 446), so that the distance between syllables is the same, regardless of whether they are stressed or not, thus creating a "machine-gun" effect (Mott 2011; Fuchs 2014; Gómez & Sánchez 2016). Although stress-timing "is only a perception 44 " (Ashby 2011: 166), the typical English rhythm produces a considerable effect of reduction of unstressed weak syllables. However, such an evident and noticeable reduction does not happen in Spanish. Consequently, in a specific time frame (the same for English and Spanish), the number of syllables uttered by a speaker will be higher in English than in Spanish, which can influence the synchronicity of the original video and the translated audio.

In English into Spanish translations in general it is common that the translated text is slightly longer in number of words than the original, "but the number of syllables

---

44 According to Ashby (2011: 166-7): " (...) J. D. O'Connor demonstrated that in fact, the actual timespan between one stress and the next was very different, even though we perceive it as being regular and equal."

increases on the order of one half to two thirds" (O'Donnell (1990) in Reynolds & Sizemore 2010: 471). On the one hand, the fact that in voice-over the translation starts and finishes a couple of seconds after and before, respectively, *"demands for lexical reduction in the translation so that the oral delivery does not sound rushed or unnatural when recorded"* (Díaz Cintas & Orero 2006: 478).

This is one of the biggest issues of voice-over translation, particularly in this language direction we are studying here, and in my experience as a voice actor, there have been times when scripts have had to be adapted (reduced) during the recording session since there was no chance to adapt the video at that stage. It will be interesting to know what the voice talents and translators interviewed for this dissertation have to say about the issue of text length and voice pace. It is not unusual for a voice talent to be requested for a retake after the recording session has finished due to pace matters. This rhythm issue - closely determining the length of the target text and the duration of its delivery - is, for some, the Achilles' heel of voice-over translation, and it can hinder action synchrony when the person speaking refers to an element on screen, since the Spanish voice will probably come a bit too late in that case.

Concerning reading speed and its relation to the length of translations, Mayoral (2001: 09) states that:

> Al igual que se habla de "voces superpuestas", en la traducción audiovisual se podría hablar de "velocidades superpuestas" …el traductor de voice-over y el de doblaje debe hacer que su texto pueda ser leído cómodamente a la velocidad de dicción de los actores de doblaje.

The issue of syllable count was one of the production elements Reynolds & Sizemore (2000) identified in their project about anti-telemarketing Spanish voiced-over videos for the elderly population in the USA. They also paid attention to the voices that would perform the script, and they realised that they had to re-edit the video track to synchronise it with the voice-over audio track. They did this in two ways:

> Two techniques were used to edit the visual track. When audio for an onscreen character was voiced-over, the editor inserted visual clips from the end of the shot to

create a freeze frame effect and extend the visual track to accommodate the longer duration of the voice-over audio. When off-screen narration communicated information on a topic or introduced a new topic, shots or continuous camera recordings were inserted or action from the same shot was repeated to expand the visual track. (Reynolds & Sizemore 2010: 473)

In projects like this, in which the video can be adjusted (extended or reduced) so that it matches the translating voice, the issue of script length and syllable count would not pose any problem. In any case, this supports Cattrysse's (2001: 03) claim for multimedia translation in general, in which translators "must collaborate closely with the graphic designer and the programmer", and this can certainly be applied to the modality of voice-over, particularly to projects in which there is the chance to have the translator and the video editor working together. Nonetheless, for multilingual projects for which a script has to be translated into several or many languages, like BBC documentaries, adapting the video for each translated language seems an impossible task.

The following example was taken from a script for an online video for an investment company, shows how translations can change, or hinder, pace in voice-over.

Example 70. Example of long translation taken from TN17-005[45]

5	Sitting in London Brexit stares us in the face	Desde nuestra sede en Londres, el brexit nos mira directamente a los ojos.
6	Will it be hard? Soft?	¿Se optará por una versión dura o blanda?
7	Orderly? Disorderly?	¿Se llevará a cabo de una forma ordenada o desordenada?

Looking at this example, is obvious that word count is a serious issue in translation. The higher number of syllables in the target text, along with Spanish syllable-timed rhythm, will require a fast pace to keep up with the original English voice, which, in turn, will make the final result rushed, "machine-gunned" and probably

45 Full text and video available in the annexe included at the end of this dissertation.

unintelligible for the audience. A fast pace will always blur intelligibility; consequently, hindering the right conveyance of the message that is tried to be transmitted in the video.

This is the challenge Matamala refers to when dealing with voice-over translations in which the speakers use different registers (2008: 121):

> The challenge lies in rewording the original to fit in the time available and, while maintaining the same register, creating a text that the voice artist will be able to read easily and the audience will understand without any problems, taking into account that they will only listen to it once.

Matamala hits the nail in the head claiming that voice-over scripts have to be easy to read and understand and reminding us that the audience will probably listen to it only once. She regards the script from the perspective of the voice talent and the audience, who are the last elements in the translation process (delivery of the script and reception), and therefore often disregarded, since translations focus more on text than delivery as is probably the case with the example above and others that will be shown later on in this dissertation.

2.5.3 **Voice quality**

Voice quality is defined by Crystal (1975: 92) as the "idiosyncratic, relatively permanent, vocal background of an individual, which allows us to recognize him, as opposed to other members of the group." Therefore, given the oral nature of voice-over discourse, it seems advisable for the distinctive oral features characteristic of human voices to be delivered in the translated voice. As Gambier & Gottlieb put it, "translation does not end with `text´ but with delivery" (2001: xix). The fact that voice-over shows the two voices (original and translated) is for Díaz Cintas and Orero (2006: 478) "the ideal translation solution if one wants to give an idea of the interlocutor's tone and way of speaking"; but this implies that voices should be taken into consideration when planning for a voice-over translation. After all, it will be the translated voice that the audience will get when they are at home watching TV. They will not have access to the

script, will not know whether the translator had enough time to do his/her job or not, whether the producer had enough money to hire the studio time the project required. The voice the audience gets at the end carries the burden of the whole translation process (transcription from original video – if no script is provided -, translation of original script, adaptation of translated script, and direction of recording at the studio). As Franco states (2000: 257), "performing voice-over is to give life to translations." If any of the translation stages fails, the voice will fail, and will be blamed by the audience at home (e.g. "it is too fast", "it's boring", "he doesn't say that", "he/she's too excited" ...).

Bosseaux (2008) considers that the issue of voice is still a "rather under-researched topic in film studies" (2008: 350). In her study of the musical episodes in the series *Buffy the vampire slayer*46 (1996-2003), she observes that what the characters feel - their inner thoughts and emotions - is conveyed by means of gestures, facial expressions and camera movements, all of which remains intact in translation. However, "what cannot stay intact is what the characters say and how they say it, a dimension of performance that can be broadly described as `voice quality´" (2008: 350). Performance in voice-over may not be considered as crucial or important as, for example, in dubbing, but still it is relevant, as long as the voice reaches the target audience.

It may be true that the vocal performance and voice characteristics in voice-over "do not need to resemble those of the original program and the voice-over speaker – revoicer – does not need to have trained as an actor" (Díaz Cintas & Orero 2006: 478). However, this statement may be debatable when one considers the kind of voice-over category being translated (news interview, documentary, reality shows, etc). What seems clear, though, is that, voices do not matter much in audiovisual translation so far, which is rather surprising, considering that they carry the weight of the whole translation process both in dubbing and voice-over.

There are, fortunately, some exceptions, that show that a specific voice has been regarded as highly relevant for a film, at least both in the original and the dubbed

---

46 Series information available here: https://www.imdb.com/title/tt0118276/

translation, such as Darth Vader's original English voice in the film *Star Wars* (1977), performed by James Earl Jones and not the actor who played Vader's character, David Prowse. Apparently, George Lucas wanted a deeper voice than Prowse's for the movie to give the character the gravitas and fearsome nature he deserved. Accent was also relevant for the role since Darth Vader has an English accent, as opposed to the Rebel Army (the "good guys" in the film), who speak American English. Endowing the *baddies* with a British accent seems to be a common strategy in American productions; even Disney films use is to differentiate good from evil characters as observed by Fattal[47] (2018).

Another slightly different but also very relevant example could be Spanish Buzz in *Toy Story 3* (2011). In the original film there are around 15 minutes in which Buzz Lightyear, one of the toys in the story, is reset in Spanish mode by mistake and the other toys do not understand a word of what he says. In the original American film, Buzz Lightyear was dubbed by Tim Allen, and in the translated dubbed Spanish film, the character is performed by José Luis Gil, the voice actor that has dubbed the whole trilogy in Spanish. One would have thought that if Gil is the voice of Buzz Lightyear in Spanish films, he should have been Spanish Buzz in this episode of the original American version of, since the producer is the same, the director is the same, and the character is the same. Despite all this, for the original version of *Toy Story 3* the voice actor chosen was Javier Fernández-Peña, rather than José Luis Gil, which shows that for some projects choosing the right voice matters. In the process of selecting Javier's voice, the production team listened to hundreds of demos until they decided to go with Javier's tone of voice, as explained in the interview[48] with the casting agency. What is interesting about the fact that Buzz speaks Spanish in the original film is that for the translated Spanish European version they decided to make Buzz speak in `andalú´, Andalusian Spanish, for which Pixar selected the flamenco singer Diego Jimenez, a.k.a. *el cigala*. In the Spanish Latin version, however, Buzz speaks Castilian Spanish with a poor vocalisation in that episode, voiced by José Luis Orozco. In total there were four different voices in the Spanish versions for this single character. However, the use of

---

[47] Online article link available in the bibliography section.
[48] Interview available in Annexe 6.

sociolinguistic markers, like accents, in the translation of "'serious´ genres" such as dubbing, is quite rare (Chiaro 2009: 159).

It is impressive how voice quality can determine the nature and idiosyncrasy of a character in a film. There is a large list of American actors who have devoted much effort to cultivating and widen their range of voices, such as Robert de Niro, Philip Seymour Hofmann and Meryl Streep, "the vocal chameleon par excellence" (Gorbman 2014: 09). This proves that voice quality is a relevant factor in the creation of an audiovisual product, and, therefore, it should be considered once this product is to be translated.

Returning to the role of voices in voice-over translation, it is worth pointing out that in Poland, where voice-over is used to translate most of the audiovisual products, the voice used is normally neutral and according to Sepielak (2016: 1057), "all practitioners emphasise [however] that voice artists should have clear elocution and read continuously with a monotonous tone and a stable reading tempo." Robert Paquin seems to agree with the situation in Poland: "the studio actor/interpreter does not play a role with emotions as it were acted out. He or she is merely an interpreter who repeats what the person is saying" (1998: 01). Nonetheless, Paquin appears to contradict himself in the same article, since when treating the issue of dramatic synchronism, he states that:

> It's important that the characters speak with a certain amount of realism […] Language level, use of idiomatic expressions, realism are all factors to be taken into account. The character has to "sound" real (1998: 04).

It is difficult to imagine how anyone can sound real without being emotional. If someone on screen were showing happiness or sadness, would the audience not expect some sort of high pitch and energy for the first context, and the opposite for the latter? Admittedly, different voice-over genres (interviews, news reports, reality shows, etc.) have different types of delivery, but that does not necessarily mean that there should be no emotion at all in some of them. Franco highlights the importance of performance in voice-over translation in order to make documentaries less dull and boring: from the audience's point of view, a good performance is needed, in terms of stress and

intonation, to break the boredom taboo generally associated with the voice-over mode; therefore, Franco insists, "stress and intonation are important factors in translation performance" (2000: 259), and, we may add, they are indeed so during the whole translation process.

Referring to Zabalbeascoa (2001), Barzdevics (2012: 60), agrees with Franco's view on voice-over performance, claiming that voice-over has something in common with acting (although not as much as dubbing), since it is necessary to observe the speaker's body language and paralinguistic features. This somehow clashes with Díaz Cintas and Orero's view that the vocal performance in voice-over is simplified, so that the revoicer/speaker does not need to be an actor. Whether the revoicers are actors, or not, can be debatable and may depend on several factors: first, some of the revoicers are journalists, in which case they may not have acting skills, second, if the person in charge of doing the voice-over is an aficionado[49], or if it is a freelance voice talent with a basic studio fee of €200 per hour (plus buyout), in which case it will normally be a voice actor.

What seems clear is that a poor voice-over performance will probably contribute to the failure of the programme. One cannot imagine a neutral, inexpressive voice performing Bear Grylls' adventures in a voiced-over version of *Ultimate Survival,* considering how expressive he is in the original. Voice-over performance is directly linked to authenticity and the "realism" illusion (Franco 2000); and it is also linked to the success of the product, be that a reality show, a corporate video, or a documentary, as was the case with the translated Spanish version of the BBC documentary series *Life* (2009)*,* which was studied by Garrido (2013: 213), who concluded that:

> Habiendo tenido acceso únicamente a la reproducción oral de la narración, registrada en el video, y no a su plasmación escrita, nos hacemos cargo de que, en algunas ocasiones, las inadecuaciones aquí reseñadas podrían deberse, no a la impericia del traductor, sino, acaso, al desempeño deficiente del respectivo locutor o actor de voz.

---

[49] The burst of "professional" online voice talents in websites such as Fiverr.com, gives the impression that, nowadays, anyone with a computer and a microphone is a voice actor.

This cannot but reinforce the idea that a poor voice performance, caused by a poor delivery and/or a poor translation, could ruin even expensive and good quality projects like *Life*, which took four years to complete, and had a budget of £10 million Garrido (2013: 231).

The issue of what voices are chosen to perform a voice-over translation does not seem to have received much attention (Díaz Cintas 1997 and Orero 2005 in Orero 2006). However, the choice of the right voice for a project is paramount, since voices "give life" to translations. Character synchrony, one of the main characteristics of voice-over and dubbing, as already stated in the first chapter of this work, contributes to the sense of realism and authenticity necessary in voice-over. The actual choice of a particular voice seems a matter of the production company. However, in the case of news agencies, (voice-over for production), it is probably the journalist who translates the script and voices it, no matter whether the original voice is masculine or feminine. For certain post-production projects, a casting might be conducted so the client has the chance to listen to several voice actors and then choose the one they consider right for their project, be that a documentary, a film, a video-game, a TV series, etc. Castings can be conducted in a specific studio, and actors may know or not the type of product they are casting for, or the lines to voice, as sometimes there is a lot of secrecy and mystery about the final product. This normally involves signing an NDA (non-disclosure agreement) for confidentiality matters, as is usually the case with big companies producing apps, and voice recognition software, although they might not be the only ones. Sometimes, the casting is conducted online, via the so-called "pay to play" voice-over websites, which are open to "premium" voice-over members (those who pay the annual fee). For this kind of on-line casting, a piece of the script is normally given to the talent together with some instructions or an extract from the original video, so that he or she may have an idea of the right delivery. Sometimes no script is given and talents only need to send a voice demo and hope to be chosen.

The way voice-over and (multi)media/localisation agencies work differs from one agency to another. Some may conduct a casting for their projects, others not. This issue will be dealt with in detail in a later chapter, but I will here present an overview on how the process according to Orero (2006), who contacted IMS (Independent Media Support Group plc.) about it:

Our multilingual production team selected highly experienced artists who were able to convey the tone the film demanded in an appropriate manner to each language. Understanding of the cultural differences is hugely important. For example, for the French version we chose a woman's voice, whereas a male voice was more appropriate for the Turkish video. (2006: 175)

Orero concludes from this statement by this production company that "the voice – with its many qualifiers – may be one more important element to create the feeling of hyper-reality and generate what has been called the `constructed reality´ (Orero 2006: 175).

Voice, therefore, is another important element, and as such needs further research. Soh Tatcha (2009) delves into vocal equivalence in the realm of dubbing, and his findings can be extrapolated to voice-over, since voice actors do both dubbing and voice-over. For him, the search for vocal equivalence is done, to some extent, by trial and error, and depends mainly on the director's experience. For Soh Tatcha (2009: 52), the voice chosen should have some special characteristics:

La recherché de la voix de doublage équivalente á la voix orginale procéde d'une demarche similaire: il s'agit de trouver la voix "V", non identique á la voix originale, mais équivalente sur le plan de l'expressivité, de la sensibilité emotive et de son aptitude á accrocher l'image, capable de coller au corps qui parle á l'écran sans laisser l'impression de "patchwork" et d'artificialité qui se degage de certains films.

Voices are usually defined in terms of *colour* or timbre, which can be defined, as "that particular attribute of a given voice which distinguishes that voice from another when the vowel and the pitch are the same" (Cleveland 1977: 18), or as "an individual's particular repartition of acoustical energy across frequency" (Latinus & Belin 2011: R143). These *attributes* or *particular repartition of energy* refer to harmonics, which are secondary or complementary frequencies that accompany the fundamental frequency, and which are responsible for the characteristic timbre of a sound source (be that a human voice or a guitar). These specific nuances characteristic of the human voice make us recognize voices and remember them. For Latinus & Belin (2011: R143),

although a speaker may never utter twice exactly the same sound, listeners extract these invariant features in the vocal signal to build representations of a speaker's identity.

We can analyse and compare the harmonic properties of voices by means of specific software, as will be seen in chapter 4, which show graphic representations of these frequencies in a line graph. This will be highly useful in our analysis section in order to compare voices in the character-synchrony section. How voices sound depends on many different factors, starting from the physical condition of the speakers, their age, the size of the vocal folds, their health status, the breathing rhythm, etc. Some of these factors are visible, other not. This is why we will resort to the harmonic spectrum; to have a clear picture of how voices sound.

All these voice quality features and the way voices are selected deserve careful consideration, particularly if we take into account the different factors that can affect the production/post-production of a voice-over. One voice may be chosen over another for voice quality and performance reasons, but it can also be the case that one voice "gets the gig" for economic reasons. The way the industry is evolving and the recent boom of online home-based voice actors (both professional and amateur – see Mayoral (2012) makes the business a competitive challenge moneywise, which might tip the scale in favour of one actor or another.

2.5.3.1 **Pitch range**

The frequency of vibration of the vocal folds is perceived by hearers in terms of what is known as *pitch*. There are two systems of pitch: tones (fall, rise, etc.) - which have already been discussed in section 2.2.1 and pitch range (Crystal 1975: 94). English and Spanish both use pitch range for pragmatic purposes which should be worth considering in audiovisual translation. Thus, Estebas-Vilapalana (2014a: 183) observes that the use of unexpected pitch patterns can result in social and interpersonal problems, for instance if the message is perceived as offensive or impolite. This is very relevant for the performance of a voice-over translation since Spanish voice actors might strive to imitate the pitch level used by the original English voice when the values of that particular level in Spanish may be rather different from their communicative/expressive

role in English. In English, for instance, broad pitch range indicates interest and involvement on the part of the speaker (Lindsey 1981: 18), therefore signalling politeness according to Estebas-Vilaplana (2014a); narrow pitch, on the other hand, may convey rudeness and unexpectedness (2014a: 192). In Spanish, however, broad pitch range can be judged as over-excited or conveying something unexpected, since the normal pitch range associated with politeness is narrow (Estebas-Vilaplana 2014a: 192).

	Broad Pitch	**Narrow Pitch**
English	Interest, involvement, politeness	Unexpectedness Rudeness
Spanish	Unexpectedness Over-excitement	Politeness

Table 17. Pitch range and its perception in English and Spanish

Estebas-Vilaplana (2014a) exemplifies these differences with the answer to a question such as *How was the food?* The picture below shows how each answer (English *It was lovely*, and, Spanish *riquísima*) is perceived differently in each language according to the pitch level used.
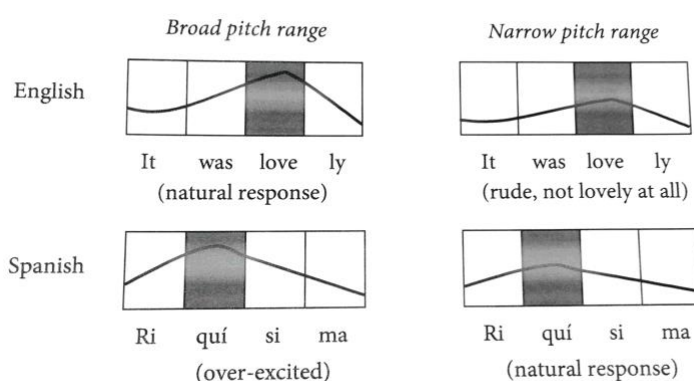
Figure 13. Pitch range and pragmatic differences in Spanish and English

As can be seen, English and Spanish use opposite pitch levels, so that the same pitch range will carry opposite conversational implicatures that trigger non-explicit

messages (Estebas-Vilaplana 2014a: 182). Therefore, this different behaviour in the use of pitch range should be another prosodic element to be considered in any audiovisual production involving English-Spanish translation if an equivalent and faithful pragmatic representation of the original message is to be achieved.

2.6 Concluding remarks

In this chapter we have analysed the pragmatic, syntactic and sociolinguistic relevance of the prosodic components of the English language in oral communication and their pertinence in English-Spanish translations, especially for voice-over. Given the oral nature of voiceover translation and the role of intonation at the centre of pragmatic meaning (Mateo 2014: 129), it seems clear that translators for this audiovisual modality should pay more attention to the communicative purposes (evident at times, subtler at others) implicit in the prosody of English. In Mateo's own words (2014: 132):

> In any case, since intonation is such an effective and common means of pragmatic value in English, it should be among the options and alternatives considered in translation to and from this language, not only in translation produced in the oral mode (e.g. interpretation or dubbing) but also when the process implies a change in the communication channel, as is often the case in AVT –e.g. subtitling, surtitling- or even when it involves the reception or production of texts in written form which are to be delivered or imagined as oral texts –e.g. drama scripts or dialogues in novels.

"The audiovisual medium enables the viewer different forms of expressions" (Pettit 2004: 34), which is why elements such as gestures, intonation and tone of voice may all affect the way in which the verbal content is perceived. It is precisely because the audiovisual medium gives us the opportunity to perceive all the emotions that the characters on screen show, that these elements should not be let aside in voice-over translation.

Intonation may be considered universal in conveying emotional content and attitude, as Patricia Ashby observes (2011: 179): "People the world over convey

something of what they are feeling through what the layperson often calls their `tone of voice´, through their intonation". However, the way in which each language uses the features of intonation to convey those feelings will vary a great deal. This is the reason why Mateo (2014) referring to House's (2001) view of teaching translation pragmatically, emphasises the relevance of including the discipline of phonetics and phonology, and the functions of suprasegmental features, in the teaching of translation, in order to provide students and practitioners "with more solid grounds for the choice of translation strategies" (2014: 131).

These strategies translators use in voice-over are at the core of the present study, and we will try to expose the lack of attention paid to suprasegmental elements by translators. during the process of voice-over translation. After examining the theoretical background of this phonetic component of language and its relevance for English-Spanish voice-over translation, we shall now proceed to state the objectives of this work.

3   VOICE-OVER TRANSLATION AND INTONATION

This chapter is devoted to commenting and highlighting further aspects of voice-over translation and is divided into two parts. The first one deals with aspects such as the oral nature of voice-over texts; the issue of faithfulness in translation and how it can be applied to voice-over so as to contribute to a more suitable audiovisual product; whether the translation and voicing process change depending on the category of the voice-over programme; and the role of the adapter in the voice-over translation process.

The second part concerns the techniques that translators use (or not) to transfer the function of prosodic and suprasegmental elements of English source texts into Spanish target texts; the equivalence achieved in the rendition of the illocutionary force carried by intonation; whether the original video is considered in the translation process; the type of hints voice talents get in scripts to use a more or less marked intonation; and finally, the kind of relationship between translators and voice talents, if any; how this can affect (for better or worse) the final product.

3.1   Further aspects of voice-over translation

3.1.1   Voice-over translation and orality

Our aim here concerns the type of translation that is done in voice-over, and whether or not the original oral text and voice(s) is/are respected (faithfully rendered in the target version) and the way this can be important at all for translation matters. As has already been discussed in section 1.3, there are two possibilities in voice-over translation: translation for production and for post-production. The difference between these two is that, in post-production, translators are given a script to work with, whereas this does not happen in the production kind and translators need to have the script transcribed (Franco, Matamala & Orero 2010). Be that as it may, for all the possibilities that may occur when getting the script to translate for voice-over, what is relevant is that there is a final script in all cases, and it will be translated, adapted (if that is the case), and voiced.

One should remember that in voice-over translation one deals with oral scripts, not just written texts, although what translators are given is a written script, sometimes accompanied by the video file. Thus, the translation of the original text should be considered from a two-way perspective: in terms of content and prosody, as pointed out by Franco, Matamala and Orero (2010: 26): "In sum, voice-over translation has to be a faithful, literal, authentic and complete version of the original audio". We can notice that these scholars do not say the translation has to be a faithful, etc. rendition of the original *written* script, but of the *audio*. In the case of dubbing, in which the original audio is not present, the original text can be manipulated, since the audience will never hear the real voice of the characters (Del Río 2001: 227). This is not the case with voice-over, for which the audio component (the prosodic component) is what has to be transferred, and that implies not only the words that are part of that oral message, but also the function of the tunes, volume, pitch, and all the suprasegmental resources of the original English message. Without translating the original message considering these two fundamental parts (written and oral) the result will never be as faithful and authentic as voice-over is considered. In addition to this, the delivery of the translation by voice actors (voice quality) is also important, since it is what gives life to the translated message, and as such, it should be considered as another element of the translation act. In this work, the role of voice quality in the translation process will play a significant role.

3.1.2 The issue of faithfulness

Fidelity, or faithfulness, is a hot topic of debate among translation theory scholars as seen in section 1.3.2, and it has been a controversial buzz word among translation practitioners too. Chaume (2005b: 11), for instance, considers that one of the standards of quality of a good dubbing or subtitling is that the target text has to be faithful to the source text. In the case of voice-over translation there seems to be a general agreement among scholars and professionals in terms of fidelity and final result; a good final voiced-over product has to be faithful to the original. There are several factors that can affect the faithfulness or fidelity that seems to be required and expected in voice-over

translation, as in any other translation modality. For this reason, let us explore in this section how a faithful voice-over translation can be achieved.

There are time restrictions that normally apply in voice-over translation. These are commonly known as *sound bites* and are a short frame of time in which the target audience can listen the words of the original speaker in the source language. Normally the first few words of the translation match literally the words of the sound bite and serve to show the target audience that what they are watching or listening is a faithful representation of the original. The length of these *bites* can range from +/- 2 seconds for Orero (2004), to - 4 seconds for Ávila, one semantic unit for Carroll, 2 seconds before and after for Espasa (2004), and even a delay of four words by Mayoral (2001) (see chapter 1.3). What these time frames show is clear: there is no agreement as to how long the restriction, or sound bite, should last. This is probably so because there is no given rule for that, and each project, maybe even each scene, has its own rules. In this work, we will try to decipher if the different time gaps or sound bites are random, or if they respond to some sort of unwritten rule and they can vary according to the category of the programme (interview, documentary, reality show, etc.) and the scene, depending on the script and time limits the video has. Nonetheless, the fact that there are time restrictions implies that scripts have to be shortened compared to the original one, which is a hard task in the case of Spanish voice-over, since translations into this language, are commonly longer in number of words and syllables than English scripts (Reynolds & Sizemore 2000, Orero 2006a).

The reduction of the translated text could be done by the same translator or by an adapter or editor (Orero 2004: 79), who is responsible for adjusting the translated script to the time limits of the original, making isochrony work in the final video. The fact that an adapter revises the translation implies that the script, which normally is longer, will be shortened or modified twofold: first by the translator, and then by the adapter. In line with this, Orero states that "the editor may alter and rewrite parts of the text according to his needs" (2004: 79). If the translator does the adaptation, then he/she will have to reduce and adapt the script as he/she considers. Now, let us remember that one of the characteristics of voice-over translation was literal synchrony (see section 1.3.3 of this dissertation), meaning that literal translation at the beginning and the end of the utterances takes place imposed by the fact that we hear the original voice before

and after the translated one. This implies that what is in between the first and last words of the original message will be condensed and adapted to be delivered presenting it as if it were actually a faithful and authentic representation of the original. It is due to this reduction of the target text in the "in between" part that some elements with substantial prosodic meaning may be deleted if they are not identified as pragmatically relevant by the audiovisual translator.

Considering the voice that performs the translation (the voice talent), character synchrony (Fodor 1976) applies here to create the authenticity feeling that surrounds voice-over. However, is character synchrony relevant in voice-over and audiovisual translation in general? For Zabalbeascoa, Izard & Santamaría (2001: 106), in the case of dubbing, voice actors are classified according to a limiting number of categories; these are: villain, hero, heroine, clown, vamp, child and so on (2001:106). Children characters are dubbed normally by women, with the exception of some characters and films, such as *El Sexto Sentido*[50] (The Sixth Sense), or *Harry Potter Y La Piedra Filosofal*[51] (Harry Potter and the Philosopher's Stone), which were dubbed by kids trained in voice acting. The fact that female voices dub, or voice, children characters is understandable considering the following facts: female and children voices have a similar fundamental frequency (as seen in chapter 2), children may take more time in the studio than a professional female voice actor (and that costs money), the restrictions imposed by law on children who can do an adult's work, and the fact that f there are not many children who are dubbing actors. Franco (2000: 259) also identified that the voice of a boy in one of the documentaries of her study was translated by a woman, which illustrates "the illusionist dimension of voice-over translations as representations of `reality´'". Some female voice actors already offer their services as children characters/voices to get more professional opportunities, which is understandable, given the fact that female and children voices have a similar fundamental frequency. However, there are projects for which children voices are specifically demanded, and

[50] Information on the Spanish version of *The Sixth Sense*:
http://www.eldoblaje.com/datos/FichaPelicula.asp?id=526
[51] Information on the Spanish version of *Harry Potter and the Philosopher's Stone*:
http://www.eldoblaje.com/datos/FichaPelicula.asp?id=40

their working conditions and rates are the same as the adults'. In fact, there are voice agencies specialised in children voices, such as, Voiceoverkids[52].

In the case of dubbing, voices are selected according to the film character, not to the original actor, as stated by Zabalbeascoa, Izard & Santamaría (2001: 106), for whom:

> (…)[M]ale characters should be dubbed with low voices, unless the character is a to be laughed at or is effeminate. This produces mismatches between the actor's original voice and the dubbed version. Clear examples can be found in actors like Tom Selleck who looks `macho´ and usually plays the `good guy´, and is given a deep mellow voice in dubbing although his own voice is quite high-pitched and creaky. In contrast to this, Dustin Hoffman is dubbed with a high-pitch voice to fit his anti-hero image, despite the fact that he has a rich deep baritone voice.

What these scholars state about the mismatch between the original voice and the translated one could encumber character synchrony in voice-over since even if the sex is the same in both the original and dubbed and voiced-over translations, there would be a dissonance between the two voices (remember that both are audible in voice-over) if the pitch and timbre differ significantly as in the cases of Tom Selleck and Dustin Hoffman in the dubbing modality. The issue of character synchrony and the voice chosen for each role/project may differ depending on the voice-over category and budget-related issues.

3.1.3 Different categories, different translations?

This third section is related to the fact that translations may change depending on the category of voice-over. All voice-over products imply that, as is common in many oral texts, there is no script "per se", since the script is created as the speakers comment, explain and express themselves about whatever matter during the filming of the programme, unless this has a planned script, as can be the case with certain web videos and documentaries. During the transcription and translation processes of these oral

---

[52] Website of the casting agency dedicated exclusively to children's voices:
http://www.voiceoverkids.co.uk/

texts, the prosodic elements and their pragmatic relevance might be lost, which is why the first quality that the translator for voice-over should have, according to Barzdevics (2012: 60), is a very good ear (*"ha de tener un muy buen oído"*), in order to transcribe and/or, translate directly from the audio track. This, in Orero's (2009) words, "constitutes a challenge to the listening and comprehension skills of the translator" (2009: 136). Let us go through the different audiovisual categories that make use of voice-over as a translation mode.

3.1.3.1 **TV interviews**

The translation of voiced-over TV interviews was studied by Orero (2005) considering the textual characteristics, the translating techniques, and the training of voice-over translations. The conclusions of her work were quite revealing. This category (TV interviews) belongs to the field of voice-over for production, which means that no script is given to the translator. In her study, Orero (2005: 215) first describes the textual characteristics of the translating material from two perspectives: formal aspects, which are directly linked with spoken/oral language and its delivery; and content aspects, which concern the topic of the interview.

The formal textual aspects are themselves divided into three main groups: prosodic elements, paralinguistic elements, and indexical features. Prosodic elements include the pace of the discourse (some people speak using a slow pace, and others a very fast one); accent, or the geographical characteristics of the language; and finally, the linguistic competence of both interviewer and interviewee is quite an important feature, since there may be an interviewee whose mother tongue is not that of the interview. The paralinguistic elements include: the voice of the interviewee, including tone, pitch, etc.; his/her diction (i.e. permanent or temporal diction problems), since this could hinder the translator's oral reception; and body language , which could constrain the reception of the message when interviewees cover their mouths with a hand, for example. Finally, the indexical features are, for Orero, those related to the speaker's tension (like pace), depending on his/her psychological state of mind (tired, excited, angry, bored, etc.), and his/her preference for certain prosodic or paralinguistic elements (2005: 216). According to this scholar, it is not very probable (though not

impossible) for these features to appear altogether; however, in most cases a couple will probably be present in each translation.

Concerning the content textual aspects (Orero 2005: 217), the topic of the interview could be, for instance, technical, thus creating some terminological and comprehension issues. In addition to this, this author acknowledges that the discourse of interviews is purely oral, and mostly colloquial, entailing some particular features, such as (referring to Halliday & Hasan 1985) a low lexical density, deixis, grammatical simplicity, and abundance of repetitions. However, Orero considers that the voice-over translation of interviews does not normally reflect this orality. In her view:

> El producto final es un discurso editado donde se mantienen algunos aspectos de la oralidad y se introducen aspectos del discurso escrito, como, por ejemplo, una sintaxis ordenada, una puntuación correcta, mayor complejidad léxica y cambio de registro de coloquial a formal (y en algunos casos científico). Si se transcribiera el discurso y se tradujera sin adaptación, el resultado sería incomprensible. (2005: 217)

What can be drawn from Orero's words on the translation of TV interviews is that the prosodic and paralinguistic elements characteristic of oral discourse in voice-over are relevant from a comprehension point of view only. The translator has to pay attention to these elements when he/she tries to grasp everything the person being interviewed says, but their pragmatic relevance and their meanings seem to be not fully considered. As it could be seen in chapter 2, prosodic features carry the illocutionary force of the utterance, and should therefore be considered not just as elements which could hinder the reception of the message, but also as powerful meaning carriers that need to be properly rendered in the translation. In addition to this, the possibility that the translated product be not a faithful representation of the original oral discourse clashes with the claimed statement that voice-over is a faithful and authentic representation of the original. Moreover, Orero (2005: 217) claims that without adaptation, the translation of the original script, or of its transcription, would be incomprehensible, and she pleads for a change in syntax, adjusted punctuation, and a change of register (formal), while keeping some oral aspects, which are not specified (probably just the voice of the translation). In her words (2005: 220): "se resume, se edita y se traduce lo imprescindible con un cambio de formato (oral a escrito) y de

registro (coloquial a formal)". In the light of these considerations, can prosodic elements be considered part of the essential parts? There is some kind of funnel effect in the translation process which gives the impression that the final translation could be just a vague reflection of the original message. One of the reasons for this could be the adaptation process and the role of the adapter.

Within the TV interviews category, we have selected for our corpus of study some programmes broadcast by the news channel Euronews. In these, sometimes we can see and hear the voice of the interviewer, but not in others. In most TV news reports (at least in Spain) we usually get the voice-over version, that is the answer, of the interviewee, and hardly ever we can hear the original question, which is, in fact, helpful to identify cases of deaccenting and reaccenting from the part of the interviewee. A full description of these interview programmes is provided in section 4.2 further on in this work.

3.1.3.2 **Documentaries**

Documentary translation is still a largely unexplored field and a neglected genre for which the literature is not extremely extensive, as scholars observed more than ten years ago (Espasa 2004: 194; Matamala 2009: 94; Kaufman 2004: 149 and Franco 2000: 235). This dissertation will not delve into concerns of whether documentaries have been more or less relevant in AVT research, but it will try to highlight the most relevant aspects of the translation of this genre and how these affect the rendition of intonational and prosodic aspects.

Different definitions of the term *documentary* have been suggested by the academic world, being, in our view, Plantinga's (2005: 114-115) the most accurate one:

> …[A]n asserted veridical representation, that is, as an extended treatment of a subject in one of the moving image media, most often in narrative, rhetorical, categorical, or associative form, in which the film's makers openly signal their intention that the audience (1) take an attitude of belief toward relevant propositional content (the "saying" part), (2) take the images, sounds, and combinations thereof as reliable sources for the formation of beliefs about the film's subject and, in some cases, (3) take

146

relevant shots, recorded sounds, and/or scenes as phenomenological approximations of the look, sound, and/ or some other sense or feel of the pro-filmic event (the "showing" part).

Documentaries can be categorised into different types, or subgenres, depending on the topic they cover and how they are done. In this respect, and following Nichol's (2001) approach there are six subgenres: expository, observational, poetic, participatory, reflexive, and performative.

According to Espasa (2002: 33 & 2004: 188), there is no correspondence between audiovisual modes (lip-sync dubbing, subtitling, voice-over) and audiovisual genres, but "in the case of documentaries, voice-over plays a remarkable role" (Matamala 2009: 95). The production studios in charge of translating documentaries remind translators "in capitals" (Espasa 2002: 33 & 2004: 190), that they should not translate texts, but dub them. There exists some terminological controversy with regard to the *voice-over* type that applies to documentaries. At this point, one should recall section 1.3 of this dissertation and the terminological blurriness around *voice-over*. Thus, Luyken et al. (1991: 80), as stated by Franco (2001: 294), affirm that documentaries should be translated by "`extended voice-over´, that is *narration*". This mode of translation refers to the monologues uttered by the narrator in the documentary, and for Franco, Matamala and Orero (2010: 39) *narration* can never be referred to as voice-over since there is no presence of the original audio track in the translated version. Chaume (2011: 03), on the other hand, sees *narration* as a kind of voice-over in which the translation has been summarized.

Another translation mode used for documentaries is *free commentary,* or just *commentary,* as seen in section 1.3, and *voice-over commentary* as is the case of Nichols (2001: 17 & 22), who uses a more industry-like approach of the term *voice-over*. Franco (2001: 297) uses the definition of *commentary* from a Film Studies perspective included in Kozloff (1988: 06), namely, "a style of *discourse* prevalent in documentaries". However, *commentaries* do not always imply any translating activity (Franco 2011: 297), and they may be generated in three different forms according to Orero, Franco and Matamala 2010: 39): "a) translated commentaries closer to the original; b) an adapted/domesticated commentary; c) created commentaries." In addition to this, due

to the absence of the original audio track, on-screen speakers, and any visual evidence of lip-sync, Orero, Franco and Matamala (2010: 41) decided to use the term *off-screen dubbing* to refer to *commentaries*.

Finally, voice-over is another AVT modality used to translate documentaries, at least as regards talking heads (Matamala 2009b: 116). This contributes to transmitting authenticity and keeping the illusion effect, although the revoicing actor normally reads the translation "without expressing emotions" (Matamala 2009b: 116). Voice-over can also be found to translate the narrator part in some documentaries.

The translation process of documentaries has several stages, according to Espasa (2002: 36-37), which include:

1. Video delivery
2. Script delivery
3. Watching the whole video
4. Checking the correspondence between video and script
5. Troubleshooting
6. Translation of the script (in the case of translation for voice-over there should be a larger compression and reduction than in dubbing due to the sound bites)
7. Final checking watching and listening to the video
8. Final orthographic checking
9. Notes (to favour the communication between the translator and other agents, such as, the adapter, dubbing actors etc.)

Kaufmann conceives documentary translation not just as the act of translating the words uttered by interviewees from one language into another, nor watching the narration of a journalist or specialist on camera, or translating the voice off camera (2004: 149) but as an objective reproduction "et non une imitation qui donne l'illusion du meme" (2004: 157) facilitated by lip-sync dubbing.

In her analysis of French documentaries for the French TV channel *ARTE* (2004), Kaufmann (2004) confirms the existence of some sort of censorship, or straitjacket effect, imposed by the TV channel. For instance, in the case of the language

of the Jewish immigrants shown in the documentary *St Jean,* there was a homogenization and standardization of the original language features, imposed both by self-censorship and a control of the submitted translation, as well as by the corrections made at the different stages of the production process. In her view, the translation strategies used in this documentary are similar to those of dubbing: linguistic standardization (reduction or deletion of idiolects, distinctive features characteristic of speech, social and geographical features), naturalization strategy (sociocultural adaptation to pretend the French translations are originals), and the use of a "bien Française" voice and pronunciation by the voice actors (2004: 156).

In line with this, in the case of Spanish documentaries, Matamala understands that the job of the translator is "to render each register faithfully" (2009: 101), considering the limitations of the transfer mode. The usual trend in documentaries is to use formal language, although these are an evolving filmic product, which have gone past the typical off-screen narration formal structure. Thus, this scholar observes that in voice-overs for documentaries, syntax is generally corrected, but colloquial features in lexis are generally kept (2009: 102). As regards the difference between the narrator's part and that of the talking heads. the latter are, according to Matamala, usually full of hesitations and oral features that have to be reworded "in order to offer a more coherent and cohesive speech, unless these elements are particularly significant" (2009b: 116). In addition, in cases in which the interviewee does not speak good English and makes mistakes, these are commonly corrected in the translated version "since what really matters is the content and not the form" (2009b: 119). However, in my opinion, if the form is not as relevant as the content, or if we consider that the way of meaning content is not relevant, how can we claim that voice-over is the best option to offer an authentic and faithful translation as many experts in voice-over claim for it? This paradox is identified by Kaufmann (2004: 157) when reflecting on the use of voice-over and subtitles as credible and faithful translation modes for documentaries:

> Mais s'ils [voice-over and subtitling] sauvegardent leur crédibilité auprès du public cible, sont-ils pour autant respectueux de l'intention des personnages et de l'auteur autant qu'il est convenu de le supposer? Peut-on affirmer que la réalité n'a pas été retouchée? Prendrait-on autant de libertés avec une oeuvre d'imagination qu'on en prend avec le réel? Considérait-on comme inconvenants les incorrections, le mélange

des langues, les accents malhabiles s'il s'agissait d'un texte littéraire, de dialogues rédigés par un écrivain?

Certain challenges have been identified concerning English-Spanish documentary translation by Espasa (2002) and Matamala (2009a and 2009b). Matamala (2009b: 117-118) explains that after conducting a survey with professional translators of documentaries, one of the main challenges in this mode of translation has proven to be keeping the target text within the limits imposed by the original. According to this survey some translators consider that length constraints should be taken into account in the process since there is no other professional who does it, while some others admit not doing it, since they do not get paid for it, or the dubbing studio or TV controllers have asked them to do so. In Matamala's view (2009: 118), an adapter was traditionally in charge of this, but she believes that this is not the case any longer, so translators are the perfect candidates to perform this task.

In addition, Espasa (2002) suggests one challenge connected to the issue of register and style in terms of oral delivery. For her, the narrator shows a rather formal register and diction, which is normally rendered correspondingly. But, talking heads can have very different registers, such as the formality characterising a specialist or the spontaneity of a child; in these cases, talking heads imply a more "oral" text than that of narrators, leading to the need for reformulation in the translation process (2002: 38 and 2004: 191).

As a consequence of the different registers present in documentaires, TV channels or audiovisual production houses may apply their own translation guidelines, as is the case with *ARTE* (as stated by Kaufmann) or *TV3* (Chaume 2001: 79-80), which recommends different registers for narrators and talking heads in terms of syntax, lexicon and pronunciation. In this regard, Pinto (2001: 297) reckons that each company, translation services provider or organization should establish their own in-house models and methods of planning and managing translations, considering their level of general and technological resources, according to the ISO norm 9000[53]. According to

---

[53] Link to ISO norm 9000: https://www.iso.org/iso-9001-quality-management.html

Pinto, the final result of any translating process, including documentaries, is a *written* text that reflects (2001: 296):

> a) [t]he tactic, rhetorical and functional competence of the translator in the process of re-elaborating documentary information,
> b) the capacity of sequential and logical text execution on the part of the translator, and to some degree,
> c) the effectiveness of the final expression of content.

While Pinto considers that the result of any translation is a written text, even in the case of documentaries, we consider that the result of the translation process in voice-over, is not a written text but an oral one. The translation process ends with the oral delivery of that written text by the voice actors in the studio; a delivery in which the oral features characteristic of the original text have been adapted and standardised in the process. It will be interesting to see how the illocutionary functions and prosodic markers are rendered in the documentaries included in our corpus.

3.1.3.3 **Reality shows**

One TV programme category that has caught the eye of the audience over the last decades is reality TV, or reality shows, which is usually rendered into a target language by means of voice-over. Despite its popularity, reality TV has not been thoroughly investigated (Baños 2019: 266), just like voice-over. According to Lundy et al (2008: 208), reality TV started to emerge as a specific genre in the late 1980s, and its main characteristic is that it is unscripted, compared to traditional scripted TV programmes. Reality shows can be defined as unscripted programmes featuring real people, not actors or athletes, active in a specific environment (Mead 2005: 03), although nowadays this type of programme can even be starred by animals (Horak: 2006: 470). In addition, Bagley (2001: 01) considers that:

> As a presentation of non-actors in legitimately natural settings and situations working without a script, reality TV stakes its claim with viewers to regard its depictions as unadorned and spontaneous truthful documentation of natural reality.

Some of the characteristics of reality TV are presented by Nabi et al. (2003) in Lundy et al. (2008: 209). For these authors, reality-based television excludes news programmes, talk and interview shows, and nonfiction narrative programmes, and its various characteristics include (2008: 209):

1- Characters are real people (not actors)
2- programs are not filmed on a set, but in natural living or working environments
3- programs are not scripted
4- events are unplanned, but evolve from narrative contexts, and
5- the primary purpose is viewer entertainment

The reason behind choosing voice-over as the preferred revoicing method for this kind of genre is not very clear. Probably, from the viewers' perspective these programmes are more in the line of documentaries, which also show spontaneous language and describe "the reality", than with cinema movies, which are dubbed. Another option could be the cost of production, and as we already know, voice-over is significantly cheaper than dubbing. These types of TV programmes, are also subtitled; this translation mode is available nowadays on most TVs thanks to DTT (Digital Terrestrial Television) and smart TVs. It is interesting to note that reality TV is defined as truthful, just like voice-over. It may be the case that the "most truthful" audiovisual translation method has been chosen to translate the "most truthful" TV genre for translation's sake, but we consider that the economic reason is a more powerful one in this case.

There is not much bibliography concerning this TV genre, and how it is to be translated (if by means of voice-over, dubbing, or subtitling). Nonetheless, in our corpus of study there are some reality TV shows translated using voice-over, and we will try to find out whether there are any differences between voice-over translation for this category and for documentaries, for example.

3.1.3.4 **Web videos**

Web videos are normally non-broadcast videos shown on company websites, or on the company channel on social networks, such as Youtube, Vevo, etc., about the activities

carried out by the company itself, related to products and services offered, news, etc. They are normally e-learning productions (training videos) and can be either in the form of an animation or motion graphics, or a film with real people. They typically show how a company, a machine, service or system works. Some of these web videos can include subtitles, but they can also be dubbed or voiced-over. Roturier (2015: 169) acknowledges the importance of synchronization in translating these types of videos when he states that "it may not always be possible to have a perfect user experience through localization" since the visual content in these videos may be referred to by the voice-over, while captions, or textual elements on screen, are sometimes not translated. In his view, the only way to guarantee that both the visual content and the voice-over are in the same language and in sync is by creating a video in a new language from scratch. This is the case for languages whose word order differs from the source language (2015: 171).

There are multiple web video categories as stated by Yang et al. (2007), but in our study we will focus on the translation of voiced-over e-learning videos.

3.1.4 The role of the adapter

The act of adapting and the role of the adapter have been the topic of hot debate for a long time, and they are still today. In translating for production and post-production, the first transcription (and translation) will, in some cases, be modified afterwards either by a journalist (Matamala 2008: 118) in the case of TV news interviews, or by an adapter in the case of documentaries and other genres like talent shows or entertaining programmes. Some scholars, like Mayoral and Chaume, prefer translators to carry out the adapting tasks by themselves as they have a better picture of the translation as a whole, which can help them avoid taking erroneous linguistic decisions (Barzdevics 2012: 59). This scholar considers that the role of the adapter, or adjuster, has to be an independent one because he/she has a better understanding of the voice actors that will take part in the production process and the possible problems that may come up in the dubbing studio. This process of reducing and modifying the translation by the adapter can be seen as a double-edged sword, since on one side relevant pragmatic information

might be deleted, and, on the other, relevant content information may also be erased or wrongly modified.

Let us assume that the translator transfers intonational aspects correctly from English into Spanish by means of syntax, and afterwards these are modified, or even deleted, by the adapter. This will probably result in the pragmatic meaning of the oral script being removed out of ignorance. Now, let us assume that the translator does not pay attention to the prosodic elements, rich in pragmatic value, of the original oral script. Would the adapter be able to include these in the script? Probably not. Del Río (2001: 226-7) discusses the figure of adapters and their relevance in quality translations for dubbing referring to Piastra (1989), who had already claimed that one of the reasons leading to bad quality translations was the lack of collaboration between translators and adapters. Del Río considers that once film scripts have been translated and returned to the dubbing studio, the adjuster is in charge of putting the text in the mouth of the actors without asking the translator nor discussing the possible mistakes that may happen in the adjusting process, which may have been inexistent in the translation phase. As a consequence of all this, the audience will be the only affected party, and will probably find parts in the film that they cannot understand.

One last change in the translated script, in addition to the translator's and the adapter's, is the one made by the voice talent at the studio. When the recording takes place outside Spain, it is very common for the voice talent to be the only speaker of the translated language. For Pageon (2007: 70-78),

> I have had on countless occasions to rewrite, edit and shorten the script while the client was getting hot under the collar (…) In most recording sessions, the voice -meaning voice talent- is the only native speaker of the language; on rare occasions there might be someone who might understand the language a little bit.

From Pageon's words one cannot but guess the different modifications that scripts suffer prior to their final recording. The fact that scripts can be altered by translators, adapters, and voice talents themselves, could imply that the fidelity to the original text is pure coincidence, not to mention the rendition of the prosodic and intonational features.

The case illustrated above is that of voice-over recordings in London, where Pageon lives and works, and the author of this work testifies to Pageon's words from his years as a voice talent based in the UK, and also working as an online voice talent for foreign (not Spanish) companies. It will be interesting to know what the professional voice talents interviewed say about this in the conclusions chapter, in which the results of the questionnaires on professional voice recording will be shown.

Adapting/modifying scripts can also result in "technical and content infidelities" (Darwish & Orero 2014: 03), in voiced-over TV news broadcasts, thus questioning the role of the translator and the channel in which the news was broadcast. In addition, Darwish & Orero reinforce this idea when they state (2014: 11):

> [T]echnical restrictions force news broadcasters to sacrifice accuracy and precision. In fact, some of the presumably translation-induced errors are essentially infidelities largely caused by technical modulations, where the script translation is crafted to suit audiovisual production requirements.

Translating for production or for post-production just differ in the fact of having, or not, a written script to work with from the beginning. However, translations are done with or without scripts, and the fact that intonational and prosodic features are obviated in the translation process can be observed both in voice-over for production and post-production, with and without a script, and for every genre (TV interviews, documentaries, etc.). In our study, the way the prosodic features are treated in each genre and translation type (production and post) will be analysed in order to search for a pattern that could explain this situation.

3.2 Intonation in voice-over translation

3.2.1 Translation techniques in voice-over

This section aims to portray the procedures used by translators in order to transfer the values of the prosodic components of the original oral message in English into Spanish

in voice-over translation. There are several methods, techniques, and norms connected with audiovisual translation related to countries, TV channels[54], written press media groups[55], and the academic community, but before delving into these them, it must be explained that voice-over can be understood as an "in between" translation method. For Orero (2005), voice-over for production (when no script is available) implies a mixture between interpreting and standard translation:

> Si bien es cierto que el material es estático (el traductor puede consultar obras de referencia o volver a escuchar la cinta), y por tanto la tarea es la propia del traductor, el texto es oral y no existe una transcripción, lo que hace de la traducción de voice-over una actividad muy próxima a la interpretación (2005: 217).

In a later study, Orero (2006a) characterized voice-over as a halfway technique between dubbing and subtitling when she was describing the synchronization features of voice-over, although she was probably referring to voice-over for post-production, and from a viewer's point of view. However, this is not specified in the article, or maybe she was referring to voice-over in general in its two variables. In her view:

> It may be said that [voice-over] is a mode in between subtitling and dubbing. On the one hand, there is subtitling, where the original message is always present and can be heard along the translation in the form of subtitles. Dubbing, on the other hand, erases all traces of the original message, though body language and in some cases lip movement can give away what is being originally said. Voice-over will be in between the two previous audiovisual translation modes because it allows for some original message to be heard, but it also crushes it when the voice narrates the translation on top of the original voice: two voices can be heard simultaneously (2006a: 256).

As regards translation techniques, in her analysis of voiced-over TV interviews (2005: 220), Orero cites Lambert and Delabastita (1996), for whom this translation modality (in general) would follow one of the five transformations suggested in their work: *repetition* (reproduction without formal modification)*, adiectio* (reproduction with addition)*, detractio* (reduction), *transmutatio* (change in order)*, and *substitutio*

---

[54] TV3 Channel norms: http://www.ccma.cat/llibredestil/manual-dus
[55] El Mundo: http://www.masmenos.es/wp-content/uploads/2002/01/librodeestilo_elmundo.pdf

(replacement). In the present work, we will use the audiovisual translation techniques proposed by Martí Ferriol (2013) to study how translators deal with voice-over in the corpus analysed, since Ferriol's proposal offers a wider range of techniques, having being developed from the contributions of Delabastita (1990), Chaves (2000), Chaume (2005a), and Díaz Cintas (2003). This taxonomy has been created for the modalities of subtitling and dubbing, and although voice-over is not mentioned in Ferriol's study, his proposal could be used for this audiovisual mode considering it has a mixture of those two, since, as stated by Orero (2006a), it features a presence of the original audio as in subtitling, and a translated voice which has to coincide (more or less) with the time limits of the original audio track as in the case of dubbing.

The taxonomy presented by Martí Ferriol (2013), shown in the figure below, is made up of twenty techniques numbered from one to twenty, for which the lowest numbers belong to techniques that favour a more literal translation and will be linked to subtitling, and the higher numbers, as one moves along the continuum, show techniques related to a more interpretative-communicative translation.

CLASIFICACIÓN DE TÉCNICAS DE TRADUCCIÓN SEGÚN EL MÉTODO DE TRADUCCIÓN

Método literal

Método interpretativo-comunicativo

Préstamo
Calco
Tr. P x P
Tr. 1 x 1
Tr. Literal
Eq. Acuñado
Omisión
Reducción
Compresión
Particularización
Generalización
Transposición
Descripción
Ampliación
Amplificación
Modulación
Variación
Substitución
Adaptación
Creación discursiva

1.	Préstamo
2.	Calco
3.	Traducción palabra por palabra
4.	Traducción uno por uno
5.	Traducción literal
6.	Equivalente Acuñado
7.	Omisión
8.	Reducción
9.	Compresión
10.	Particularización
11.	Generalización
12.	Transposición
13.	Descripción
14.	Ampliación
15.	Amplificación
16.	Modulación
17.	Variación
18.	Substitución
19.	Adaptación
20.	Creación discursiva

Figure 14. Ferriol's taxonomy of translation techniques (2013: 122).

Of these twenty techniques, only those that affect voice-over translation, from our perspective, will be explained below (each number in the continuum has been respected, and the name the techniques has been translated into English) (2013: 120-121):

a) Calque: translating a word or foreign phrase literally; it can be lexical and/or structural.

b) Word by word translation: the translated text keeps the grammar, word order and primary meaning of all the words in the original text. This can be the case of the first words we hear in voice-over translation, which commonly match the meaning and order of the original exactly.

c) Literal translation: the translated version represents the meaning of the original very closely, but the number of words does not coincide and/or the order of the phrase has been altered.

d) Set equivalent: the use of a term or expression, which is known as an equivalent in the target language.

e) Omission: the complete suppression in the target text of some information elements present in the source text.

f) Reduction: suppressing in the target text some part(s) of the informative load, or an informative element, present in the source text. This technique and the previous one are voluntarily adopted by the translator facing a problem or a specific restriction.

g) Compression: synthetizing linguistic elements.

h) Particularization: using a more precise or specific term in the target text.

i) Generalisation: using a more general term, or a neutral one, like a hypernym.

j) Transposition: changing the grammatical category, or the voice of the verb, from active to passive, and vice-versa.

k) Amplification: adding linguistic elements that fulfill the phatic function of the language, or elements which are not relevant from an informative point of view, such as adjectives that qualify something which is obvious as it is shown on the screen.

l) Modulation: changing the point of view without altering the meaning in order not to generate a sense of awkwardness in the receiver of the target text.

m) Variation: changing linguistic or paralinguistic elements (e.g. intonation, gestures) that affect the features of the linguistic variety, such as the tone, style, social dialect, geographical dialect, etc.

n) Substitution (linguistic and paralinguistic): replacing linguistic elements with paralinguistic (intonation, gestures) or vice-versa.

o) Adaptation: replacing a cultural element characteristic of the source language with and culture another one from the target culture and language.

Of these fifteen techniques that may apply to voice-over translation, only two are directly related to prosody and intonation: variation, and substitution (linguistic, paralinguistic). Nonetheless, the description of these two techniques does not really specify how the prosodic changes are actually made. As examples of these techniques, Ferriol (2006) identifies variation in the dubbing of mispronunciations from English into Spanish in the film *Lost in Translation*, such as *"Rogere" Moore* for *"Loger" Moore* (2006: 481), and *Uh, "brack" toe* for *Uh, dedo "loto"*, and the oral substitution of *Wow (guau)* for *Vaya* (2006: 452) and *Oh jeez* for *Lo que faltaba* (2006: 407). These examples of variation and substitution for dubbing do not show any change or variation in the prosodic level, since the first corresponds to a segmental variation in the pronunciation of two words ("Loger" and "loto") to match the Japanese pronunciation effect of the originals ("Rogere" and "Brack"), and the second are just linguistic substitutions. This shows that, from the strategies that can be applied to intonation, in this case, there is no rendition of a prosodic element for a linguistic one, which can be a sign of how elusive prosodic traits can be.

3.2.2 Intonation delivery

The way the intonational and prosodic features characteristic of the English language can be rendered in the Spanish translation of voiced-over texts is varied. In chapter 2 we delved into the intonational features of English and their pragmatic relevance in communication, and we saw, for example, how the function of tones can be slightly different in Spanish in terms of showing speakers' attitude, as noted by Solé (1989: 193). Nonetheless, it seems that voice actors are not aware of this intonational difference between English and Spanish, since, as stated by

Zabalbeascoa, Izard & Santamaria (2001: 106): "(d)ubbing actors do not tend to be creative but rather imitate the dominant speech and intonation patterns, which makes each category even more paradigmatic [AR1]".

Although the quote above is addressed to dubbing actors, it is also valid to voice actors in general, since many dubbing actors participate in voice-over productions. In addition, the fact that voice actors listen to the original voices while they record their lines (both for dubbing and voice-over) could sometimes make them imitate the original audio track and convey a different attitude or emotion from that shown by the original actor, which could, in turn, lead to misinterpretations for the audience watching the film. This inaccurate intonational transfer could also be present in voice-over translation, as it is in dubbing. In our view, if tone, for example, which is the most obvious and recogniseable intonation feature, and showing the greatest similarity between English and Spanish, seems to be problematic in its rendition in the target text, it is not difficult to imagine that other prosodic elements that are not so obvious (such as the deaccenting of anaphoric elements, and the placement of the intonational nucleus) will probably be more problematic or not considered at all in English-Spanish voice-over. The results of Sánchez Mompeán's research in her PhD thesis were evident in the case of English-Spanish dubbing (2017: 419-420):

1) Dubbing intonation shows vast differences with spontaneous intonation and features a great number of unnatural patterns that might reduce the quality of the final outcome.
2) Connotations transmitted intonationally by the original characters in English are often lost in translation, thus leading the target viewers to make a wrong interpretation or depriving them of a considerable amount of implicational content.

In the light of these results for English-Spanish dubbing, it is probably safe to assume that this will also be the case for voice-over. Moreover, as Matamala (2019: 70) points out, voiced-over documentaries in Spain are delivered with a non-emphatic pronunciation, unlike reality shows, which are "more and more emphatic". This emphatic polarity between voice-over subgenres could result in some relevant intonational features not being properly rendered in the case of documentaries, and an exaggerated emphasis or inadequate rendition of the prosodic-illocutionary force in the target text for reality shows.

The next chapter will show a detailed description of the programmes that make up our corpus of study, the methodology used to analyse them, the tools used, and the objectives we aim to fulfil at the end of this work.

4 CORPUS, METHODOLOGY AND OBJECTIVES

This chapter is devoted to the description of the type of corpus and the methodology used for the analysis, which will take part in chapter 5. First a justification for the texts and programmes chosen to build the corpus will be presented, followed by a description of the types of programmes, the documentation available in the annexe section, and the methodological approach and analysis tools to carry out the study of the texts. Finally, the hypothesis and objectives aimed to fulfill will be introduced.

4.1 Corpus justification

A corpus can be defined as "a collection of pieces of language that are selected according to explicit linguistic criteria in order to be used as a sample of language" (Sinclair 1996)[56]. In our case, the corpus is formed by voiced-over programmes which have been broadcast on TV and/or the Internet. Since we are dealing with a corpus for translation analysis, its nature will be bilingual. Bilingual corpora can be of two types according to Rabadán & Fernández Nistal (2002: 51): parallel and comparable. A parallel corpus is formed by a group of source texts in a source language and their translations in a target language. On the other hand, a comparable corpus is formed by texts which have been created originally in two or more languages and that, despite the fact that they are not translations of one another, they have certain similarities. Therefore, our corpus will be bilingual and parallel, as we have the original programmes in English (ST) and their translation by means of voice-over into Spanish (TT).

The perfect representation in a corpus is impossible (Rabadán & Fernández Nistal 2002: 59) since it would require an infinite number of texts; so, what a researcher is expected to do is to work with a corpus large enough to achieve the desired results and which includes a varied set of

---

[56] Sinclair, 1996, EAGLES. Preliminary Recommendations on Corpus Typology. http://www.ilc.cnr.it/EAGLES96/corpustyp/node5.html

samples. With this in mind, and considering that voice-over is a translation modality which can be found in different types or categories of audiovisual programmes, I have included a few programmes per category so there is a general representation. Thus, my corpus includes three documentaries, three short interviews and news programmes, three reality shows, and two internet videos.

Collecting the programmes for the corpus was not an easy task. The first option was to request them directly from some of the TV channels and production houses that produce and broadcast these types of programmes, but all of them refused to collaborate allegedly due to copyright issues and in-house policy issues. As can be seen in the annexe section, a series of emails asking for collaboration and help for academic research were sent to the following companies: AMC Networks Iberia, FOX Spain, The History Channel Iberia, RTVE. None of them, not even the public TV company in Spain, agreed to collaborate with this research. RTVE simply directed me to the website of "Docufilia", one of their programmes, in which some documentaries are available in Spanish, for me to see if there was anything of interest. Concerning the translation of the documentaries, they explained that the documentaries they broadcast are translated in-house by some RTVE translators but, due to copyright issues they could not disclose any further information. The History Channel Iberia company also refused to collaborate due to copyright reasons and gave me the names of the studios in charge of the Spanish recording of the programmes they broadcast, SDI Media and Studio 21.

Because of these difficulties in building a corpus for study, I had to resort to the place where one can access almost a copy of every programme: the Internet, and more specifically, Youtube.

4.2   Description of the programmes analysed

This description will be divided into four sections, depending on the type of voiced-over programme category: documentaries, news channel interviews, reality shows, and web videos.

4.2.1 **Documentaries**

The Last Days of Anne Boleyn / Los últimos días de Ana Bolena

The Last Days of Anne Boleyn is a documentary directed by Rob Coldstream, produced by Oxford Film and Television, and distributed by the BBC in 2013. It has a duration of 59 minutes and describes, in a mixture of historical dramatization and interviews of renowned scholars and historians, the Queen's last days in 1536. The translation of this documentary (at least of the script into Spanish) was carried out by the RTVE in-house translation department, while the studio in which it was recorded is unknown (but it was probably recorded within the RTVE premises). It was broadcast in 2014 and could be accessed online in the "Docufilia" (now called "Documenta2") website until January 2015. Although the video was supposed to be available online in the RTVE website, it was retrieved and downloaded for the present study from a Youtube channel[57] which no longer exists.

Image 16. Screenshot of the BBC website displaying "*The Last Days of Anne Boleyn*" documentary.

---

[57] The channel was Doctv Doctves, and it was terminated by Youtube because of third-party claims of copyright infrigment.

Tasteology / Saborología

Tasteology, *Saborología* in Spanish, is a documentary about the food industry produced by AEG[58] in 2016 and it is divided into four episodes: *Source (Origen), Chill (Frío), Heat (Calor),* and *Experience (Experiencia).* It was not broadcast on any Spanish TV channel and is only available online on the AEG Youtube channel[59] both in the original version and the Spanish voiced-over translation. From the four episodes available, only two were analysed here: *Chill (Frío)* and *Experience (Experiencia)* since they included more people being interviewed than the other two and offered more material to analyse. The duration of the episodes is the following: 13 minutes in the case of *Chill* and 19 minutes in that of *Experience*.

Jaimie's Comfort Food / La Comida Reconfortante de Jamie

Jamie's Comfort Food is a TV show produced in England for Channel 4 and aired for the first time in 2014 in the UK. In the programme we can see chef Jamie Oliver preparing three "comfort food" dishes, including starters, mains and desserts. Each episode lasts around half an hour (24 minutes approximately) in the original version. In Spain, the translated version is called *La Comida Reconfortante de Jamie*, and it has been broadcast on two channels: RTVE and Canal Cocina.

---

[58] AEG is an electrical equipment company famous worldwide for its home appliances.
[59] AEG Youtube channel: https://www.youtube.com/user/AEG

Image 17. Screenshot of TVE's Jamie's Comfort Food episode list.

The Spanish version of this reality tv show is quite different from the original in two respects: content and length. While in the original English version there are three recipes per episode, the Spanish version shows four, and instead of having a duration of approximately thirty minutes, they last for almost one hour. This means that, apart from enlarging the content of the programme with a new recipe in the translated version, as well as in the version to export to foreign channels outside the UK, a lot of footage which had been cut in the original English programmes has now been included. This led to a mismatch in terms of time code reference between source and target texts, as shown in the spreadsheet used for my analysis. Due to this mismatch in timing and content, only the parts that match have been transcribed and analysed in this dissertation. Two columns have been included in the spreadsheet to the right of the Spanish translation in order to show the "in/out" time codes for the translated lines, so that the reader can find these parts easily in the target video.

4.2.2 News TV interviews

The texts within this category have been produced and broadcast by the news channel *Euronews*, and they belong to the programmes *Le Mag, Reporter,* and *I-Talk.*

Le Mag

Le Mag is a magazine programme which covers news about different topics, such as sports, music, art, etc. The clip under study selected here covers an interview with Bruce Dickinson, lead singer of the heavy metal band Iron Maiden; and was broadcast in September 2015. It has a duration of 02:33 minutes and the original English video includes a narration of a journalist together with the interviewee's testimony, which is the part voiced-over into Spanish in the target version.

Reporter

Reporter (now called *Insiders*) was a programme devoted to covering relevant topics and it included the narration of a journalist and several interviews with people talking about the specific topic of debate. The video clip analysed in our corpus is titled *What's it like for EU migrants in Britain amid UKIP's success?* (in the Spanish version *El "Turismo social", ¿un problema en Gran Bretaña?*) and lasts for 08:21 minutes.

I-Talk

I-Talk was a *Euronews* programme about interviews with some of the world's most iconic political figures, discussing current hot topics such as the institution of the monarchy in European countries, the possibility of the UK leaving the EU, LGTBI rights, etc. The programme analysed here is: *Is it time to scrap the CAP? (in the Spanish versión ¿Debería el Reino Unido quedarse en la Unión Europea?)*, and it has a duration of 09:48 minutes.

Image 18. A frame of a Euronews I Talk TV show.

4.2.3 **Reality shows**

The three programmes selected for this category are quite different but all share the requirement which is common to all reality shows: they are unscripted. They are the following:

a) Cupcake Wars - Season 6 Episode 13 (Guerra de Cupcakes – Temporada 6 Episodio 13).

b) Ramsay's Kitchen Nightmare's – Season 1 Episode 1 (Pesadilla en la Cocina UK – Temporada 1 Episodio 1 "Bonaparte's").

c) Teen Mom OG - Season 6 Episode 6 (online trailers 1 and 2).

Cupcake Wars / Guerra de Cupcakes

Cupcake Wars is a TV show produced in the USA by Super Delicious Productions and Wut It Is, and distributed and broadcast in that country by Food Network channel. It first aired in 2009 and saw its last episode in 2016, after 11 seasons and 150 episodes. In this TV show, four bakers and their teams compete to make cupcakes, the winner getting a prize of $10,000. The winner will also have the chance to display their products in a major event in the USA. The duration of each original episode is one hour (50 minutes approximately). The Spanish version of this TV show was broadcast on Divinity, a channel belonging to the Mediaset España media group. The duration of the translated episodes for the Spanish market differs from that of the original American ones in approximately ten minutes, part of the content having been edited and cut out in the target version. This is shown in the spreadsheet used for analysis in the present study, which includes two columns for each time code (original/target), so the reader can access the lines analysed easily.

Ramsey's Kitchen Nightmares / Pesadilla en la Cocina

Ramsey's Kitchen Nightmares is a TV show produced by Granada Television and Optomen Television, which was broadcast by Channel 4 in the UK from 2004 to 2009 and included 36 episodes (28 original and 8 revisits). The programme shows chef Gordon Ramsey visiting failing restaurants in an effort to help them improve and succeed. In Spain the TV show was broadcast by channel Nova, which belongs to the Atresmedia group. The episodes, both in the original and translated versions, last for 50 minutes approximately. In this case, the episodes distributed in the Spanish market show no mismatch in content or timing with the original ones.

Teen Mom OG

Teen Mom OG is a reality show produced for MTV and distributed by Viacom Media Networks, which is a media company that owns TV channels such as MTV, Nickelodeon or VH1, among others. The show portraits the life of four 16-year-olds and the motherhood struggle of USA TV celebrities Farrah Abraham, Maci Bookout, Catelyn Lowell and Amber Portwood. The episodes have a duration of approximately one hour. The TV show was broadcast in Spain by MTV and the two clips included for the analysis here are the trailers of one of the episodes from season 6. The full episodes (both original and translated) were impossible to retrieve so I had to get the trailers using a screen capture video recorder on the computer. The duration of the trailers differs, as can be seen in the spreadsheet for analysis, which shows time codes both for the original English and the Spanish translated version.

4.2.4 **Web videos**

The web videos selected for analysis in this work belong to Columbia Threadneedle, and have been retrieved from this company's Youtube channel and website. The videos show different content; video 1 (*TN Europe Jan 17*) is a report on Europe's economic situation as per January 2017, presented by one of the investment analysts of the firm, and video 2 (*Investment Philosophy*) is a talking heads[60] video on the company's philosophy of investment.

TN Europe Jan 17 is rather interesting since it shows how script translation can affect the production of the video for the target locale. As can be seen both on the videos and the spreadsheet, there is a difference of one minute and forty seconds between the two versions, since the video was enlarged in the target version to fit the Spanish audio recording. As can be observed, the video starts and ends with voice-over; at the beginning we can hear the presenter's original voice but just after a couple of lines, the original voice has been removed completely and we only hear the translated Spanish voice. The same thing happens at the very end of the video when we get to hear the original voice for a few seconds. This could have been done to give some realism at the beginning and at the end, but it is puzzling to see this kind of voiced-over dubbed with no lip-sync

---

[60] According to the online Collins dictionary of English: *"talking heads" are people who appear in television discussion programmes and interviews to give their opinions about a topic*.

video production. This shows that some companies prefer to adapt the video to a long translated text rather than adjusting a long translation to the original video. Adapting and editing videos to make them fit into TL translated scripts can be quite expensive since hardly any translated script will have the same length as the corresponding source text.

Investment Philosophy is a standard voice-over translation. Both videos (original and translation) have the same length (10 minutes), which means that, on this occasion, the translation of the script was adjusted to the time codes of the original video.

4.3 Documentation

The cloud-based folder and the CDs attached to this PhD dissertation include all the documents that have been used for this research, namely:

- A folder with all the video files, both in English and Spanish, that have been analysed: *The Last Days of Anne Boleyn; Tasteology: Chill; Tasteology: Experience; Jamie's Comfort Food* episode 1*; Euronews Le Mag* (Iron Maiden interview); Euronews *I Talk: Is it time to scrap the CAP?;* Euronews *Reporter: What's it like for EU migrants in Britain amid UKIP's success?; Cupcake Wars* Season 6 Episode 13*; Teen Mom OG* trailers for Season 6 Episode 6*;* Web video *Investment Philosophy;* Web Video *TN Europe Jan 17*.
- A folder with two spreadsheets: one including all video transcripts (one sheet per video) and analysis, with comments and notes; and one with all the quantitative analysis.
- A folder with all the audio files and pitch contour images for the examples shown in the analysis section, together with the harmonic spectra of voice quality analysis from the character synchrony section.
- A folder with the questionnaires sent to translators, voice actors and casting directors in order to assess how they value and consider the role of prosody in voice-over translation as well as audiovisual translation in general.
- Emails sent and received from TV channels.

- A video interview with Javier Fernández-Peña, a Spanish voice actor whose voice was chosen for Spanish Buzz in Oscar-winning Toy Story 3.
- UNE norms for subtitling and audio-description.

4.4 Methodology of the study

4.4.1 Research Model

After having introduced the corpus of study, I will explain the way in which this material will be analysed: first, describing and justifying the methodological approach, and secondly, presenting the tools for the analysis, so that the reader can use them too if s/he wishes to check the analysis and get an insight in order to apply these tools in future research.

The model of analysis used in this work is based on the one proposed by Rabadán (1992: 208), which, in turn, derives from De Beaugrande & Dressler's model (1981). These scholars define and understand a *text* as a communicative unit that satisfies seven criteria of textuality: cohesion, coherence, intentionality, acceptability, informativity, situation, and intertextuality. These criteria constitute the communicative principles of texts, and the failure to satisfy one or more of them will compromise communication, obtaining, as a result, what Bertucelli (1996: 258) considers a "no-text".

In her methodological frame for translemic analysis, Rabadan's (1992: 208) model uses and applies four of the seven parameters included in De Beaugrande & Dressler's: *intentionality*, which refers to the attitude of the speaker, that is, the goal the speaker tries to achieve; *acceptability*, dealing with the relevance the text can have for the target audience; *situation*, which refers to the relevance of the text within a particular communicative situation; and finally, *intertextuality*, concerning the possible dependence of the ST on previous texts, and which requires the audience's awareness of these in order to understand the text. For Bertucelli (1996: 258), this last parameter makes it possible for the target audience/reader to identify texts as interviews, or sports chronicles, or essays, etc.

The Figure below shows how these parameters are set within this methodological framework:

SOURCE POLYSYSTEM	TARGET POLYSYSTEM
LINGUISTIC SOURCE SYSTEM	LINGUISTIC TARGET SYSTEM

TRANSLATOR

ST   TT

INTENTIONALITY - - - - - - - - - - - - - - - ACCEPTABILITY

COHESION

INTERTEXTUALITY

SOURCE SOCIOCULTURAL CONDITIONS	SITUATIONALITY	TARGET SOCIOCULTURAL CONDITIONS
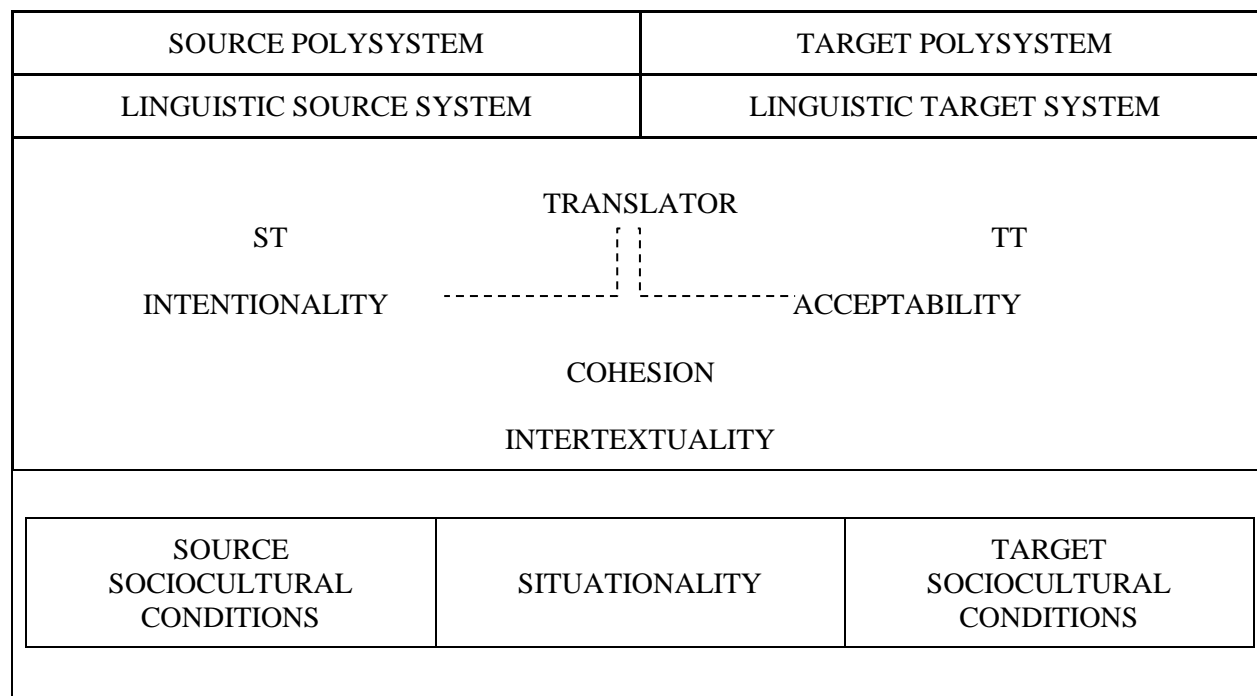
Figure 15. Rabadan's (1992: 208) framework for translation analysis method (my translation)

Thus, for each text analysed in the present dissertation (both in the source and target language) a brief description of these four parameters (intentionality, acceptability, intertextuality and situation) will be included, followed by an analysis and comparison of both texts.

Rabadán proposes a method of analysis including four steps that are taken from Toury's method (1980):

1. ST Analysis.

2. TT Analysis.

3. Comparison of both texts (setting the translation units and checking the translemic relationships between the two poles).

4. Evaluation (assessment of the data analysed in step 3).

The first two phases will allow us to get an idea of the "changes" that take place in the target text. The third phase will show us the functional equivalence between them and the sort of translational relationship between the units of the two poles. Next, these equivalent translation units will be evaluated. One has to remember, as noted in chapter 1.3.2, that we are dealing here with equivalence in audiovisual translation, which means that the visual and aural non-verbal elements of both target and source texts will have to be considered, since they play a key part in any form of audiovisual product.

In order to develop phases 1, 2, and 3 in our analysis, an Excel sheet has been created with all the possible intonational and oral variables of both texts (source and target) worth considering (see Chapter 2). The aim is to get both quantitative and qualitative data using the spreadsheet and specific software to analyse intonation. A quantitative analysis is necessary to measure the way intonation is translated in terms of tangible data, which will show us, for example, the percentage of ellipsis and reaccenting used in the target text for deaccenting cases in the source text. The qualitative analysis will consider whether there is a loss or change of meaning in the target text for the translation of the intonational features under analysis; it will be used for a series of questionnaires aimed at audiovisual translators, casting directors, and voice actors.

4.4.2 Resources for analysis

4.4.2.1 Excel spreadsheet

The above-mentioned spreadsheet is rather large, so the reader might not be able to see it very well in the picture below. I will therefore explain the variables under analysis which are shown in it and how to interpret them.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1																						
2																						
3											Subject vs predicate division											
4							yes	yes	Deaccenting	Intonational subordination	Definining / Non-definining relative clause	Narrrow focus	Fall		Yes	Yes	Yes	Yes	Yes	Yes	Yes	
5								no	Reaccenting	Quotes	Topics	Broad focus	Rise		No	No	No	No	No	No	No	
6											Cleft sentences		Fall rise		n/a	n/a	n/a	n/a	n/a			
7											Tag Questions		Rise fall									
8							ITEMS OF STUDY															
9							Discourse Markers				3 TS				SYNC				Rendition in Spanish voice-over			
10		In	Out	Character	English Original	Spanish Translation	Engish Source Text	Spanish Target Text	STRESS	PITCH-RELATED	TONALITY	TONICITY	TONE	SPANISH TRANSLATION STRATEGY	ISOCHRONY	LITERAL	CONTENT	KINETIC	CHARACTER	Rendition of prosodic illocutionary force	Transfer of semantic meaning	COMMENTS
11																						
12																						

Figure 16 Screenshot of the spreadsheet used to analyse the corpus

The items of study are highlighted in yellow on the spreadsheet and include a dropdown menu to select the variables that may apply to each item. The distribution of the spreadsheet is as follows:

- Starting from the left, we see that columns 2 and 3 show the "in" and "out" timings of the speaker's words on screen. This is useful for isochrony, and also to show the reader where to go in the video file to analyse other data.

- The next column, number 4, is for characters, and shows the names of the people speaking on screen, or off screen, but whose voices are being translated through voice-over. Most of the times narrators are translated using commentary, or off-screen dubbing, and there is no voice-over for these, so there is no analysis of their parts.

- Columns 5 and 6 show the English source text and the Spanish translated text respectively.

- Some oral features are analysed in columns 7 and 8, which deal with discourse markers. What column 7 shows is that there are discourse markers in the source text with a *yes*, and column 8 shows whether these have been kept in the translation (*yes*) or not (*no*).

- Column 9 deals with the possible features of stress in the source text, and there are two possibilities: deaccenting, and reaccenting.

- Column 10 sets two pitch-related phenomena: intonational subordination and quotes.

- Columns 11,12, and 13 deal with the 3 Ts. Thus, column 11 analyses tonality and all the features that are relevant in translation from our perspective, which are: subject-predicate division, defining/non-defining relative clauses, topics, cleft sentences, and tag questions. Column 12 deals with cases of narrow focus and broad focus, and column 13 deals with tone (fall, rise, fall-rise, rise-fall). All these columns (7 to 13) and their respective variables are analysed in column 14 "Spanish translation strategy", which shows how the target text has rendered the features mentioned in the different columns.

   The next five columns concern the synchronisation features stated in chapter 1.3, for which the possibilities are *yes*, *no*, and *n/a* for those examples in which sync is right, wrong, or it does not apply.

The next two columns, numbers 20 and 21, belong to the category "Rendition in Spanish voice-over". These two columns analyse whether there is a full rendition of the source text, a partial one, or no rendition at all. Column 20 is labelled "Rendition of the prosodic illocutionary force", which aims at reflecting whether the illocutionary force of the original text plus the prosodic traits that are characteristic of the delivery and that also carry pragmatic load have been rendered in the target text. The possibilities here are two: *yes* or *no*. Column 21 is labelled "Transfer of semantic meaning" and its purpose is to reflect whether the basic message of the original text has been transferred[61] to the target text. The options here are again two: *yes* and *no*.

- Finally, column 22 shows some comments that together with those of column 14 (Spanish translation strategy) give the reader our view regarding the suitability of the translation for each line under analysis.

All the headers for all the items of study (columns 7 to 21) have filtering possibilities so the readers can decide how to display the results depending on what they want to review. This is also quite practical to get quantitative results at the end of the analysis. We understand this spreadsheet might not be perfect or have the best possible layout, but it is the most useful and manageable document that I have been able to devise for this task.

4.4.2.2 **Speech analysis software**

Together with the spreadsheet, a speech analysis software has been used to analyse the intonation patterns of both the source and target texts so we can provide graphic evidence of certain elements which are relevant for our study. The software is Praat[62] and it allows us, among other things, to show the audio wave, spectrogram, and pitch contour from a given utterance, as well as to write the words of the utterances and place

---

[61] Although the term *transfer* is a common one in Translation Studies and is understood as the stage in which the analysed material is transported from the source language into the target one in the translator's mind by Nida and Taber (1969: 33), and also as the "dynamic process of `reconfiguration´ in the TL of sets of SL semantic and structural components" by Hatim & Munday (2004: 46), our use of *transfer* here relates to the generic transfer of meaning from the source text to the target text.

[62] Praat is a free speech analysis software developed by Paul Boersma and David Weenink at the University of Amsterdam (1995). The current version is 6.0.46 (2019).

them accordingly on the horizontal (time) axis shown on screen, so that we can export the whole "picture" of the utterance in different formats, such as pdf, png, and eps. The examples below show how this works.
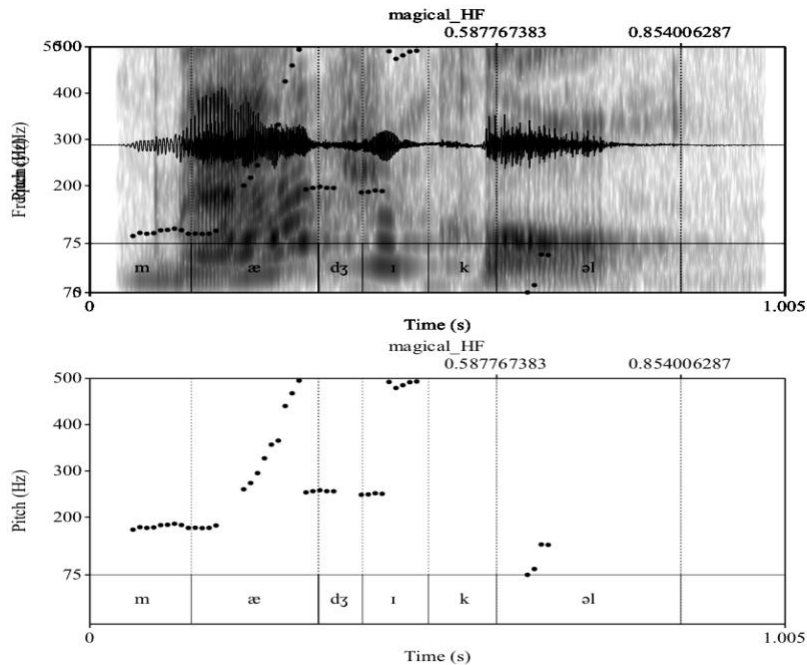
Figure 17. Spectrogram and pitch contour

The figure above shows two graphic representations for the pronunciation of the word "magical". The picture on the top shows a black and white spectrogram with the audio wave and intonation contour superimposed with the transcription of the word at the bottom, in which each phoneme stands in its time frame during the utterance. The dark parts of the spectrogram show phonation (vibration of vocal cords), so the darker the part the more phonation there is, as in the case of vowels, and nasals. The whiter parts show, on the contrary, the voiceless parts (no phonation), as in the consonant phoneme /k/. The spectrogram, audio wave, and intonation contour are shown on two axes: the horizontal one shows the time frame in seconds, and the vertical one shows the frequency, i.e. the pitch, in Hertz. The picture at the bottom shows the utterance of the same word ("magical") but without the spectrogram and the audio wave. This time there is only the pitch contour and the phonemes at the bottom. Here we can clearly see that the intonation is a high fall nuclear tone, in which the tune goes from as high as 500Hz in /æ/ to nearly 100Hz in /əl/. This is the typical intonation used in exclamatory utterances, as is the case in the example. There is no pitch contour for /k/ because it is a voiceless phoneme and there is no vibration of the vocal folds. Note that the dots just

above /dʒ/ and /k/ are background noise and do not represent intonation.

Nevertheless, the use of spectrograms is really useful when analysing segments, i.e. phonemes, and their distribution along the axes. Its use in the example was motivated by the nature of the utterance (one word) and the fact that the pitch contour could seem unclear with the /k/ in the last syllable of the word. The examples shown on the figure below, on the other hand, do not show the spectrogram and will probably be easier to understand, since they only display the pitch contour and the audio wave.
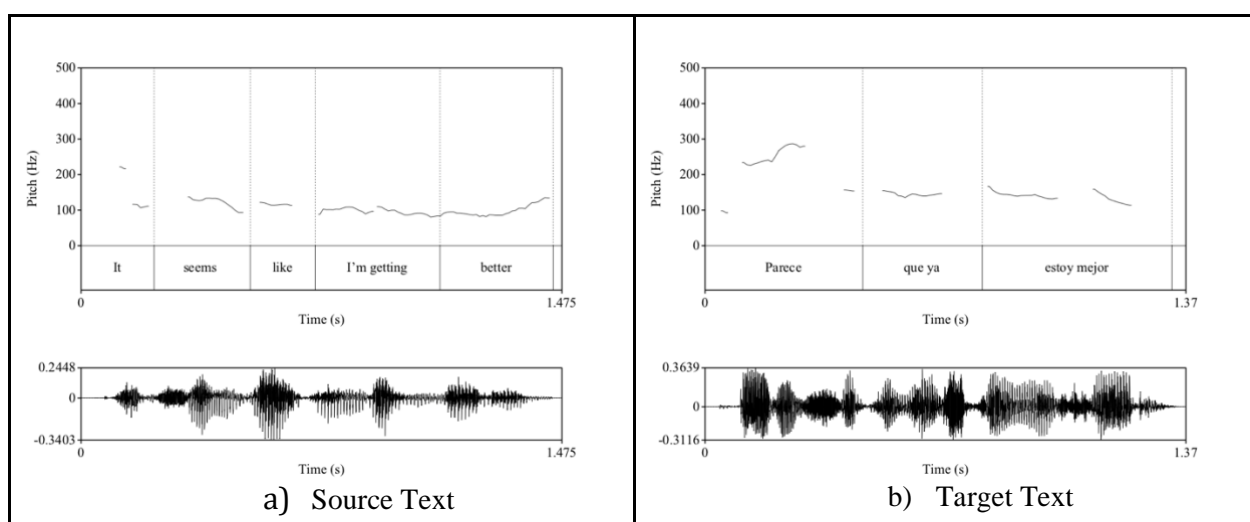
Example 71. Pitch contour and audio wave comparison from Euronews *Le Mag*

These two examples belong to the Euronews *Le Mag* interview with Bruce Dickinson[63] and show the original and translation of the sentence "It seems like I'm getting better". We can see (and hear) that in example 70a the sentence is divided as follows:

| it   ˈseems like I'm getting   ˈbe   tter |

Pre head   **Head**   **Nucleus**   **Tail**

Example 70a shows a pitch contour within a range of 100Hz - 200Hz, which matches the fundamental frequency typical of male adults, and shows pitch variation

[63] See Euronews Iron Maiden Interview on the corpus (spreadsheet and mp4 file).

on the word "better", where the nucleus is located, showing broad focus or neutral tonicity (i.e. all the information is new), with a rising tone, which indicates that more information is to come.

If we now take the translation for that line, as shown in example 71b, we can see (and hear too) that there is a significant change in terms of nucleus placement and frequency.

| pa   ˈre   ce que ya estoy mejor |

Pre head   Nucleus   tail

Although the structure of example 71b could look similar to that of 71a (since it only lacks a head), the pitch contour, together with the IP division, shows us that the nucleus of the IP is in narrow focus (*parece*) since it is not the last lexical item of the IP, but the very first one. In addition, the frequency of the voice that delivers the translation is set in a range between 300 Hz and 150 Hz, and as we can hear it is the voice of a female adult, thus hindering character synchronization.

These two examples (71a and 71b) show how the translation of tonicity can sometimes be complicated and deliver a different type of message. As can be seen, the nucleus of 71a is "better", while the nucleus in 71b is "parece" (it seems), which is not what the original speaker intended to convey in the first place. Here, despite the fact that the overall message is translated and that the basic semantic meaning is not compromised, the specific pragmatic meaning and emphasis is altered and as a consequence the rendition of the original message is only partial.

There is a second software we have used in order to analyse the voice quality of the characters who show no character sync in our corpus: Logic Pro X[64]. This is a music recording software provided by Apple, which has multiple equalizing options. One of these is *Match EQ[65]*, a plug-in that allows to analyse the frequency spectrum of an audio

[64] Logic Pro X website: https://www.apple.com/logic-pro/
[65] An overview of the Match EQ plug-in can be found here:
https://support.apple.com/kb/PH27199?locale=en_US&viewlocale=en_US

file. By activating this plug in we can see the fundamental frequency and the vocal harmonics displayed on a graph. These harmonics are, in fact, the ones responsible for the characteristic timbre of a voice or a musical instrument.

The table below shows two harmonic spectra. The horizontal axis (in yellow) shows the frequency (Hz) and he vertical axis (in green), the amplitude of the audio wave measured in decibels (dB). The first peak of the graph stands for the fundamental frequency, and the subsequent peaks are the harmonics, which give colour to the voice. These are multiples of the F0 (fundamental frequency), so for example, assuming that F0 is 120 Hz, the second harmonic will happen at 240 Hz, and the third at 360 Hz, and so on. Normally, harmonics above the fifth one are inaudible to the human ear. It is worth mentioning at this point that in order to compare voices, we will analyse similar vowel sounds retrieved from the two texts, and in the graphs the lowest three major peaks are the primary correlates of vowel quality (Rose 2002: 206); they are what make a vowel sound like an /e/ and not like a /u/.

The harmonic spectrum

| **Source text** | **Target text** |

Figure 18. Harmonic spectra of two voices out of sync in our corpus

As we can see, the lines depicting the two voices are quite different. The range of lower frequencies, below 100 Hz, is higher in the target text voice, which means that it is a relatively deep resonant voice, compared to the source text voice, which barely

has no range below 100 Hz. This, in addition to the harmonic differences, implies that the two voices are very different.

4.4.2.3 **Questionnaires**

A series of questionnaires have been prepared so that voice-over professionals, including actors, translators and casting directors, answer them to provide a professional insight to this dissertation. The aim is to get a complete picture of how voice-over and prosody are treated from an AVT industry perspective. Thus, we have contacted over 40 voice-over companies obtaining 25 answers from casting directors; more than 20 professional voice actors, of whom 14 answered our survey; and over 10 audiovisual translators, getting only 4 answers. We understand that the number of translators who participated in the survey is not significant, but given the fact that participating in our study is voluntary, and that some companies do not want to share their provider's contacts for confidentiality and business-related matters, there was nothing we could do to increase the number of answers. Nonetheless, we believe that the number of voice-over actors and casting agents who took part in the study is relevant, and they will provide a clear picture of how the industry conceives prosodic traits in the voice-over translation process.

The questionnaires for voice actors and translators are in Spanish since most of them are Spanish or from a Spanish-speaking country, which means that the information provided by these professionals will be presented in Spanish in the conclusions section. On the other hand, the questionnaire devised for casting agents is in English, since most of the companies asked to participate in our survey are international and based outside Spain. Therefore, the information provided by casting directors will be presented in English in section 6.1.

The questionnaires include multiple choice answers for the sake of clarity in the analysis, although some open-ended questions are also provided so that these professionals can express themselves freely when asked about prosody, voice quality and industry-related questions.

CUESTIONARIO PARA LOCUTORES				
#	PREGUNTA	RESPUESTA		
1	¿Qué tipo de entonación usa para grabaciones de voice-over (voces superpuestas)?	Sigo la entonación original	Uso una entonación neutra	La que me diga el cliente
2	Para locuciones de personajes con acentos extranjeros…	Uso acento neutro español	Uso acentos regionales españoles	
3	¿Ha hecho voces de personajes que no tienen su edad y rango vocal (adolescentes, mayores...)?	Sí pero cambio el tono para imitar el original	No, solo dentro de mi rango de voz y edad.	
4	¿Ha puesto voz alguna vez a un personaje del sexo opuesto (no valen voces de niño hechas por mujeres)?	Sí	No	
5	¿Alguna vez se ha cambiado el texto a locutar en la sesión de grabación por problemas de sincronización?	Sí, es frecuente.	No, nunca me ha pasado.	Raras veces.
6	¿Cree que su voz se elige en función del personaje al que va a dar voz?	Sí	No	A veces
7	¿En las grabaciones para voice-over (voces superpuestas) hay un supervisor lingüístico?	Sí, siempre	Sí, a veces.	No. Nunca.
8	¿Ha grabado algún proyecto de voice-over (voces superpuestas) desde su propio estudio?	Sí, muchos	Sí, algunos.	No, ninguno
9	¿Si tiene estudio propio, qué tipo de proyectos de voice-over suele grabar en su estudio de casa?	Documentales	Talking heads	Reality Shows….
10	¿Qué tipo de proyectos de voice-over (voces superpuestas) suele grabar en estudio profesional externo?	Documentales	Talking heads	Reality shows, etc.
11	¿Cuánto tiempo suele dejar para empezar a locutar después de oir al original?	Nada, lo hace posproduccion	1-2 seg	Depende de si encaja o no el español

Table 18. Questionnaire for voice actors

CUESTIONARIO PARA TRADUCTORES			
#	PREGUNTA	RESPUESTA	
1	Durante su formación (carrera, máster, etc.) ¿estudió aspectos de fonética y fonología inglesa?	Sí	No
2	¿Cómo utiliza los aspectos de fonética y fonología inglesa en la traducción audiovisual al español?	(respuesta abierta)	
3	¿Tiene conocimientos de prosodia/entonación inglesa?	Sí	No
4	¿Qué aspectos de la prosodia inglesa considera más relevantes a la hora de traducir?	(respuesta abierta)	
5	Para traducir acentos regionales:	Los neutralizo	Utilizo otro recurso (ej. acentos/dialectos regionales del español).
6	Si traduzco para voice-over (voces superpuestas)	Me ocupo del ajuste para sincronización	El ajuste lo hace otro profesional en el estudio
7	Una frase como, por ejemplo: "she's the new teacher, isn't she?" (tag question) la traduzco siempre como una pregunta al español.	Sí	No
8	Cuando un personaje introduce una cita en su discurso, (`and she said "are you crazy?"´) suelo:	Parafrasear	Mantener la cita en la traducción
9	¿Sus tarifas para traducir voice-over son más baratas que para doblaje?	Sí	No
10	¿Qué considera más relevante a la hora de traducir para voice-over/voces superpuestas?	(respuesta abierta)	

Table 19. Questionnaire for voice-over translators.

Questionnaire for casting directors & voice agents			
1	In a voice-over project, do target voices have to be similar to source/original voices?	Yes	No
2	What is the reasoning for choosing voices for a voice-over programme? (tone of voice, pitch, availability, performance, cost, etc.)	(open-ended answer)	

3	Is the adjusting done in post-production? Or is the translator who adapts the script to fit the video?	(open-ended answer)	
4	In voice-over (like a documentary or talking heads), how long can we hear the original voice before the translation is given?	(open-ended answer)	
5	Children voices...	Are recorded by children	Are recorded by female voices
		Are recorded by both	
6	Does a voice actor voice more than one character in a voiced-over programme?	Yes	No
		Sometimes	
7	In documentaries, Spanish target voices have to sound.... (like the original voices, neutral, act but not too much, etc.)	(open-ended answer)	
8	Are Spanish voice actors required to act in certain categories/genres of voice-over programmes? Which programmes are these? (for example: documentaries, reality shows, web videos, talking heads, etc.)	(open-ended answer)	
9	Should Spanish voices imitate the original intonation?	(open-ended answer)	
10	10. Why do you think there is a large number of programmes that are translated using voice-over and not dubbing, for example?	(open-ended answer)	

Table 20. Questionnaire for voice-over agents.

4.5 Hypothesis

The main challenge of this work is to get to understand how aware audiovisual translators (and adapters) are in general, and translators for voice-over in particular, of the relevance that all the prosodic features described in chapter 2 have in English communication and therefore in the process of delivering a faithful and reliable translation to the final audience. Our starting point here is that audiovisual translators in general, as well as the professionals involved in the voice-over translation process,

are not normally fully aware of these prosodic traits, and consequently the pragmatic value of the source text is not adequately conveyed in voice-over, resulting in a final audiovisual product which is not as faithful as is claimed by many scholars and professionals.

One of the reasons for this hypothesis is that most undergraduate translation programmes in Spanish universities lack a module or course devoted to the study of the phonetics and phonology of English (as a source language). There are modules in foreign languages, but these do not necessarily focus on or include the phonetic and phonological component of the language. Phonetics and intonation are, in fact, often neglected in EFL teaching, as stated by several scholars, such as Ur (1996: 55), Kelly (2000: 13) and more recently Scrivener (2011: 271).

The evidence for this lack of attention to phonetics and phonology in AVT undergraduate programmes can be found in Cerezo Merchán's PhD dissertation (2012), in which she analysed the didactics of AVT in Spain. Two pieces of evidence concerning our hypothesis can be found there: the first one is that there is no module devoted to phonetics in any of the AVT undergraduate courses she studied; the second is shown in a survey conducted among university lecturers, translators and employers concerning the key competences that translators must have, which those taking part in the survey should arrange in terms of relevance. The two most voted items are shown in the chart below:

Competencias específicas más importantes por grupo de población		
Profesores	**Traductores**	**Empleadores**
Conocimiento exhaustivo de la lengua meta en todas sus dimensiones (ortográfica, fonética, morfológica, sintáctica, léxica, fraseológica y textual) (79,4%)	Conocimiento exhaustivo de la lengua meta en todas sus dimensiones (ortográfica, fonética, morfológica, sintáctica, léxica, fraseológica y textual) (78%)	Conocimiento exhaustivo de la lengua meta en todas sus dimensiones (ortográfica, fonética, morfológica, sintáctica, léxica, fraseológica y textual) (76,5%)
Dominio de la lengua origen (excelente comprensión escrita y oral, reconocimiento de registros y variedades coloquiales, dialectos, etc.) (50%)	Dominio de la lengua origen (excelente comprensión escrita y oral, reconocimiento de registros y variedades coloquiales, dialectos, etc.) (68,3%)	Dominio de la lengua origen (excelente comprensión escrita y oral, reconocimiento de registros y variedades coloquiales, dialectos, etc.) (70,6%)

Image 19. Table of key competences for professors, translators and employers in Cerezo
Merchán (2012: 388)

As the image above shows, the second most voted competence is: source language proficiency (excellent comprehension skills –written and spoken-, command of different registers and colloquial varieties, dialects, etc.), which was ranked second by 50% of the academic lecturers, 68.3% of the translators, and 70.6% of the employers. Given the fact that only 50% of the lecturers surveyed considered that source language proficiency is one of the top priorities, and that the phonetic and phonological component is usually neglected in language courses, it seems that prosody might not considered very relevant.

In section 3.1.4 of this dissertation we could see that adapters, most of the time, do not have a good linguistic background with regard to the source language, which may also be the case for translators as regards the suprasegmental component of English when doing English-Spanish voice-overs. In our opinion, however, in order to produce a good-quality translation, a thorough linguistic knowledge of the source language is paramount. In this sense, Rabadán & Fernández Nistal's (2002: 17), consider linguistics as one of the fundamental pillars of translation, so that in order to understand the translating process, translators cannot ignore the linguistic conceptions and

epistemologies of the language. We will borrow their quote from Mona Baker (1992), even if it is an "old" one, to back up our hypothesis:

> (i)f translation is ever to become a profession in the full sense of the word, translators will need something other than the current mixture of intuition and practice to enable them to reflect on what they do and how they do it. They will need above all to acquire a sound knowledge of the raw material with which they work; to understand what a language is and how it comes to function for their users (1992: 04).

Rabadán and Nistal understand that the phonetic and phonological level of the source text is paramount in certain modes of subordinate translation (2002: 18), as is the case with voice-over. Similarly, Hatim & Munday (2004: 65) observe that certain prosodic features, such as stress, can be used for focal effects (emphasis, etc.) in some languages, but not in others, so these communicative clues may be replaced by means of syntax in the target language. This is exactly what we discussed in section 2.4 of this dissertation.

With this in mind, the fundamental question in our hypothesis is: are translators, and also adapters, fully aware of the relevance of these prosodic qualities (if they are at all)?

As stated before, our starting point will be to assume that the prosodic component of the source language (English in our case) is not fully considered in voice-over translation. In addition, given the fact that the only phonetic element in audiovisual translation which seems to have been considered worthy of relevance so far has been at the segmental level - what is known as phonetic synchrony or lip sync - and that the oral language synchrony as suggested by Mayoral et al. in Martí Ferriol (2006: 124) is too general for our purposes, we will here propose *prosodic-illocutionary force* rendition as a new feature to be considered in the translation of audiovisual products, in order to duly grasp and convey the intonational communicative richness of English when translating from English into Spanish in AVT.

4.6 Objectives

This dissertation has two main goals: the first is to get to know whether English prosody is fully considered in voice-over translation, and the second is to highlight the communicative value of English intonation and the different communicative resources used in English and Spanish in order to improve the quality of the final voiced-over translation.

In order to achieve these main objectives, a series of specific objectives are set:

a) To identify the way prosodic features related to stress, pitch, the 3Ts (tonality, tonicity and tone) are translated into Spanish and whether the original prosodic illocutionary force and the communicative value implicit in the source text have been rendered or lost in the final translation.

b) To identify the way discourse markers are considered in voice-over translation.

c) To check how synchronic features can affect or be affected as a result of the translation process.

d) To check how voice quality is treated in voice-over translation.

e) To check how voice-over professionals (translators, actors, adjusters, directors) conceive prosody and prosodic traits in the process of a voice-over production.

The next section is devoted to the analysis of the corpus selected for this dissertation, in which the most significant utterances which, in our view, are worth of study will be studied and compared. For some of them a graphic illustration will be provided, as in the examples above; in others, we will just discuss how the translation applies to the overall message intended by the original speaker in terms of transfer of semantic meaning and rendition of the prosodic illocutionary force.

5   ANALYSIS AND RESULTS

5.1   Analysis

This part is devoted to describing the most interesting and revealing findings during the analysis process, which are included on the spreadsheet. The structure of this chapter is as follows. First, we will describe the variables of analysis in full as a general overview, that is, including the four categories/genres (documentaries, reality shows, web videos and news interviews) within each section. Then, each variable will be described according to each category/genre.

The order in which the information provided by the analysis is presented in this chapter will be as follows:

a) Intonation:
   - Stress: deaccenting; reaccenting
   - Pitch: intonational subordination; quotes
   - 3 Ts: tonality (subject VS predicate division; defining/non-defining relative clause; topics; cleft sentences; tag questions); tone (fall, rise); tonicity (narrow focus)

b) Synchronisation

c) Voice quality

d) Discourse markers

The analysis presented here shows both quantitative and qualitative information, and aims at quantifying the most common resources translators use in Spanish when dealing with intonational and prosodic elements in the English source text. Thus, we will get an idea of how these features are rendered into Spanish for a functional and suitable translation, and, in addition, raise awareness for those translators that are not fully conscious of the importance that prosody has in audiovisual translation.

The examples within each variable for categories a) and b) above (Intonation and Synchronisation) have been analyzed and categorised taking into consideration two possibilities: the rendition of the prosodic-illocutionary force and the transfer of semantic meaning. These categories are similar but do not mean exactly the same thing. Thus, *transfer of meaning* implies rendering the semantic content of the message into the target language. It can be done, for example, using a literal translation, using different words with a similar meaning to that of the source text, and even reorganising the structure and word order of the sentence in the target version. On the other hand, *prosodic-illocutionary rendition* has to do with the rendition of the illocutionary force (i.e. the purpose or function of the textual unit) and how the message is delivered by the voice actor in the target version.

With this in mind, the possibilities available in the examples analysed are the following:

Source Text	Target Text	Transfer of Semantic Meaning		Rendition of prosodic-illocutionary force		Result
		YES	+	YES	=	Full rendition of original message
		YES	+	NO	=	**Partial rendition of original message**
		NO	+	YES	=	**Partial rendition of original message**
		NO	+	NO	=	No rendition of original message

Table 21: Possible results from a prosodic and communicative perspective.

The options shown in the table above are set in order of preference, from our point of view. Thus, the first option, YES - YES, is the most suitable, since we understand that rendering the original semantic meaning together with an equivalent prosodic-illocutionary force is the most important thing in translation. From the two options in the middle, YES - NO (or Semantic Meaning Only) and NO - YES (or

Prosodic-Illocutionary Only), we consider that the rendition of the intention is more important than not doing it, so Prosodic-Illocutionary Only, NO - YES, is our second-best option, followed by Semantic Meaning Only, YES - NO. In our view, in line with Reyes (1995: 34-35), communication succeeds when the recipient gets the intention of the sender, and not just the literal semantic meaning of what is said. Therefore, we consider that the rendition of the pragmatic meaning is more relevant than that of the semantic one. Finally, the least possible alternative is None (or NO – NO), which means that neither the basic semantic meaning nor the prosodic-illocutionary force have been rendered in the target version.

It is important to bear in mind that some cases might be entwined; for example, we will see examples of narrow focus which have been properly translated by means of a suitable strategy to highlight the focus of information in the target text, but that have been delivered during the recording of the script using an inequivalent tone in Spanish. Or, it could be the other way around: for example, a declarative sentence with narrow focus and rising tone in the source text, translated as a question but without highlighting the same nucleus of information in the target text. In this case the meaning of the tone would be rendered but not the nucleus. We cannot obviate that these variables are possible, and that in order to achieve a full rendition of the original English source text into Spanish we need to consider all the intonational features, both individually and also as a whole.

In addition to this, we will also show and comment on the translation strategies used for each intonational feature. Like this, we will see which are the most recurrent strategies translators use to deal with these intonational aspects, which of them work, and which do not and why.

We have included a minimum of 5 examples for each item of study. There are items which show more than 5, such as *deaccenting* and *reaccenting* with 10 examples each. This is due to the fact that we have found numerous cases of these features in our corpus and, in our view, it seems pertinent to include s few more examples so as to have the most varied analysis possible. Within each example analysed the item of study (*deaccenting, reaccenting, subject-predicate division*, etc.) will be underlined, and the nucleus will be highlighted in bold.

With this in mind, let us show the results of our analysis starting with the intonational features of stress, pitch, tonality, tonicity and tone.

5.2 Results

5.2.1 Stress

Within stress, the two variables analysed are deaccenting and reaccenting. Initially, we intended to search for accenting features as well, but in the end, we considered that would not be as relevant here as the other two variables, so it was ruled out. In any case, in the narrow focus section, we will find relevant accenting examples.

5.2.1.1 Deaccenting

The pie chart below shows the total number of examples with deaccenting found in our corpus and the percentage for each variable: Full rendition (grey); No rendition (yellow); Semantic Meaning only (orange); and Prosodic-Illocutionary only (blue).

Total cases of DEACCENTING (36)

- 3%
- 3%
- 25%
- 69%

Legend:
- Prosodic-Illocutionary only
- Semantic Meaning only
- Full rendition
- No rendition
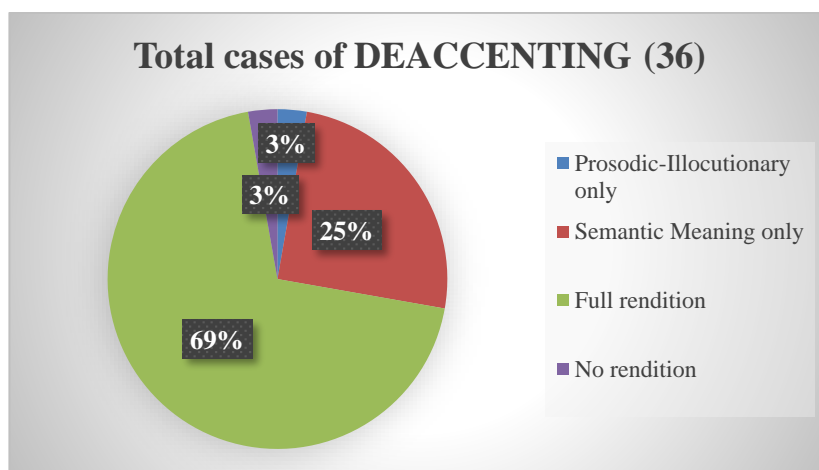
Figure 19. Total number of deaccenting examples and percentage of variables.

We can see that for the 36 cases of deaccenting, 69% were fully rendered, 28% were partially rendered (most of them are cases of semantic meaning only) and 3% were not rendered at all.

193

Moving now to the strategies used to render this stress-related feature in the target text, the bar graph below shows the different options.
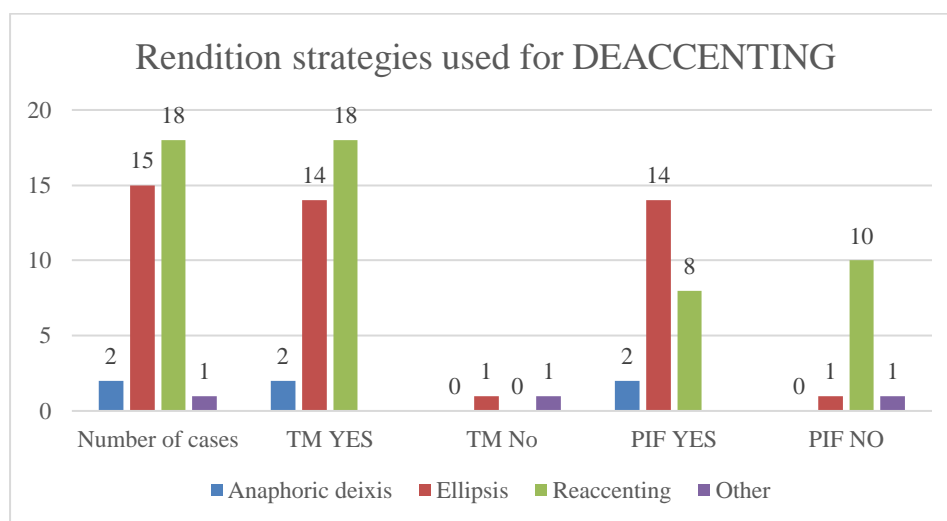
Figure 20. Strategies used to render deaccenting in the Spanish target version

As can be seen in the graph, the different strategies are: anaphoric deixis (2 cases), ellipsis of deaccented word (15 cases), reaccenting of deaccented word (18 cases) and other (1 case), which refers to an example that has been reformulated so much in the target version that we consider it does not render the original message or its prosodic value.

We observe that, from the strategies identified, reaccenting and ellipsis are the most recurrent ones with 18 and 15 cases respectively. From the 18 cases of reaccenting, all of them render the basic semantic meaning of the original message, but only 8 do so in terms of prosodic-illocutionary meaning. On the other hand, from the 15 cases of ellipsis, the semantic meaning of the original is rendered in 14 cases, and also 14 carry the prosodic-illocutionary value. Finally, there are 2 cases in which the deaccented material in the source text has been rendered in the target version by means of anaphoric deixis, which fully renders the original message.

Let us now see which are the most recurrent strategies per category.

DOCUMENTARIES			
Category	**Number of Cases**	**Strategy**	**number of cases**
Deaccenting	25	Anaphoric Deixis	2
		Ellipsis	13
		Reaccenting	10
		Other	0

INTERVIEWS VIDEOS			
Category	**Number of Cases**	**Strategy**	**number of cases**
Deaccenting	3	Anaphoric Deixis	
		Ellipsis	1
		Reaccenting	1
		Other	1

REALITY SHOWS			
Category	**Number of Cases**	**Strategy**	**number of cases**
Deaccenting	8	Anaphoric Deixis	
		Ellipsis	1
		Reaccenting	7
		Other	

WEB VIDEOS			
Category	**Number of Cases**	**Strategy**	**number of cases**
Deaccenting	0	Anaphoric Deixis	0
		Ellipsis	0
		Reaccenting	0

		Other	0

Table 22. Number of deaccenting examples and translation strategy into Spanish per genre

As can be seen from the table above, out of the three genres/categories that show examples of deaccenting, the commonest strategy is reaccenting for reality shows and interviews videos, with documentaries showing more cases of ellipsis than reaccenting. The least preferred strategy is anaphoric deixis with only two examples. These strategies coincide with what was explained in section 2.5.2 about the different resources Spanish has in order to render such a prosodic feature as is the deaccentuation of old information. We can now recall Caldiz (2012: 01) and García Lecumberri (1995: 197), who stated that given information in English is deaccented but in Spanish it is, in fact, reaccented. We can also observe, from the data above, that anaphoric deixis, is not very much considered although all its examples show a full rendition of the original message. If we consider Mateo's (2014: 120) suggestions in order to render successfully the deaccenting of old information in English into Spanish, deixis is a great strategy as it allows the translator to use pronouns, for example, to refer to the deaccented material using fewer words, and syllables, which can be beneficial for synchronisation purposes.

Let us now see some of the examples that make part of the figures mentioned in the graphs and tables above, starting with those that achieve a full rendition of the original message, then partial, and finally, no rendition at all.

a) **Full rendition.**

D.Ex. 1[66] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:12:37	00:12:39	Hilary Mantel	Henry is beginning, it seems,	Es porque está empezando a

[66] D.EX. stands for Deaccenting Example.

00:12:39	00:12:43	Hilary Mantel	to think about annulling his marriage to <u>Anne</u>.	pensar en anular su matrimonio con <u>Ana</u>.
00:12:43	00:12:47	Hilary Mantel	He cannot imagine what he ever **saw** in <u>Anne Boleyn,</u>	No entiende **qué vio** en <u>ella,</u>

This example, taken from the documentary *The last days of Anne Boleyn* shows an evident case of deaccenting in English translated by means of anaphoric deixis into Spanish. We can see that "Anne Boleyn", which is repeated information as it comes from the previous line "Anne" is deaccented by means of the anaphora rule. In the Spanish translation, we see that the first "Anne", which is new information, is given as "Ana", and the second time it is given as "ella". In addition, the nucleus in the source text falls on "saw" as a consequence of the deaccenting "Anne Boleyn" and this is reflected in the target text with a subtle emphatic intonation on "qué vio" since the use of exclamative "qué" in Spanish usually has prosodic inflexion to differentiate it from the atonic relative pronoun "que" (Alarcos, 1994: 135).

D.Ex. 2 *The Last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:46:07	00:46:10	Greg Walker	And if you happen to be an egotistical monster,	Y si resulta que eres un monstruo ególatra
00:46:10	00:46:14	Greg Walker	as Henry VIII was, you want to <u>act decisively</u>.	como era el caso de Enrique VIII <u>actúas con contundencia.</u>
00:46:14	00:46:16	Greg Walker	And he **does** <u>act decisively.</u>	**Eso** fue lo que hizo.

This example, taken from the same documentary as the previous one, shows, again, how anaphoric deixis ("eso") in the target text serves to render the deaccented old information ("act decisively") given the in the source one. As can be seen, the nucleus in the source text falls on the emphatic particle "does" (default accent), and in

the Spanish text it falls on the deictic particle "eso" which reinforces the original speaker's intention (to act decisively). Here the translator has used a copulative relative construction[67], also known as *estructura ecuacional* (Gutiérrez Ordóñez 2017: 37), which serves to highlight the relevant information in Spanish as cleft sentences do in English.
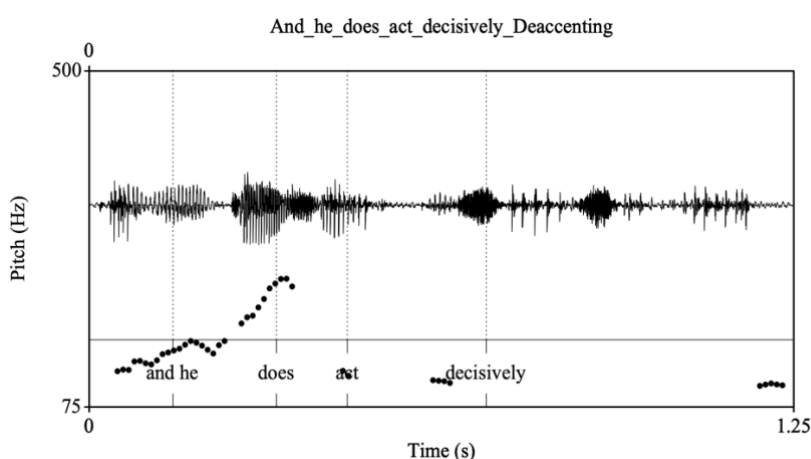
Image 20. Pitch contour and sound wave of D.Ex.2

The image above shows the pitch contour of the source text for this passage. We can see how the pitch increases on "does" and falls on "act decisively" as a result of deaccenting. This example shows how intonational variables are quite often combined in fluent spontaneous speech (narrow focus and deaccenting) and serves to remind us that all intonational features are relevant for a successful translation both individually and as a whole. The example, however, has one downturn, which is literal synchronisation at the end since "decisively" and "hizo" do not match. As we saw in section 1.3, literal synchronization serves to endow the target text with credibility and realism. Nonetheless, in this scene, Greg Walker continues talking and this mismatch is almost unnoticeable for the audience.

D.Ex. 3 *Tasteology: Experience*				
In	Out	Character	Source text	Target text

---

[67] This construction includes the verb "ser" plus a relative clause (que, quien, etc) which highlights one of the components of the sentences.

198

00:05:23	00:05:46	Charles Spence	I'm a gastrophysicist, which means I'm interested in systematically studying any associations, **surprising** associations between the senses.	Soy gastrofísico, es decir, mi interés radica en estudiar sistemáticamente las **asociaciones sorprendentes entre los sentidos.**

This example taken from the documentary *Tasteology: Experience* shows how the deaccenting of old information in English can be rendered in the Spanish target text by means of ellipsis. In the source text we have "associations, surprising associations" and in the target text we have "asociaciones sorprendentes". In the source text the first "associations" is accented, while the second one is deaccented since it is given or old information and, therefore, the accent falls on "surprising". Here the translator decided to omit the second "associations" since it would be a bit redundant; moreover, the strategy works very well because Charles Spence, the speaker on screen, speaks at quite a fast pace, so and the fewer words the target text has, the easier it is to voice for the voice actor. In addition, as we can see, both beginnings and ends in the source and target versions match literally, which means that there is literal synchrony in this case, enhancing audience credibility.

D.Ex. 4 *Tasteology: Chill*				
In	Out	Character	Source text	Target text
00:09:27	00:09:56	Para	A banana that has dots in it is more tasty than a **green** banana, you know.	Un plátano con pintas es más sabroso que un plátano **verde**.

In this example of deaccenting, taken from the documentary *Tasteology: chill,* we can observe how the word "banana" is deaccented in the source text the second time it appears, since it is anaphoric, and the nucleus thus falls on "green", thus resulting in narrow focus (or default accent). In the Spanish target text, the translator decided to reaccent "plátano", a strategy that works well this time because the nucleus falls on "verde", since this adjective is placed at the end of the phrase, so it gets the focus. One

could think that, in this case, the second occurrence of the word "plátano" is redundant, and we could just get rid of it by means of deixis, using "uno verde", but considering there are no timing issues that require any reduction in the word count and that repeating the word serves to fit literal synchrony, this reaccenting of "plátano" makes sense this time.

b) Partial rendition.

b1) Semantic meaning only

In	Out	Character	Source text	Target text
D.Ex. 5 *The last days of Anne Boleyn*				
00:41:12	00:41:15	Greg Walker	In a sense, he would have hoped Smeaton hadn't <u>confessed</u>,	que tenía la esperanza de que Smeaton no <u>confesara</u>
00:41:15	00:41:18	Greg Walker	but Smeaton **did** <u>confess.</u>	pero **confesó,** y después de aquello
00:41:18	00:41:20	Greg Walker	And after that, he had to follow it up.	tenía que seguir adelante con el proceso.

This example shows how the basic semantic meaning of the message can be rendered without fully considering the illocutionary force of the original. We can see how "confess" is deaccented in the source text since it anaphoric and the nucleus falls on the auxiliary verb "did", which works both as a marker of past tense and also as an emphatic particle. This is a clear example of deaccenting with narrow focus, since the nucleus does not fall on the last lexical item, which, as we see, is deaccented. The Spanish version shows reaccenting of the word "confesó"), which is a usual strategy for the cases of deaccenting, as we have seen, but there is no emphatic particle to reinforce the original pragmatic value, as could have been the word "sí", for instance, "pero **sí** confesó" or "pero **sí** lo hizo". In addition to this, there is no pause after "confesó" in the delivery, which would have helped to mark the emphatic intention of

the original source text. What we have is a rather flat delivery which, in our view, does not fully render the original prosodic-illocutionary force.

D.Ex. 6 *Kitchen Nightmares*				
In	Out	Character	Source text	Target text
00:04:55	00:04:57	Gordon Ramsay	Wow, fuck me, that's nice! Who's doing what? Who's doing the fish?	Ah, qué bien. ¿Quién hace eso? ¿Quién hace el pescado?
00:04:58	00:04:59	Tim	**I'll** do the fish.	Yo haré **el pescado.**

The example above shows again how cases of deaccenting accompanied with narrow focus can be tricky to translate if one wants to achieve a full rendition of the original message. As we can see, "do the fish" is the deaccented part in the source text since by means of anaphora rule it is considered old information and, consequently, the nucleus falls on the preceding content word, in this case "I". In the target version, the original deaccented items are reaccented ("haré el pescado" and the nucleus of the IP falls on "pescado" since it is the last word of the utterance and there is no emphatic intonation on "yo" that could signal that the nucleus is not the last word of the sentence. Besides, displacing the nucleus of intonation is not so common in Spanish. If we wanted to mark "yo" as the nucleus of the sentence, we could either mark it with emphatic intonation or place it at the end of the phrase, as in, for instance: "el pescado lo haré yo", or "lo haré yo", or simply, "yo". In these three examples we can see how we can highlight the same nucleus as in the source text reorganising the syntactic structure of the sentence and resorting to reaccenting and ellipsis to render the deaccented material.

D.Ex. 7 *Teen Mom2*				
In	Out	Character	Source text	Target text
00:12	00:14	Female	Yeah. And she went 'cause she wanted to go.	Y… ¿ha ido por voluntad propia?

| 00:15 | 00:52 | Male | She **so** ~~wanted to go~~. | Quería **ir**. |

This example, again, shows that the deaccented anaphoric material in the source text is reaccented in the target text. Nonetheless, in the Spanish version, the nucleus is not highlighted as it is in the English one and, consequently, the illocutionary and pragmatic load is not achieved. In the original text there is narrow focus, with the nucleus falling on "so", an emphatic particle, meaning that she, indeed, wanted to go to the place they were talking about. In the translated text, this emphatic load is inexistent. The nucleus falls on the last lexical item "ir", which coincides with the deaccented part in the source text. In addition, there is no emphatic particle to support the will to go to that place of the person they are talking about. Two alternatives here could have been to use "Claro que quería ir", and "Claro que sí": in the first one we have the emphatic particle "claro" plus the reaccenting of the deaccented part (*ir*) and in the second one, we have two emphatic particles (*claro/sí*) to highlight even more emphasis, and ellipsis of the deaccented words.

D.Ex. 8 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:57:37	00:57:41	Hilary Mantel	I don't think it does any favours to Anne	No creo que a Ana le favorezca en nada
00:57:41	00:57:44	Hilary Mantel	to cast her as a ~~victim~~.	que se la muestre como la ~~víctima~~.
00:57:44	00:57:46	Hilary Mantel	She was **not** ~~a victim~~.	No fue una **~~víctima~~**.

What can be drawn from the example above is similar to the other two previous ones. The deaccented information in the original text is reaccented in the target text and the nucleus falls on the "wrong" word. As we can see, "víctima" is the last word of the sentence and it gets the nuclear accent, while we can observe that in the source text, the nucleus falls on "not" to highlight the new information. Again, the basic semantic meaning of the message is rendered but the illocutionary load is not. Some possible

alternatives for this phrase could be: "no lo fue", "porque no lo era", or adding "de ninguna manera" to the translation provided. By omitting the reaccenting of the deaccented word "víctima" we can use it to our advantage and get the audience to understand that the focus here is the negation of being a victim, as in "she was not a victim whatsoever".

b2) Prosodic-Illocutionary only

| D.Ex. 9 Euronews: is it time to scrap the CAP? |

In	Out	Character	Source text	Target text
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**?

This example, taken from the Euronews programme *I-Talk: is it time to scrap the CAP?*, shows the only example found in the whole corpus in which the prosodic-illocutionary meaning is rendered but the semantic meaning of the original is not, mainly because, in this case, a relevant part of the message has been omitted. As we can see, in the source text, the last line has two instances of the word "whole", being the first one accented and the second one deaccented by means of anaphora rule. The nucleus in this last IP falls on "powerful" marking the contrast between the two "wholes". In the target text there is ellipsis of the second deaccented word, and we get just one "whole", in Spanish, "Unión", which refers to the European Union. The strategy to render the deaccenting by means of ellipsis is suitable, and the placement of "fuerte" at the end of the sentence in order to get the nucleus is perfect. However, the translator leaves the word "larger" that goes with the first accented "whole" untranslated; by doing this, relevant information in the original message contextualizing the deaccented part is excluded and, therefore, the rendition of the original meaning is only partial. By making reference to a larger and stronger whole,

the interviewer in the ST is trying to convince Daniel Hannan, a Brexiteer and Eurosceptic MEP within the Conservative Party, of the benefits of remaining in the EU, which is stronger and also larger than for example the United Kingdom, which is formed by four countries only (England, Scotland, Wales and Northern Ireland). In addition to this, the word "compromise" has been translated as "compromiso" in Spanish, which is a clear case of *false friend*[68] and the meaning is different.

c) No rendition.

D.Ex. 10 Euronews: What's it like for EU migrants in Britain amid UKIP's success?				
In	Out	Character	Source text	Target text
00:04:56	00:05:14	Marcus Reckless	I think many of the people here are making a huge contribution to the country and I wouldn't want to single out any, any nationality. It's a question of overall numbers and at the last elections the Conservatives said they would cut immigration from hundreds and <u>thousands a year</u> to **tens** of <u>thousands a year</u> on a net basis. It's now gone back up to the levels we saw under labour. But we want to be fair and treat people with a commonwealth heritage, <u>people from outside the EU</u>, on the same basis as <u>people from</u> **inside** <u>the EU.</u>	Muchas de las personas que viven aquí están haciendo una enorme contribución al país y no quiero señalar a ninguna nacionalidad en concreto. Hablo de cifras globales. En las últimas elecciones los conservadores dijeron que querían cortar la inmigración en cientos <u>de miles</u> de personas al año, en **decenas de miles**, hablando en cifras netas. Pero queremos ser justos y tratar a todos los inmigrantes de la misma manera porque tradicionalmente siempre hemos sido un país de acogida.

[68] According to the Collins English Dictionary, a false friend is a word or expression in one language that, because it resembles one in another language, is often wrongly taken to have the same meaning. https://www.collinsdictionary.com/dictionary/english/false-friend

This example shows the only evidence found in the whole corpus in which there is no rendition at all of the basic semantic meaning or of the prosodic-illocutionary load. It belongs to the Euronews programme *Reporter: What's it like for EU migrants in Britain amid UKIP's success?,* and, as we can see, there are two instances of deaccenting worth commenting. The first one is the deaccenting of "thousands a year" in the source English text since it is anaphoric and, as such, treated as old information. In the Spanish text we can notice that there is reaccenting of "de miles", which could make sense, but the meaning is not rendered properly since we have "cortar la inmigración en cientos de miles, en decenas de miles" (literally in English: cut immigration in hundreds of thousands, tens of thousands), whereas in the source text we are told "from hundreds and thousands a year to tens of thousands" which is clearly the opposite. A suitable translation in this case would be, for instance: "cortar la inmigración de cientos y miles de personas al año, a decenas de miles".

The second deaccenting case in this example is in the last sentence when the speaker, Marcus Reckless, talks about people from the EU. The table below shows these lines again.

Source English Text	Target Spanish Text
But we want to be fair and treat people with a commonwealth heritage, <u>people from</u> outside <u>the EU</u>, on the same basis as <u>people from</u> **inside** <u>the EU</u>.	Pero queremos ser justos y tratar a todos los inmigrantes de la misma manera porque tradicionalmente siempre hemos sido **un país de acogida.**

As we can see, in the target text there is no reference to migrants both from inside and outside the EU, which in this case is, in our view, quite relevant since the programme is about people from the EU coming to work in Britain and the political vision that the UKIP party has towards immigration and immigrants taking advantage over the social benefits programme offered in the UK. In the Spanish target text, because the topic of the programme and the interviews are about EU migrants, we get the impression that the translation "todos los inmigrantes" refers to EU migrants only. In addition, by substituting the difference between the original "inside /outside the EU"

with "tradicionalmente siempre hemos sido un país de acogida", the focus is set on the the UK as a country rather than on the migrants, whoch is the purpose of the speaker on the source text. Mr Reckless is, in fact, emphasizing that his party (UKIP) wants to treat people not only from inside but also from outside the EU fairly. Thus, we consider that neither the semantic meaning nor the prosodic-illocutionary force are rendered in this case.

5.2.1.1.1 Concluding remarks on deaccenting rendition

As can be observed from the examples above, dealing with deaccented elements can be tricky when in addition to the deaccented material we have a narrow focus possibility. In the Full Rendition section we can see that the different strategies (anaphoric deixis, ellipsis and reaccenting) available to render the deaccenting feature in the Spanish target are satisfactory when they are accompanied by a suitable syntactic reorganization to keep the focus of information at the end of the phrase, or there is some prosodic highlighting, in the delivery of the target text, of the words corresponding to the word(s) carrying the informative and illocutionary load in the source text. The rest of the examples in this section have failed to either reorganize the text from an informative perspective or deliver it in a way in which the prosodic-illocutionary force renders the same pragmatic value as the original English text.

5.2.1.2 Reaccenting

Reaccenting, as we saw in section 2.2.2.2, is another feature of English stress that is used by English speakers, for instance, when they reuse the other speaker's words, as in the case of *echoes*, or when they want to insist on a piece of information already given in order to highlight it again. This English prosodic feature is quite recurrent in audiovisual texts, and, just like deaccenting, the commonest strategies to render it in Spanish are: deixis, ellipsis, reaccenting, and cataphora.

Let us see the overall figures of the deaccenting examples in terms of rendition.
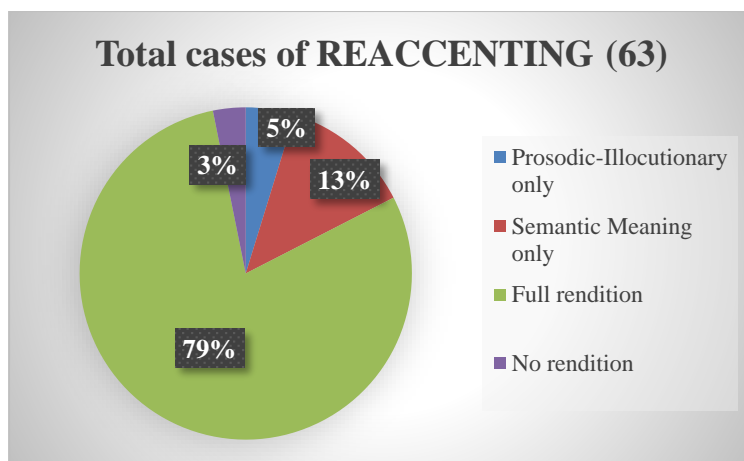
Figure 21. Pie chart showing the total number of examples with reaccenting.

As we can see, 79% of the examples show full rendition, 18% show partial rendition (5% prosodic-illocutionary and 13% semantic meaning), and in just 3% there is no rendition at all. Before delving into the analysis of the examples for each section of the pie, let us check which are the most repeated translation strategies as shown in the bar chart below.
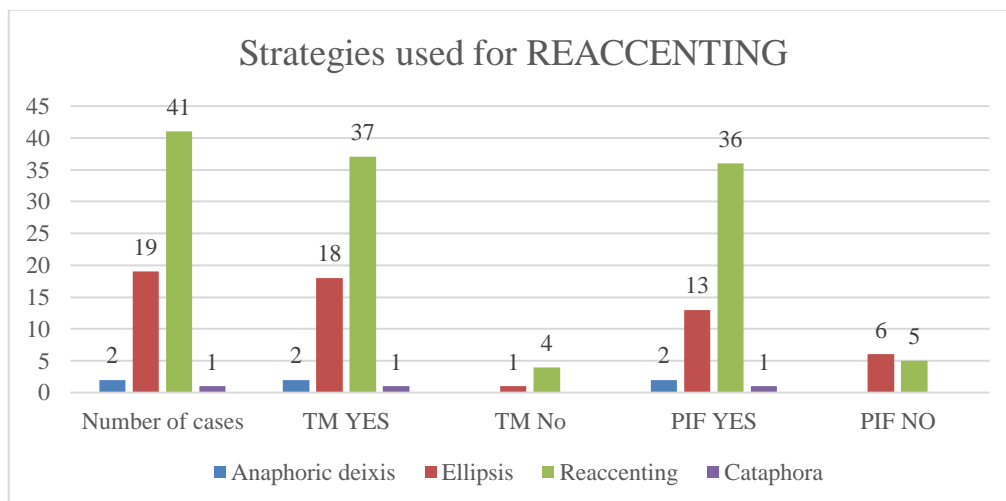
Figure 22. Translation strategies to render reaccented material

As we can observe from the chart above, the strategies used for the 63 cases of reaccenting are: anaphoric deixis with 3 examples, cataphoric deixis with 1 example, ellipsis with 19 examples, and reaccenting with 41 examples, more than double the other most resorted strategy. In addition, almost all the examples with reaccenting in the target text show rendition of both the semantic meaning and the prosodic-

illocutionary force. On the other hand, we see that in the case of ellipsis, from the 19 examples, 18 transfer the semantic meaning, but only 13 do so together with the prosodic-illocutionary force, which gives us 6 cases around ⅓ which do not render the prosodic-illocutionary force in the target text. If we now take the examples of deixis (anaphoric and cataphoric), we clearly see that both examples show full rendition, just like the deixis case in the deaccenting section.

The preferred strategies in terms of audiovisual genre are shown in the table below.

DOCUMENTARIES			
Category	**Number of Cases**	**Strategy**	**number of cases**
Reaccenting	37	Anaphoric Deixis	1
		Ellipsis	12
		Reaccenting	23
		Cataphora	1
INTERVIEWS VIDEOS			
Category	**Number of Cases**	**Strategy**	**number of cases**
Reaccenting	9	Anaphoric Deixis	1
		Ellipsis	3
		Reaccenting	5
		Cataphora	0
REALITY SHOWS			
Category	**Number of Cases**	**strategy**	**number of cases**
Reaccenting	13	Anaphoric Deixis	0
		Ellipsis	4
		Reaccenting	9
		Cataphora	0

WEB VIDEOS			
Category	**Number of Cases**	**strategy**	**number of cases**
Reaccenting	4	Anaphoric Deixis	0
		Ellipsis	0
		Reaccenting	4
		Cataphora	0

Table 23. Number of reaccenting examples and rendition strategies per genre

The information provided by the table above shows that, for each genre, the preferred strategy to render English reaccenting in Spanish is also reaccenting the corresponding word. Ellipsis is the second commonest strategy for each genre, and we have evidence of deixis just in documentaries. There is no sign of this strategy neither in interviews videos, reality tv or web videos. This coincides, more or less, with the data provided in deaccenting, since deixis was not present in reality shows and web videos.

Considering that Spanish is a language that reaccents information, the data obtained in this section is not surprising. Let us now move onto the qualitative analysis and see how these examples show full rendition, partial, or none.

a) **Full rendition.**

We will here cover the different translation strategies found for full rendition of the reaccenting value of the original message. Example 1 below shows reaccenting also in the target text.

R.Ex. 1 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:18:16	00:18:19	Hilary Mantel	But Anne continued to be **Anne.**	Sin embargo Ana era mucha **Ana.**

This example belongs to the documentary *The last days of Anne Boleyn* and is a clear example of reaccenting both in the source and target texts with full rendition. As we can see in the image below the audio wave clearly shows two IPs and the pitch contour reveals that both "Anne" are stressed and accented since there is pitch variation. Both are the nuclei of their IPs and both have a high fall nuclear tone, which is evident in the graphic illustration below. The Spanish version shows a similar structure with almost a word by word translation but with an idiomatic expression in Spanish. Both "Ana" are present and in the exact same position within the sentence as in the original, which works perfectly in terms of literal and content synchrony; and since the nucleus in Spanish falls at the end of the phrase both instances of "Ana" are in focus, just as in the source text.
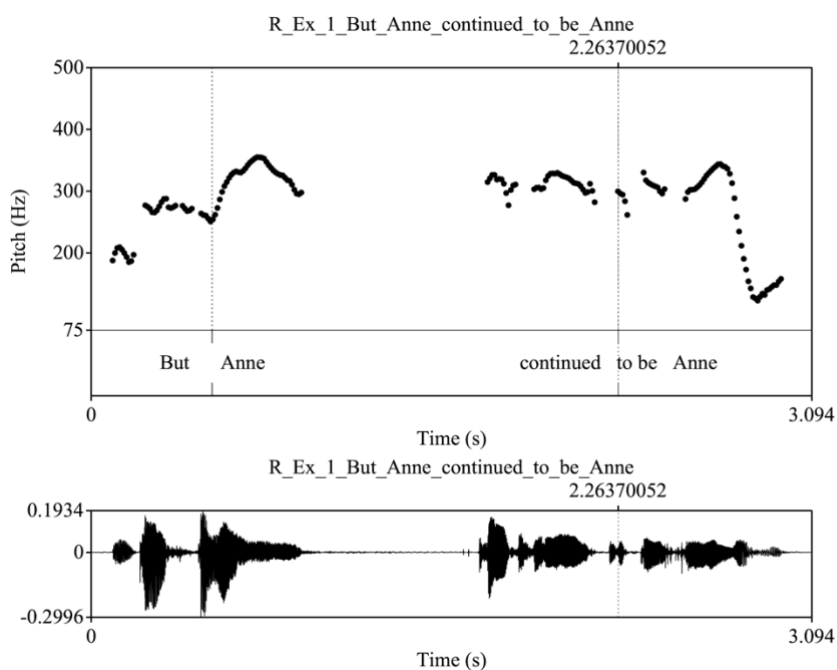
Image 21. Audio wave and pitch contour of R.Ex.1

R.Ex. 2 *Tasteology: Experience*				
In	Out	Character	Source text	Target text

| 00:01:51 | 00:02:14 | Christine Flynn | I developed Jacques in sort of, like, a perfect storm. I mean I'd left my fine dining role behind, but yeah, I missed plating, I missed a lot of the stuff from <u>restaurants</u> that I used to do, because the <u>restaurants</u> I run now are very different. | Creé a Jacques tras lo que podría considerarse como una tormenta perfecta. Ya no trabajo en la alta cocina, pero echaba de menos emplatar y muchas cosas que solía hacer en ese tipo de <u>restaurantes</u>, porque los que dirijo ahora son muy distintos. |

This example shows full rendition of the original message in which the reaccented word in the source text (*restaurant*) has been rendered in the target text by means of ellipsis, which means that the word is inexistent in the target text, and as a consequence, there is no reaccenting in the target text. It has been taken from the documentary *Tasteology: Experience*, and in the source English text we have the word "restaurants" accented twice; the first one because it is the first time she talks about restaurants (the ones in which she worked in the past) and that is new information, and as such, it is accented. And the second time, "restaurants" is reaccented because the speaker is now talking about the restaurants she runs now in the present time, so she treats it as new information, and reaccents it. In the Spanish target text, we can see that "restaurantes" is used just once, and there is no such word after the comma. What we find is the relative pronoun "que" referring back to "restaurantes", its antecedent. In this example, reaccenting the word "restaurantes" could be redundant in Spanish and so the translator has decided to leave it out. "...solía hacer en ese tipo de <u>restaurantes</u>, porque los (restaurantes) que dirijo ahora son muy distintos." We can see that literal and content synchrony work well here because apart from the omitted reaccented word, it looks like a literal, almost word by word, translation.

R.Ex. 3 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:42:55	00:43:00	David Starkey	The final driver of everything under <u>Henry</u>, is **Henry**.	Durante su reinado, <u>quien</u> tomaba las decisiones en última instancia era **Enrique**.

This example, taken from the documentary *The last days of Anne Boleyn*, shows the rendition of reaccenting by means of cataphoric deixis. As we can see, in the source text, the word "Henry" is accented twice. Here, David Starkey is reusing his own words to emphasise that the guilty party in the beheading of his wife was Henry VIII alone. This type of reaccenting and its usage has been explained in section 2.2.2.2 (see also Wells 2006: 178). In the Spanish target text, we can see that there is only one "Henry/Enrique" and that it is the final word therefore receiving the nucleus. However, the second "Henry" can be traced back in the sentence, in fact it is there and is referred to by means of cataphoric deixis in the relative pronoun "quien", the antecedent of "Enrique", its postcedent. This strategy to translate the source text with such an example of reaccenting works well because we do not get the word repeated, which sometimes sounds redundant in Spanish, and it is in focus having been placed at the end of the phrase, as in the source text; because they are in broad focus, both names "Henr/Enrique" carry the nuclei in both texts (source and target). In addition, there is literal synchrony in the final sound bite.

R.Ex. 4 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:37:24	00:37:31	Hilary Mantel	It's a short step from saying, "One day Henry will die,"	Entre decir: "un día el rey morirá"
00:37:31	00:37:34	Hilary Mantel	to saying, "And I hope it's soon."	y decir "y espero que sea pronto" no hay más que un paso.
00:37:34	00:37:38	Hilary Mantel	And it's a short step from saying, "I hope it's soon",	Y entre eso y decir
00:37:38	00:37:40	Hilary Mantel	to saying, "Let's accelerate it."	"vamos a acelerarlo un poco", otro paso.

This example shows two different cases of reaccenting within the same intervention. The extract is taken from the documentary *The last days of Anne Boleyn*, and in it we can clearly see that the words "It's a short step from saying" and "I hope

it's soon" are repeated and also reaccented. The former is reaccented (present) in the target text since we get the word "paso" twice and the verb "decir" three times, while the latter is reaccented by means of anaphoric deixis with the neutral pronoun "eso". This strategy works nicely to render the reaccented word in the source text, because, by omitting it in the target, it does not sound redundant, while the number of words is reduced, which gives more time to the voice actor to deliver the rest of the phrase without rushing the pace. The fewer words we have in the target text, the more time there is available to fit in the video file. We can see that there is no literal synchrony during this scene. The translator has reorganised the structure of the text to highlight "un/otro paso" (it's a short step) so that the audience realise there was a thin line between Anne wishing Henry's death and the possibility of making it happen.

b) Partial rendition

We shall now move on to the partial rendition examples. As we saw in the pie chart above, the partial rendition examples account for 18% of the total, with 16% being Semantic Meaning Only, and just 2% Prosodic-Illocutionary. We will start with the Semantic Meaning Only examples and finish this section with the only case for Prosodic-Illocutionary one.

b1) Semantic meaning only

R.Ex. 5 Euronews. *Reporter: What's it like for EU migrants in Britain amid UKIP's success?*				
In	Out	Character	Source text	Target text
00:03:33	00:03:43	Woman 1	I don't know a lot about how the positives of being in the EU. All I really heard about is, yeah, the <u>negatives</u>, and again, the <u>negatives</u> of **immigration**.	Yo no sé mucho acerca de los puntos positivos de estar en la UE. Solo he oído puntos <u>negativos</u>, **aspectos** <u>negativos</u> de la inmigración.

This first example, obtained from the Euronews programme *Reporter: What's it like for EU migrants in Britain amid UKIP's success?* shows reaccenting in the source

213

text of the word "negatives" and we can see that the equivalent word in Spanish for this reaccented word is also present in the target text ("negativos"). However, the word "negativos" seems to be deaccented the second time in the target text since its preceding word "aspectos" carries an emphatic accent which, in our view, should not be there. We can see that the target text shows an almost word by word translation, which is perfect for synchronic features (literal and content). Moreover, repeating in the target text the word corresponding to the ST's reaccented word makes sense, since this is a strategy that works well to fully render the original message, as we have seen in the previous section. Nonetheless, the delivery of the target text hinders the full rendition and makes it only partial (semantic meaning only) because the reaccented word is somehow deaccented in it, and the focus, which is broad in the source text with the nucleus falling on "immigration", seems to be narrow in the target text with "aspectos" carrying the informative load, rather than "inmigración", as should have been the case.

The images below show the pitch contours and audio waves for the two lines in this example.
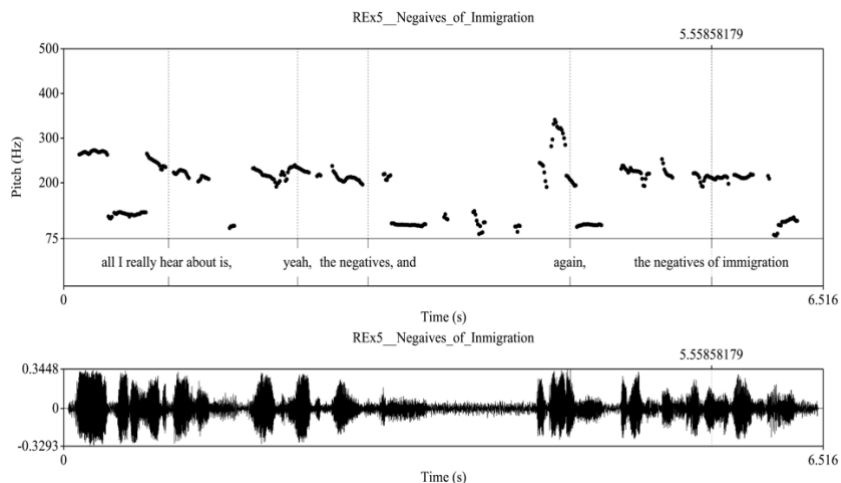
Image 22. Pitch contour and audio wave of the source text in R.Ex.5

In the image above we can see that the first "negatives" is accented; it is the nucleus of the IP | yeah the ˈ**ne**gatives and |, as we observe that the pitch contour rises slightly over the first syllable of "negatives", and then falls all the way until the end of "and", which is the tail of this IP. The second "negatives" is accented and marks the beginning of the head in the IP | theˈnegatives of immiˈ**gra**tion |. The nucleus in this IP,

214

as we said, is the third syllable of "immi**gra**tion", which is clearly visible in the pitch contour above with a high-fall nuclear tone.
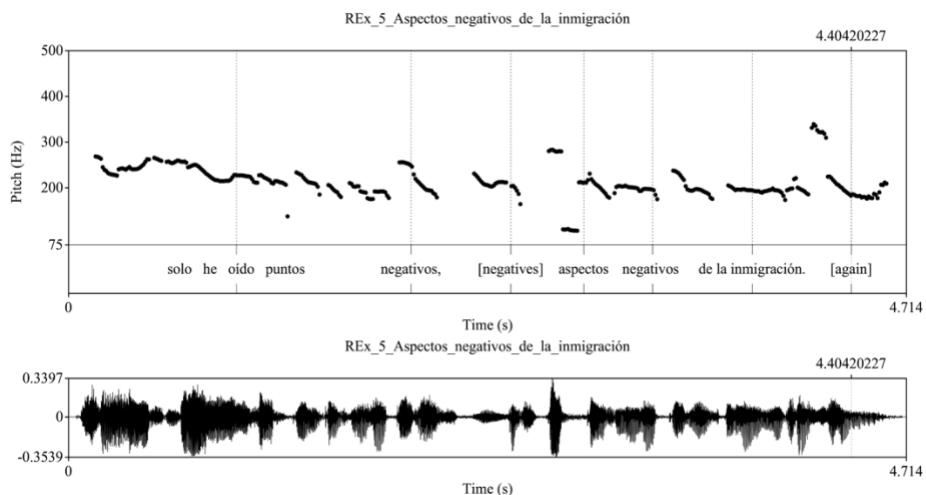
Image 23 Pitch contour and audio wave of the target text in R.Ex.5
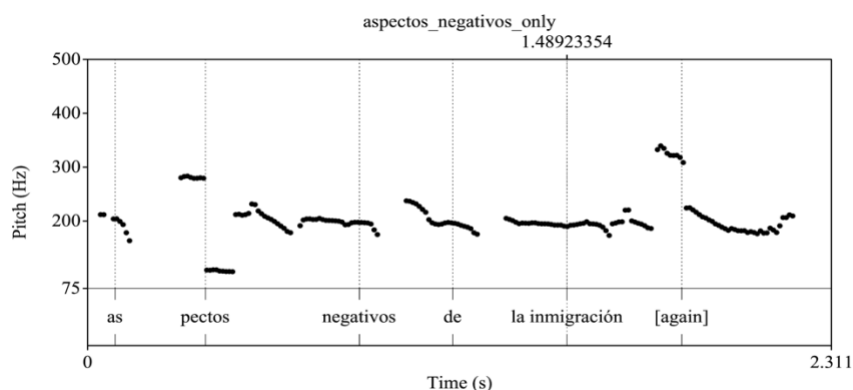
Image 24 Pitch contour for the second part of the target text in R.Ex.5

Image 24 above shows the pitch contour and audio wave for the target text in this example. Note that the words in brackets ("negatives", "again") belong to sound bites we can hear in the audio and video tracks, and there is a pitch line for these too. We see that the first "negativos" carries the nucleus in the IP because there is an inflexion in the pitch contour and, also, it is the last word of the IP (the common position for the nucleus in Spanish). Now, if we look at the next line, we can notice how the word "aspectos" has a pitch peak around 300 Hz, which is higher than the rest of the words. We can see this better in image 24. The second syllable of "aspectos" carries an emphatic stress which turns it into the nucleus of the IP, thus making it a narrow focus

IP. Consequently, we can only say that there is no prosodic-illocutionary rendition in this example.

R.Ex. 6 *Cupcake Wars*						
ST in	ST out	Character	Source Text	Target Text	TT in	TT out
00:19:42	00:19:43	Janell Brown	Well, I like Yoda. You said **Yoda**.	Y Yoda, ¿has dicho **Yoda**?	00:11:44	00:11:45

The example shown here has been retrieved from the reality show *Cupcake Wars* and we can see that the reaccented word in the source text is "Yoda". In the target text, this word is present twice, so it is also reaccented. In addition, the translation of the sentence "You said Yoda / ¿has dicho Yoda?" is literal, word by word, and the two nuclei ("Yoda") coincide since both are in broad focus this time. However, the target version shows something the source one does not have: question marks. Since the transcriptions of the programmes in our corpus (source and target texts) have not been provided by the production companies or translation agencies which produced the broadcast versions, but by a transcriber hired for the present study, we do not know whether this question mark was present in the original script to voice at the studio. Nonetheless, judging from the intonation used by the voice actor in this scene, we understand the question marks were present. By translating a sentence as an interrogative, when its source version was a declarative statement, we consider that the prosodic-illocutionary force is not rendered, given the fact that, as according to Leech and Svartvik (1994: 125) the pragmatic nuances between a statement and a question are different.

The table below shows the pitch contours for both sentences (source and target). As we can see in the picture on the left (source text), the pitch movement falls on the two "Yoda", which means that the speaker is not asking, as could have been the case if the tone was rising. On the other hand, the picture on the right (target text) shows a rising tone on the second "Yoda", which is typical of Spanish interrogative constructions (RAE 2011: 467- 468).

Source English text in R.Ex.6	Target Spanish text in R.Ex.6
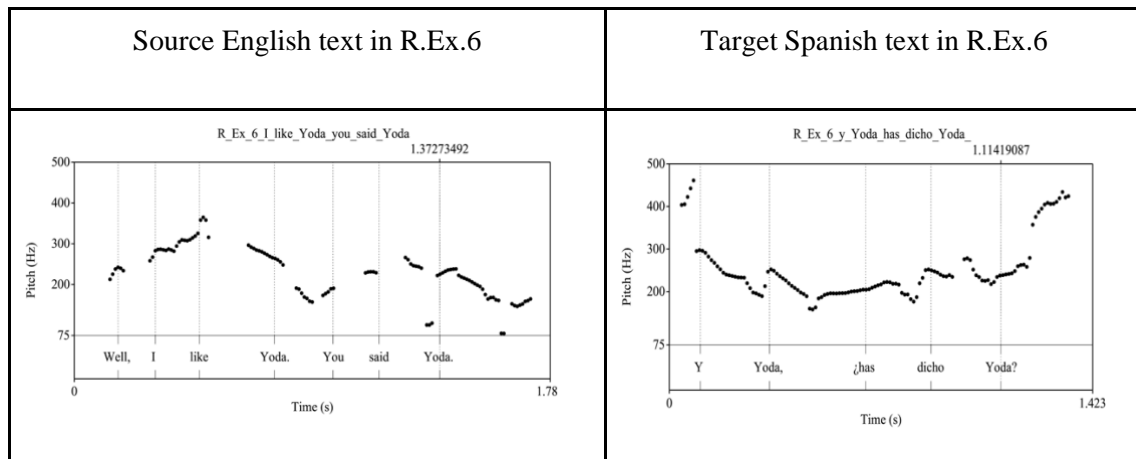
Figure 23 Pitch contours in R.Ex. 6

Now that we have analysed this example in detail, we can say that there is only partial rendition of the original message since the prosodic-illocutionary force has not been successfully rendered.

R.Ex. 7 *Teen Mom (clip 2)*				
In	Out	Character	Source text	Target text
00:12	00:14	Female	Yeah. And she went 'cause **she** wanted to go.	Y… ¿ha ido por **voluntad propia**?

This example, retrieved from the reality show *Teen Mom* (clip 2), shows a clear case of reaccenting of the word "she" in the source text. We can see that the translation strategy here is the deletion of the reaccented word, which is not present, not even once, in the target text. This strategy is quite common in Spanish because, as we saw in section 2.4, Spanish has a very rich morphology, which in turn allows it to have a free word order. In this example the translator can delete the subject of the sentence ("she") since it is already marked in the morphology of the verb "ha ido", which shows the third person singular, as stated by Alarcos (1994: 174). In addition, by using "voluntad propia" (on her own) the translator is reaccenting "she" as in the source text, thus reinforcing the idea that she went there because she had decided so. Nonetheless,

despite the fact that the reaccenting of the word "she" has been rendered satisfactorily, again, we observe that a declarative statement has been rendered as a question, as in the previous example, which consequently hinders the rendition of the prosodic-illocutionary force in the target text. The melody of both sentences is clearly visible in the pitch contours shown in the table below.

R.Ex.7 Source English text	R.Ex.7 Target Spanish text
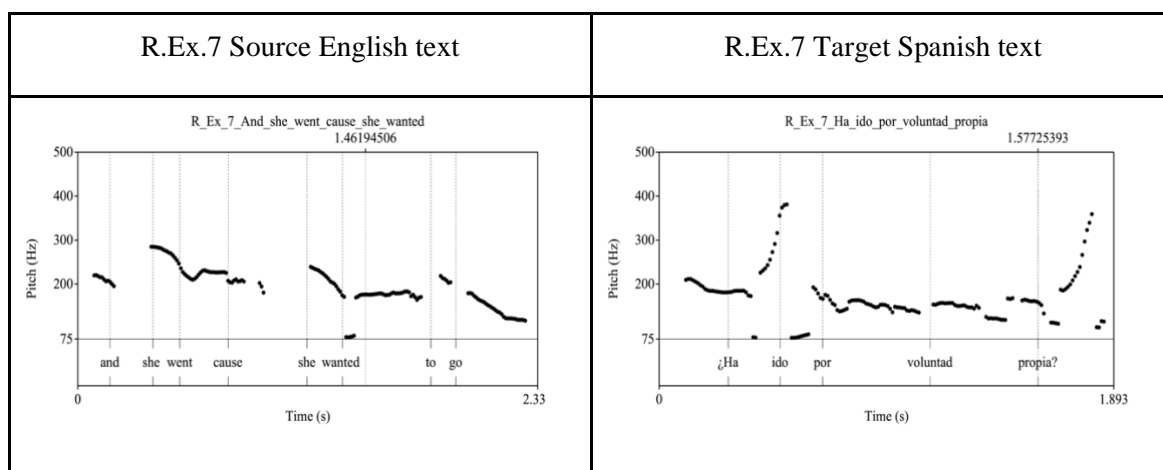
Figure 24 Pitch contour in R.Ex. 7

As can be seen in the table above, the pitch movement in the left graph shows a falling tone, while the tune in the image on the right is rising. In fact, there are two rising tunes, one on "ido" and another on "propia", which can be a consequence of the voice actor wanting to stress the word "ido", or probably to lengthen this line a bit with the characteristic intonation of interrogative sentences in Spanish.

Although falling tones can be characteristic of certain types of questions in English, as seen in section 2.2.1, and as stated by Wells (2006: 91), this example does not fall in any of those categories, which consequently implies that it is not a question, and that the intention of the speaker is not requesting information but confirming her thoughts. If the speaker wanted to ask a question or even request confirmation by checking if what she said was right, then she would have used a different grammatical construction, or use rising intonation, or a tag, or a "right?", to mean it. However, in this example, none of these are present and we understand that the rendition of the prosodic-illocutionary force is therefore not complete.

R.Ex. 8 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:42:13	00:42:16	Suzannah Lipscomb	But the problem is, historians don't always tell you	Pero la cuestión es que los historiadores no siempre nos dan <u>toda la información</u>
00:42:16	00:42:18	Suzannah Lipscomb	everything <u>you need to know</u>, and <u>what you need to know</u>	y en este caso lo que nos falta es el contexto,
00:42:18	00:42:21	Suzannah Lipscomb	in this instance is <u>the line before</u>, the context.	la <u>línea anterior</u>,
00:42:21	00:42:24	Suzannah Lipscomb	The **crucial** <u>line before</u>, is that he **<u>himself, Cromwell</u>**,	y <u>esa línea</u> aclara que **Cromwell**
00:42:24	00:42:29	Suzannah Lipscomb	had been commissioned by the King to put to an end the mistress's trial.	había recibido órdenes del Rey para poner fin al juicio.
00:42:29	00:42:33	Suzannah Lipscomb	So it actually means that the context is very much	Eso significa que
00:42:33	00:42:38	Suzannah Lipscomb	that <u>Cromwell</u> **himself** admits that Henry has told him to do it.	el **propio Cromwell** admitió que fue Enrique quien le pidió que lo hiciera.

This extract has been taken from the documentary *The last days of Anne Boleyn*, and shows 3 different examples of reaccenting in the source English text: "need to know", "the line before" and "Cromwell himself". Here we will only comment on the first and third, since the second belongs to the next section (b2), devoted to prosodic-illocutionary rendition only. Let us see how these two have been rendered in the target Spanish text. In the first case "you need to know" is reaccented in the source text because the speaker is reusing her own words to emphasise her point; by saying "and

what you need to know" she is giving new information, which she therefore reaccents. In the Spanish target text, we can see that "everything you need to know" has been translated as "toda la infomación", which is omitted by ellipsis in the next line. The fact that this part is deleted from the target text is not relevant for the communication of the whole message. We can see and hear in the clip that the second reaccented "what you need to know" is inserted in an intonationally subordinated IP, and therefore it is not that relevant for the rendition of the message in the target text. However, since intonational subordination is a topic which will be covered in the next section, we will analyse this in more detail further on in this dissertation.

The third example concerns the accenting of the reflexive pronoun "himself" which reaccents the pronoun and noun that goes with it, "he" and "Cromwell", which are also accented. Reflexive pronouns are not normally accented, only when they are used for emphatic purposes (Wells 2006: 130). As we can see, in the target text we have two different strategies to render this emphatic meaning. The first time we get just "Cromwell", while the second we have "el propio Cromwell". We understand that the second version is perfect to render the pragmatic value of the reaccenting, which in the source text is also reinforced by means of an emphatic intonation with a high fall tone. Nonetheless, the first time, by using "Cromwell" alone, the illocutionary force implicit in the source text by means of accenting the reflexive pronoun "himself" is lost.

Source English text	Target Spanish text
The crucial line before, is that he **himself,** Cromwell, had been commissioned by the King to put to an end the mistress's trial.	y esa línea aclara que **Cromwell** había recibido órdenes del Rey para poner fin al juicio.

As we can see in the table above, in the source text we have three accented words, being one the reflexive emphatic particle, whereas in the target text that emphasis is not achieved neither by syntactic means nor by an empathic intonation. A possible alternative to mark the emphasis as in the second time could be: "...aclara que **el mismo/mismísimo** Cromwell…" or "aclara que **fue Cromwell quien** había…".

2) Prosodic-Illocutionary Only

The only example in which there is just prosodic-illocutionary rendition and no rendition of the original message in this section is the second case of reaccenting in example R.Ex.8 above. As we can appreciate, the second time "the line before" is accented, it is accompanied by an adjective "crucial" ("the **crucial** line before"), which gives more emphasis to the reaccented material and defines that "line before" as relevant. In the target Spanish text, "the line before" is rendered as "línea anterior" and it is reaccented again in the next sentence "esa línea". However, we do not have a translation for the adjective "crucial", which in our view is essential for the audience. Thus, in this case, the transfer of basic semantic meaning has not been fully achieved, although the rendition of the prosodic element has. An alternative translation to include "crucial" in the target text could be "... y en este caso lo que nos falta es el contexto, la **línea anterior**, que **es crucial,** porque aclara que el mismísimo Cromwell …".

c) No rendition

There are two cases of no rendition in the reaccenting section, which will be analysed in the following pages. Our first example has been retrieved from the Euronews programme *Reporter: What's it like for EU migrants in Britain amid UKIP's success?*.

R.Ex. 9 *Reporter: What's it like for EU migrants in Britain amid UKIP's success?*				
In	Out	Character	Source text	Target text
03:17	03:23	Man 1	Seems too late really. Jobs been taken, houses **been taken**, aren't they?	Ya es demasiado tarde. Se quedan con los puestos de trabajo, tienen **casas**.

In this example we can observe how "been taken" has been reaccented by the male speaker to reinforce his idea that migrants have grabbed the jobs and housing benefits that belong to Britons by right. This man reuses his own words to support his

statement and emphasise his claim. However, the Spanish target text does not show this intensity. As we see, the first "been taken" has been rendered as "se quedan", literally meaning "they take (our)." which makes sense. Nonetheless, the second "been taken" is rendered as "tienen", literally "they have", which does not have the same emphatic weight as the first one, nor does it render the same semantic meaning as "been taken/se quedan". So, consequently, we can say that this time there is no rendition of either the prosodic-illocutionary function or the meaning of the original source text. A possible alternative here could be: "se quedan con los puestos de trabajo **y con nuestras** casas." Like this, we could use ellipsis for the reaccented material leaving some trace in the repetition of the preposition "con" and include "nuestras" to emphasise that the migrants are taking advantage of the British system, from the perspective of the interviewee.

R.Ex. 10 *Teen Mom (clip2)*				
In	Out	Character	Source text	Target text
00:00:54	00:01:15	Male	Noon! I opened the door and there she is in a blanket watching Netflix in our room. And I was like, "What are you doing?" I said, "You know, it's OK that maybe you weren't ready to be a mom." And she just kept saying, "I felt so bad because she kept crying so why am I so (beep - fucked up-)? And look at me. It is affecting you, and it's affecting **my mom**, and it's affecting..." And she said, "I just (beep -fuck-) everybody up."	Mediodía. Abría la puerta y me la encontraba en la cama viendo la tele de la habitación. Y no entendía qué estaba pasando. Le dije que no pasaba nada por admitir que no estaba preparada para la maternidad. Y ella decía que a mí me molestaba que llorara tanto. Veía que nos estaba afectando y que le afectaba a **su madre**. Sentía que molestaba a su madre. Sentía que molestaba a todo el mundo.

This example of reaccenting, taken from *Teen Mom* (clip 2), shows two intonational features, one in terms of stress, reaccenting, and another in terms of pitch, quotes. Here, again, we see how the different prosodic intonational features can be

222

combined in a single sentence. In the source text we have a male speaker quoting his girlfriend and there is a time in which he reaccents "affecting". It is important that we understand this line as if it was the girl speaking, because once we see the target text, we will get the impression that it is the boy who talks, so the meaning, consequently, changes.

Source English text	Target Spanish text
And she just kept saying, "I felt so bad because she kept crying so why am I so (beep - fucked up-)? And look at me. It is <u>affecting</u> you, and it's <u>affecting</u> **my mom**, and it's <u>affecting</u>..."	Y ella decía que a mí me molestaba que llorara tanto. Veía que nos estaba <u>afectando</u> y que le <u>afectaba</u> a **su madre**. Sentía que <u>molestaba</u> a su madre.

Focusing on how the Spanish translation here rendered the original text, we can see that the line "it is <u>affecting</u> you, and it's <u>affecting</u> **my mom"** has been translated as "Veía que nos estaba <u>afectando</u> y que le <u>afectaba</u> a **su madre**". The reaccented word in the source text is also reaccented in the target text, which is a usual strategy in this case. However, by reporting the quote and using the word "veía", we get the impression that this time the subject is `I´ (yo - the man speaking) and not `she´ (ella). This is due to the fact that, as has been already stated in this dissertation, Spanish has a rich word order, and we can leave out the subject, which is implicit in the morphology of the verb. Thus, "veía", has two possible grammatical subjects (`yo´/`I´, and `él´/he - `ella´/`she´).

We can notice that in the previous sentence the translator makes it clear in the reporting that it is `she´ who speaks as we get "Y **ella** decía…" but we do not get this in the following lines, which can lead to misinterpreting the target text. This confusion is even more evident in the last line "Sentía que molestaba a su madre". In addition, the word "molestaba" is also used to translate "felt so bad" two lines before and in that precise line the subject is the man speaking "a **mí** me molestaba"; therefore, the impression we get is that it is the man speaking his own words all the time, that he was bothering his girlfriend's mom.

To conclude, we can assume that neither the prosodic-illocutionary force nor the original semantic meaning of this example have been rendered in the target text.

5.2.1.2.1 **Concluding remarks on reaccenting rendition**

After reviewing how English reaccenting can be rendered in Spanish, we see that, as in the case of deaccenting, deixis (both anaphoric and cataphoric) work well to deliver full rendition of the original message. The other two possibilities observed, namely, ellipsis and reaccenting, are more numerous and also work well. However, one should also bear in mind that, when dealing with prosodic-illocutionary features, there are other factors that play the game, as is the case of tone in both the source and target texts. As will be seen a few sections further on in this chapter, misinterpreting a tone direction in English can lead to a translation error, and vice versa, as can happen with the delivery of a tone which does not render the same illocutionary force as the one in the original text, or with the placement of an emphatic accent on the wrong word.

Now that we have gone through the prosodic features related to English stress (deaccenting/reaccenting), we have seen that the strategies to render these in Spanish are similar: deixis (anaphoric/cataphoric), ellipsis and reaccenting. It is up to the translator to judge which of these works best for the text and whether there are other prosodic variables to consider before delivering the script to the recording studio where it will be voiced.

5.2.2 **Pitch**

In this section we will analyse the rendition of two features that English signals by means of pitch in connected fluent speech: intonational subordination and quotes (see section 2.2.2.4). Here we will not comment on tone pitch (whether the nuclear tone is high or low), as that will be dealt with in the tone section.

5.2.2.1 **Intonational subordination**

The total number of examples with intonational subordination, which normally provides superfluous information, in our corpus is 12 and, as can be seen in the pie chart below, only two options were found for its rendition: full and none.
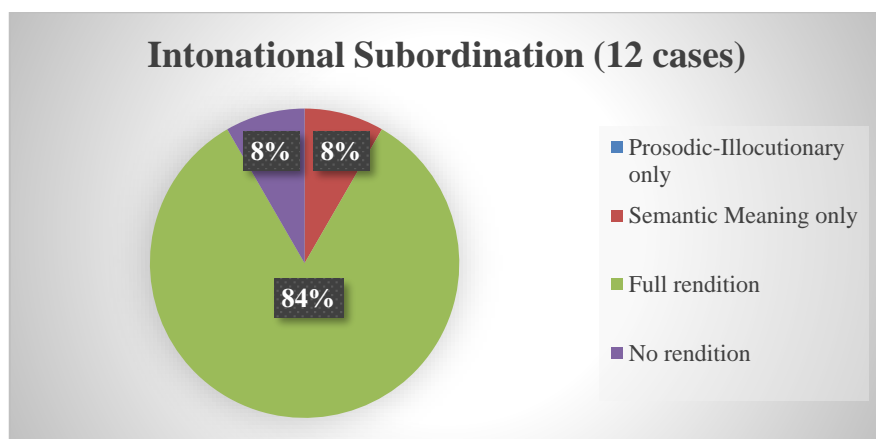
Figure 25 Examples with intonational subordination and their rendition in Spanish

The graph shows that almost all of the examples, 84%, were fully rendered, with 8% showing no rendition at all, and the remaining 8% rendering the semantic meaning only. The bar chart below shows the two possible strategies we have found in our corpus that have been used to render this pitch-related feature.
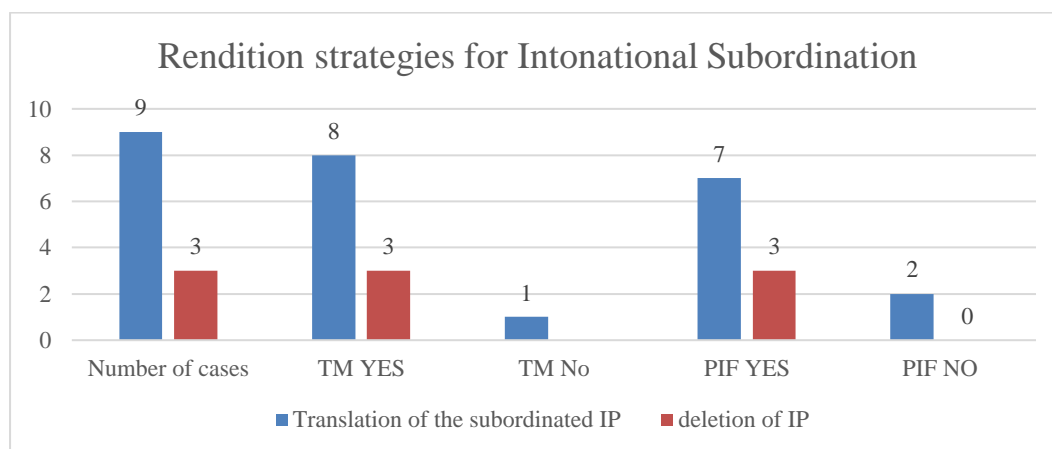
Figure 26 Bar chart with the rendition strategies for intonational subordination

As can be seen, there are just two options: translation of the original subordinated IP, and its deletion or omission in the target text. We notice that the commonest strategy, with 9 examples, is to translate the subordinated IP; whereas its deletion in the target text is just resorted to three times. Nonetheless, we can also identify that both the semantic meaning and the prosodic-illocutionary force can be rendered by deleting the subordinated IP, a good alternative in these cases.

Let us now see how these examples and strategies relate to the different genres within our corpus.

DOCUMENTARIES			
Intonational subordination	10	Translation of the subordinated IP	8
		Deletion of IP	2
INTERVIEWS VIDEOS			
Intonational subordination	0	Translation of the subordinated IP	0
		Deletion of IP	0
REALITY SHOWS			
Intonational subordination	0	Translation of the subordinated IP	0
		Deletion of IP	0
WEB VIDEOS			
Intonational subordination	2	Translation of the subordinated IP	1
		Deletion of IP	1

Table 24 Intonational subordination examples and their rendition strategy per genre

Only two of the genres analysed in our corpus show examples of intonational subordination: documentaries and web videos. Let us move on to the analysis of these examples starting with those which show a full rendition of the original message and finishing with the one example that fails to do so. An image with the pitch contour is displayed below the text for each example to show how intonational subordination works in English in terms of pitch, and can be measured with the right software.

a) **Full rendition**

I.S. Ex. 1[69] *TN Europe video Jan 17*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:00:54	00:01:02	Francis Ellison	Fund managers, <u>ourselves included</u>, have no better or more accurate answers <u>to these questions</u> than anyone else.	Los gestores de fondos, <u>entre los que nos incluimos</u>, no pueden responder ni de manera más satisfactoria ni más precisa que otros observadores.	00:00:5 4	00:01:0 2
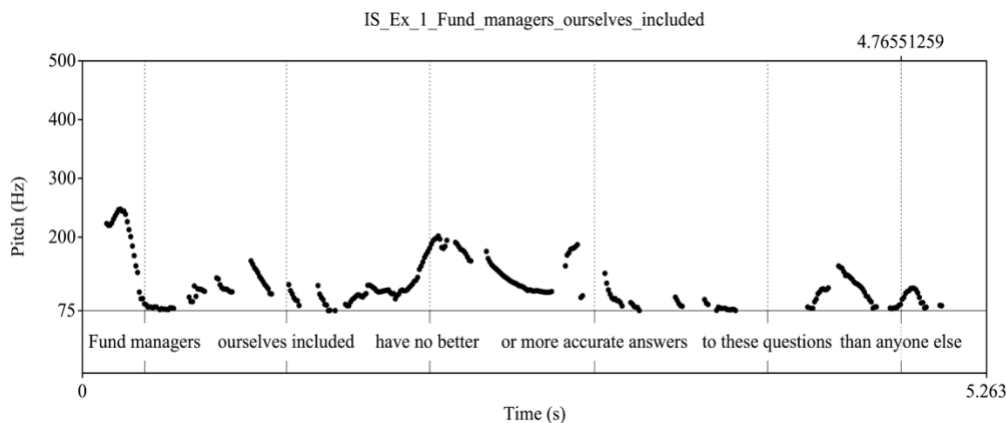
Image 25 Pitch contour in IS.Ex1

This example has been retrieved from the web video *TN Europe*, and we can see two instances of intonational subordination in the source text which are rendered in the target version by means of the two strategies identified for this prosodic feature: translation of the IP and deletion. The IPs in question are "ourselves included" and "to these questions", which are superfluous and do not provide any new information to the audience: regarding the first subordinated IP, as we know, or at least, as the usual audience of these videos knows, the company that produced the video and broadcasts it on its website, Columbia Threadneedle[70], is an investment company; the second

---

[69] IS.Ex. stands for Intonational Subordination Example.
[70] The website of Columbia Threadneedle can be found here: https://www.columbiathreadneedle.co.uk/

subordinated IP makes reference to a series of questions that appear a few lines before in the text.

The picture above has the pitch contour for this example, clearly showing how the average pitch range is around 200-100Hz while for the IPs under study it goes as low as 75Hz.

For the first subordinated IP the translator has opted to include it in the target text, which is a valid option, given the fact that these web videos are edited and, therefore, the target Spanish video clip is enlarged to match the target Spanish translation. One should note that in cases for which isochrony is tough to match, these sorts of IPs can be ruled out and left untranslated. Both the source text and the target one in this IP share the same nucleus "included/incluimos", so there is no dissonance recording the prosodic-illocutionary force, and since it is almost a word by word translation, the basic semantic meaning is rendered. Consequently, this example shows full rendition of the original meaning.

The second subordinated IP has been omitted from the target text and there is no loss of semantic meaning or illocutionary force. The questions which are referred to in the subordinated IP are given previously in the text, and rendering this IP in the target version could sound repetitive. The whole sequence of questions up to the line under analysis is shown below.

Source text time frame		Character	Source English text	Target Spanish text	Target text time frame	
00:19:00	00:21:00	Francis Ellison	Sitting in London Brexit stares us in the face.	Desde nuestra sede en Londres, el Brexit nos mira directamente a los ojos.	00:24:00	00:27:00
00:22:00	00:23:00	Francis Ellison	Will it be hard? Soft?	¿Se optará por una versión dura o blanda?	00:28:00	00:30:00

00:24:00	00:26:00	Francis Ellison	Orderly? Disorderly?	¿Se llevará a cabo de forma ordenada o desordenada?	00:31:00	00:34:00
00:27:00	00:29:00	Francis Ellison	Will it hurt us? or Europe? Or Both?	¿Perjudicará al Reino Unido o a Europa, o a ambos?	00:35:00	00:38:00
00:30:00	00:33:00	Francis Ellison	Did the UK electorate really want to leave the single market? And will we?	¿Quería el electorado británico abandonar el mercado único y se materializará esta salida finalmente?	00:39:00	00:44:00
00:34:00	00:36:00	Francis Ellison	What relationships will we build with the US and China in the future?	¿Qué relaciones se entablarán con EE. UU. y China?	00:45:00	00:48:00
00:37:00	00:42:00	Francis Ellison	Will Trump's isolationism, on the face of it aimed at China and Mexico, affect this too?	¿Afectará también el aislacionismo de Trump, aparentemente dirigido hacia China y México?	00:49:00	00:53:00
00:43:00	00:48:00	Francis Ellison	Fund managers, ourselves included, have no better or more accurate answers to **these questions** than anyone else.	Los gestores de fondos, entre los que nos incluimos, no pueden responder ni de manera más satisfactoria ni más precisa que otros observadores.	00:54:00	01:02:00

As can be seen, there are numerous questions, so the decision to eliminate that IP from the target text seems to be a good one since it provides superfluous information. Let us now move on to our second example for this category.

I.S. Ex. 2 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:13:47	00:13:49		The belief was, in the medieval world,	En el mundo medieval, cuando una mujer daba a luz
00:13:49	00:13:53		was that if a woman gave birth to a deformed or a malformed foetus,	a un feto deformado con malformaciones,
00:13:53	00:13:58	Philippa Gregory	then what everybody would genuinely, thoroughly and sincerely believe	la creencia generalizada y con plena conciencia de la gente
00:13:58	00:14:01		is that she's done a truly awful sin.	era que había cometido un pecado terrible.
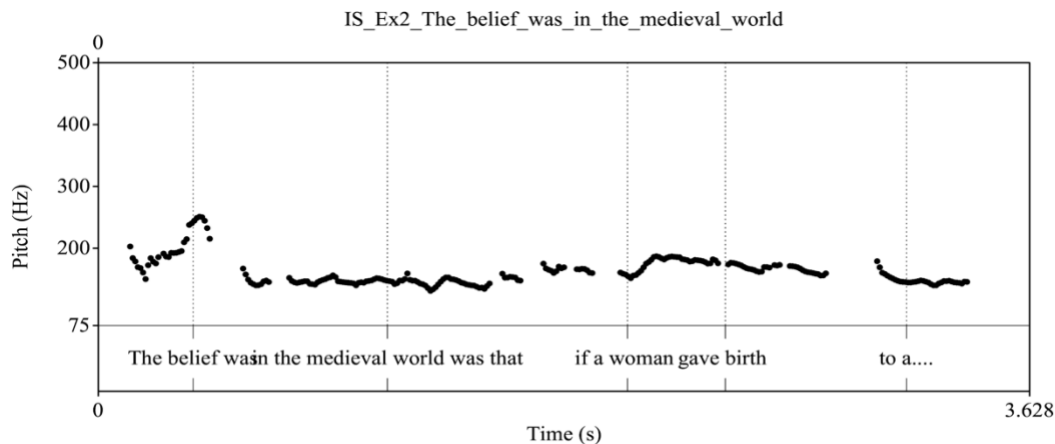
Image 26 Pitch contour for IS.Ex2

This example, extracted from the documentary *The last days of Anne Boleyn*, shows, again, that the subordinated material has been translated. The subordinated IP is "in the medieval world", and as we can see in the pitch contour image shown above, the pitch level decreases from around 200 Hz to around 100Hz during the time of the IP, to rise again over "woman". The IP under analysis is superfluous and could even sound redundant here, since the documentary is based on the Tudor court and the tragic events that took place in 1536, which, as we know from the context and the comments in the documentary, is set in the Middle Ages. With this in mind, the translator has

rightly decided to include this information in the target text and use it to open the sentence so as to set the context of what comes next. The result is nice since this initial position is often used for given information or relate back to something previously mentioned; and, since there are no isochrony issues, this rendition makes sense. Nonetheless, by starting the line with this IP, literal synchrony is sacrificed.

I.S. Ex. 3 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:12:32	00:12:37	Hilary Mantel	When Henry talks about enchantments, charms, magic tricks,	Cuando Enrique habla de encantamientos, hechizos, conjuros, magia,
00:12:37	00:12:39		Henry is beginning, it seems,	Es porque está empezando a
00:12:39	00:12:43		to think about annulling his marriage to Anne.	pensar en anular su matrimonio con Ana.
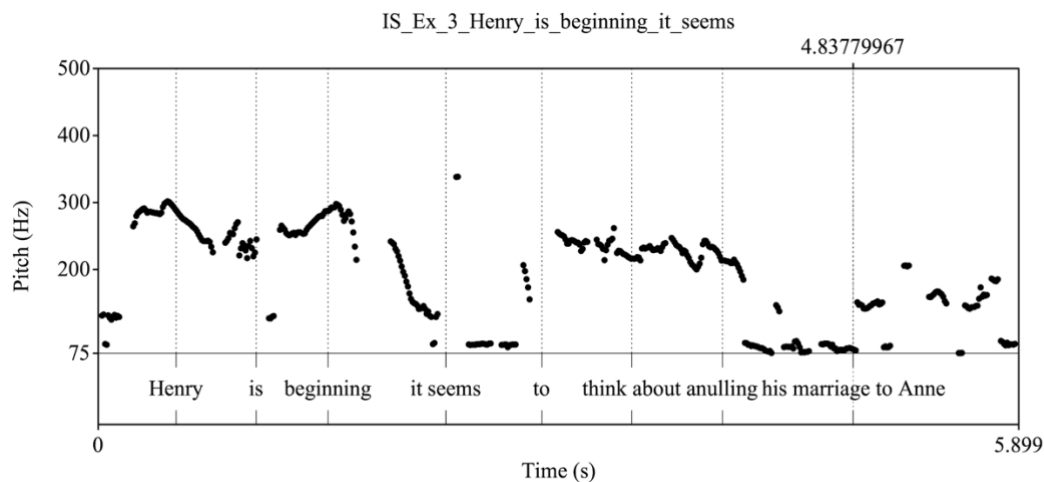
Image 27 Pitch contour for IS.Ex 3

This example, taken from the same documentary as the previous one, shows that the subordinated IP, "it seems", is a discourse marker or filler, and falls within the *mainly interactive* category of markers (Leech and Svartvik 1994: 13), as seen in section 2.3 of this dissertation. As we can see in the picture showing the pitch contour, the subordinated IP shows a lower pitch (200-75 Hz) compared to the rest of the sentence (200-300Hz), with the exception of the last part of the whole sentence, which

is deaccented. The intonationally subordinated IP in the English source text is rendered in the target Spanish text by means of ellipsis, that is, we do not have such IP in the target version. Since this IP does not provide any relevant information and can be considered as superfluous, the translator opted not to render it in the target text, as suggested by Chaume (2001: 81), Matamala (2008: 118) and Barzdevics (2012: 67). Therefore, since there is no modification of the semantic meaning of the message and no change in the illocutionary force, we understand that there is a full rendition of the original source text.

| I.S. Ex. 4 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:40:54	00:40:57		You know, he had been handed, <u>if I can change metaphors,</u>	En mi opinión, muy al contrario, <u>y si se me permite la expresión,</u>
00:40:57	00:41:02	Greg Walker	a really hot potato, and he wasn't happy to be investigating adultery	le habían pasado la patata caliente y no le hacía ninguna gracia tener que estar en los aposentos
00:41:02	00:41:03		in the Queen's Privy Chamber.	de la Reina investigando un presunto adulterio.
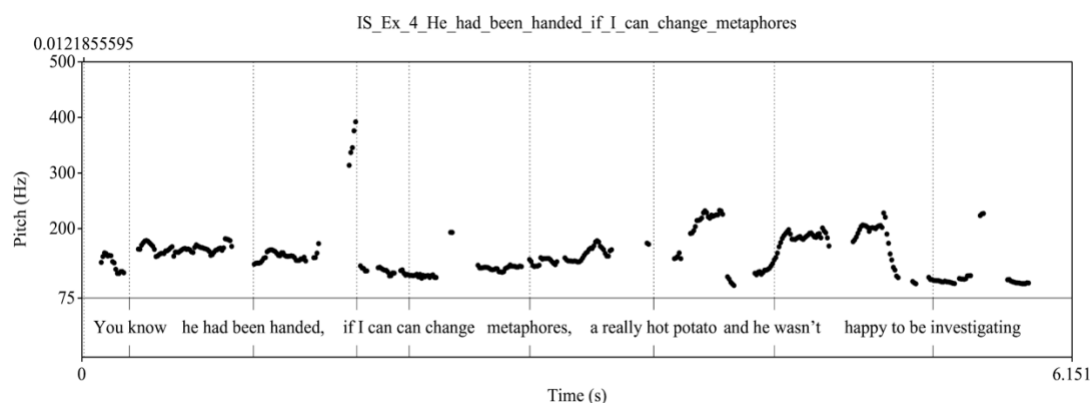
Image 28 Pitch contour for I.S.Ex.4

This last example for the full rendition section shows that, again, the subordinated IP has been translated into the target language. The audio clip is a bit long

(6 seconds) which makes it difficult to appreciate the lowering of the pitch in the image above, but we can see how there is a pitch increase both over "handed" and "really", which are the preceding and succeeding words of the subordinated IP.

This time, the speaker is off camera, so it would not have mattered if, instead of including this IP in the target text, the translator had not. However, considering that it serves to introduce a rather informal metaphor, "hot potato", which has also been rendered with an informal expression in Spanish in the target text and might not be very usual in history documentaries, the inclusion of the subordinated IP in the target text makes sense in our view.

b) No rendition

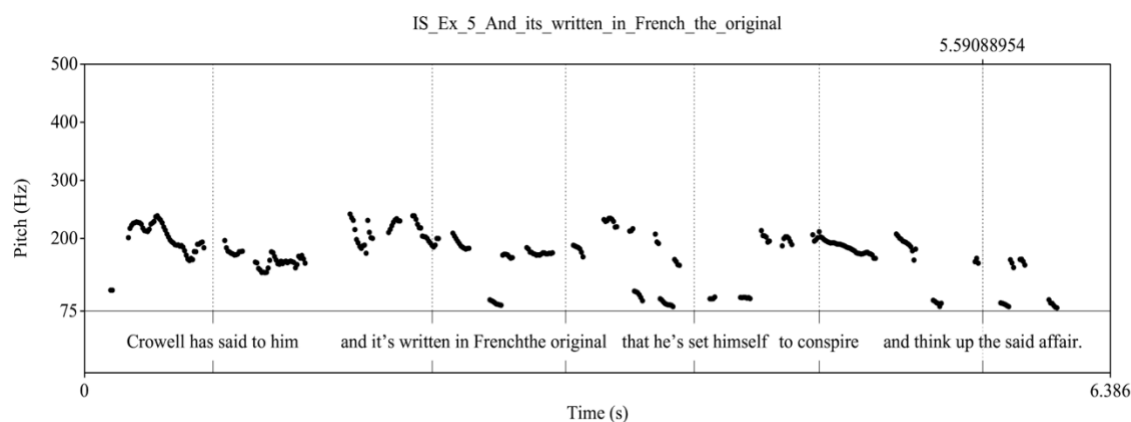| I.S. Ex. 5 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:42:06	00:42:09	Suzannah Lipscomb	Cromwell has said to him, and it's written in French, the original,	Cromwell le ha dicho "el original está escrito en francés",
00:42:09	00:42:13		that he's set himself to conspire and think up the said affair.	que se dispone a conspirar y urdir el complot.

Image 29 Pitch contour in IS.Ex.5

This is the only example in the intonational subordination category which does not render the original message, and has been retrieved from the documentary *The last days of Anne Boleyn*. Again, the extract for this example is a bit long, over 6 seconds, so the pitch contour is not very illustrative; however, once we hear the original audio file or watch the source video, we notice how, for this IP, the speaker increases her pace, so words are compressed, and there is a subtle lowering of the pitch.

We can see that the target version for this example shows an almost word for word literal translation, as in the previous times in which the translator opted to translate the subordinated IP. However, the issue with the rendition of this IP in the target text is the fact that, in its delivery, it sounds as if the speaker was quoting Cromwell, which is not the case in the source text. The inverted commas shown in the target text to indicate the quote have been included during the transcription process of the target Spanish video. We understand that the original script did not show the quotation marks and that there was an error during the voicing in Spanish of this scene; probably due to the large amount of text to voice in such a short time frame. Given the facts that the original speaker (Suzannah Lipscomb) speaks at quite a fast pace sometimes in her interventions and that the Spanish text is usually longer in terms of word count, the target voice actor fitted as many words as possible in a short time frame to match the source text.

In addition, this subordinated IP makes reference to a previous line of Suzannah Lipscomb's intervention. We have the whole sequence below.

In	Out	Character	Source text	Target text
00:41:59	00:42:02		The only piece of evidence that's used to say that it was a coup,	La única prueba en la que se apoya la teoría del golpe a la autoridad,
00:42:02	00:42:06	Suzannah Lipscomb	is a line in a letter from Eustace Chapuys, the Imperial Ambassador.	es una línea de una carta escrita por Eustaquio Chapuys, el embajador de España.
00:42:06	00:42:09		Cromwell has said to him, and it's written in French, the original,	Cromwell le ha dicho "el original está escrito en francés",

			that he's set himself to conspire and think up the said affair.	que se dispone a conspirar y urdir el complot.
00:42:09	00:42:13			

As we can see (in bold), she introduces the information that Chapuys wrote a letter in which Cromwell confessed he had conspired to frame Queen Anne, and then, by subordinating the IP under study, "and it's written in French, the original", the speaker introduces a side comment about the language in which the letter is written, but this fact is not that relevant. In order to include this IP in the target version without making it sound as a quote, we could place it in the part in which the letter in mentioned, e.g: "es una línea de una carta escrita en francés por Eustaquio...". As it stands, however, since the subordinated IP comes after the verb "dicho/said", we understand that, in this example, the basic semantic meaning of the original has not been transferred since Cromwell has not said anything in French, which is the impression we get from the target version. The prosodic-illocutionary force is not rendered either, since the original intention of the speaker is to make a simpl side comment on the original language of the letter Chapuys wrote.

5.2.2.1.1 Concluding remarks on the rendition of intonational subordination

In this section we have seen how intonational subordination works in context. As we know, these IPs have a parenthetical nature, since they provide little or no relevant information to the context of the conversation. Thus, it is up to the translator to include them or not in the target text, depending on the scene, or whether the speaker is on camera, and how close to the original text he/she wants the translation to be. Some of these IPs may include discourse markers, which, according to some scholars (Chaume 2001: 81; Barzdevics 2012: 67) are not to be rendered in the target tex. As we have seen, by not including these IPs in the target text, the original semantic meaning and the prosodic illocutionary force are left untouched.

5.2.2.2 **Quotations**

This section is devoted to the rendition of English quotations in Spanish voice-over translation. As we saw in section 2.4, quotations carry a significant load of prosodic-illocutionary force since they are signalled by means of intonation (pitch variation) and work as demonstrations. Therefore, as relevant communicative acts, they deserve their part in our analysis.

A significant number of quotations have been found in our corpus, which means, in our view, that this type of communicative act is a recurrent one in native English speakers' discourse. As can be seen in the pie chart below, there are 69 cases of quotations, with 59% showing full rendition, 6% no rendition at all, and 35% partial rendition (33% semantic meaning only, and 2% prosodic-illocutionary only).
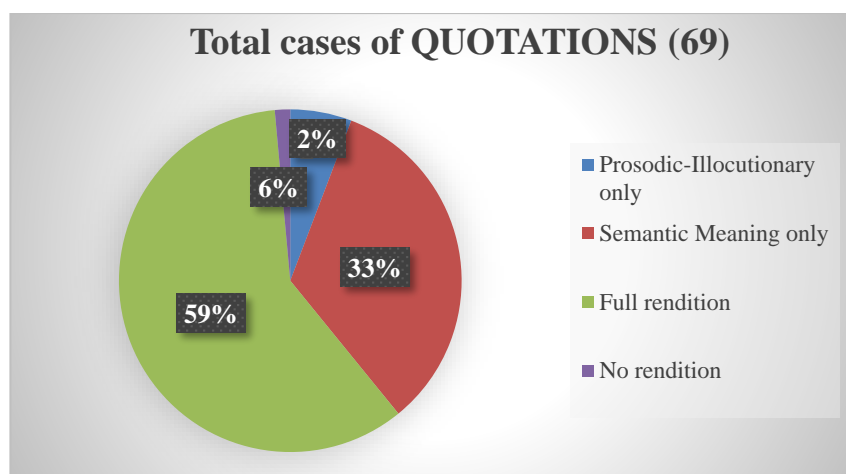

Figure 27 Pie chart showing the rendition percentage for quotations

We have identified several ways of rendering English quotations in Spanish voice-over texts, as can be seen in the bar chart below, namely: reporting the quote, keeping the quote in the target text, paraphrasing, and deixis (in one case in which the quotation was repeated).
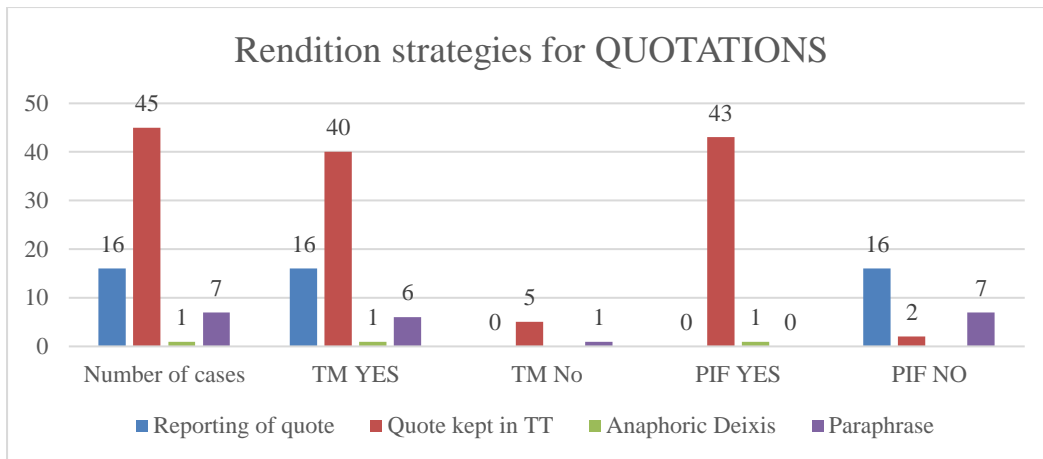
Figure 28 Bar chart with the rendition strategies for quotations

The chart above shows that the most commonly resorted to strategy, with 45 examples, is to keep the quotation in the target text for which 5 do not render the basic semantic meaning of the original, and 2 do not do so with the prosodic-illocutionary force. The second most used strategy is to report the quotation, with 16 examples. All of these succeed in transferring the original basic meaning in the target text, but none get to do it concerning the prosodic-illocutionary force. And this is also valid for the third strategy, paraphrasing, for which 7 out of 7 examples do not render the prosodic-illocutionary force, 6 transfer the semantic meaning, and 1 does not. Finally, there is one case of a quotation that is repeated twice; the first time the quotation is kept in the target text, but the second time it is rendered by means of anaphoric deixis in the target text, as will be seen in the qualitative analysis.

The preferred strategies in terms of genres, as per our corpus, is the following.

DOCUMENTARIES			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Quotes	51	Reporting of quote	15
		Quote kept in TT	33
		Anaphoric Deixis	1
		Paraphrase	2

INTERVIEWS VIDEOS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Quotes	3	Reporting of quote	0
		Quote kept in TT	2
		Anaphoric Deixis	0
		Paraphrase	1
REALITY SHOWS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Quotes	15	Reporting of quote	1
		Quote kept in TT	10
		Anaphoric Deixis	0
		Paraphrase	4
WEB VIDEOS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Quotes	0	Reporting of quote	0
		Quote kept in TT	0
		Anaphoric Deixis	0
		Paraphrase	0

Table 25 Number of quotation examples and rendition strategies per genre

The information shown in the table above tells us that, in all genres, the commonest strategy is to keep the quotation in the target text, followed by reporting and paraphrasing. However, we do not get to know when translators do prefer to keep the quotation and when to use reporting or paraphrasing. It can probably be due to a desire to introduce variation in the target text, as not all quotations are rendered exactly, but we cannot confirm this. From the figures above, we see that, in the case of

documentaries 33 out of 51 quotations are kept and 17 are reported/paraphrased; for interviews videos, 2 out of 3 are kept and 1 paraphrased; and for reality shows, 10 out of 15 are kept and 5 reported/paraphrased. This shows that more than 60% of quotations in our corpus have been rendered as quotations in the target text, while the remaining 40% of these are either paraphrased or reported.

Let us now move onto the qualitative analysis of the translation of some of these quotes, starting with those that show full rendition of the original English text.

a) **Full rendition**

Q. Ex. 1₇₁ *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:43:47	00:43:52	David Starkey	As relations between Anne, Henry and Cromwell	La relación entre Ana, Cromwell y el Rey
00:43:52	00:43:57		all become increasingly fraught, the signal must come from Henry.	era cada vez más tensa y quien tenía que tomar cartas en el asunto era Enrique.
00:43:57	00:44:01		"I am fed up. We want this woman out of the way."	"Estoy harto, quiero desacerme de esta mujer de una vez por todas."

This first example of full rendition of the original quotation has been retrieved from the documentary *The last days of Anne Boleyn*, the programme which, within our corpus, provides the most number of examples. This one example shows David Starkey, a renowned historian, commenting on how King Henry VIII made up his mind and committed himself to getting rid of his second wife (Anne Boleyn). As we can see, the line preceding the quote states "the **signal** must come from Henry" after which we get the quotation indicated in bold. So, what Starkey is doing here is giving us the signal (probably those were not the exact words of King Henry VIII) or the order the King gave Cromwell to go on and orchestrate the set up against Anne. Starkey even uses the

---

₇₁ Q.Ex. stands for Quotation Example.

royal we or majestic plural to give more realism to his quotation, thus giving the impression that the king Henry VIII is actually talking.

In the Spanish target version, there is no word to introduce the quotation as "signal" does in the English original on the preceding line. However, the quotation is there and its delivery shows the audience that the speaker is showing how King Henry VIII gave Cromwell the order. The voice sounds more in character, as impersonating, and uses a more emphatic tone, mostly on the words "harto" (fed up) and "de una vez por todas" (out of the way), compared to the previous lines when David Starkey is talking. Therefore, we can say that this rendered quotation is also a demonstration, so the prosodic-illocutionary force of the source text is fully transmitted, together with its basic semantic meaning, resulting in full rendition.

Q. Ex. 2 *Kitchen Nightmares*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:11.20	00:11.27	Gordon Ramsay	All this fucking fridge jammed up with shit and we're standing here saying, 'Put it in the bin, put it in the bin.' Well, get it in the fucking bin!	Toda esa comida, sólo tienes mierda y estamos aquí diciendo: "Tírala a la basura, tírala a la basura". Pues tírala a la maldita basura.	00:11:0 4	00:11:1 1

This extract has been retrieved from the reality show *Kitchen Nightmares*, and we can see Gordon Ramsey talking to his apprentice. This time, the quotation "put it in the bin" is preceded by the reporting verb "saying" and this is also reflected in the Spanish target version with the verb "diciendo". Gordon Ramsey is a very expressive man and he shows a lot of passion and intention in almost everything he says on camera, as can be seen in this quotation. He is criticising his pupil for not keeping the fridge clean and he is also mocking him somehow by quoting his words with a tone of voice which mimics the apprentice's. This is a clear example of quotation as demonstration.

The Spanish translation for this line shows a word by word translation with emphatic intonation on the quote. Since the original northern English accent of the apprentice is neutralised in the Spanish voiced-over version of the show, the delivery of the quotation by the actor voicing chef Ramsey does not show funny accent or tone of voice, since both voice actors (for Gordon Ramsey and the apprentice) use neutral Castilian Spanish and this nuance is lost. Nonetheless, the irony and reprimanding intention achieved with this quotation, in our view, equals that of the source version, and the original quotation is also rendered as a quation in the voice-over, so we can say that there is full rendition of the original message.

Q. Ex. 3 Cupcake Wars						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:13:40	00:13:47	Ashley Eckstein	While I did enjoy the flavor of the cupcake, as Vader would say: "Your lack of icing is disturbing".	Aunque el sabor me gustó, como diría Darth Vader: "Tu falta de icing es inquietante".	00:06:18	00:06:24

This third example of full rendition has been retrieved from the reality show *Cupcake Wars* (Star Wars episode) and shows how one of the members of the jury that evaluate the cupcakes, Ashley Eckstein, gives some feedback to one of the contestants. This episode is dedicated to the film saga *Star Wars* and Ashley Eckstein is the voice actor behind Ashoka Tano, one of the characters in *Star Wars: The Clone Wars*. In the original English text, we can see and hear how Ashley gives her feedback quoting one of Darth Vader's most popular lines "I find your lack of faith disturbing" but changing "faith" for "icing", which is the downside of the cupcake she evaluates. In addition to the quote, she also changes the speed and tone of her voice to imitate Vader's. This is, again, another example of a quotation as a demonstration since she is demonstrating, or imitating, how Darth Vader himself would judge the cupcake.

The Spanish target text renders the prosodic-illocutionary force of the original perfectly, since the Spanish voice delivers the text as if Darth Vader was speaking through her. In addition, there is a little pause before the quote, which gives a bit of suspense to the feedback (the quote itself). In terms of transfer of semantic meaning, there is a literal translation (word by word), which makes sense and works perfectly. It is curious that while in the source English text we get one of Darth Vader's famous quotes, the Spanish did not use the exact words Vader says in the Spanish translation of the film, shown in the table below. This, in our view, would have given more realism to the translation and would also have rendered the intertextuality present in the source text. Nevertheless, we can still consider the translation as a case of full rendition, since this is something that only Star Wars mega fans can notice and the line has probably gone unnoticed by the general public.

Darth Vader original English quote	Ashley Eckstein's quote	Darth Vader Spanish translated quote	Ashley Eckstein's quote in Spanish
Your lack of faith is disturbing	Your lack of icing is disturbing	Su carencia de fe resulta molesta.	Tu falta de icing es inquietante.

Table 26 Darth Vader's quote and translation.

Q. Ex. 4 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:31:08	00:31:10		And to cut a long story short she said,	Para resumirlo un poco le dijo:
00:31:10	00:31:12	Suzannah Lipscomb	"If you think I'm bad, you should see the Queen.	"Si consideras que mi comportamiento no es adecuado, deberías ver a la Reina,
00:31:12	00:31:15		She entertains men late at night, including Mark Smeaton,	recibe a hombres a altas horas de la noche, Mark Smeaton es uno de ellos,
00:31:15	00:31:17		who's a musician at the Queen's court."	uno de los músicos a su servicio."

In	Out		Source text	Target text
00:31:19	00:31:23		And this was pyrotechnic intelligence.	Fue una maniobra de inteligencia pirotécnica.

This last example of full rendition has been found in the documentary *The last days of Anne Boleyn,* and shows Suzzanah Lipscomb quoting Lady Worcester when she told her brother about Anne Boleyn's alleged infidelities with different men in the court. This time, the original English quote is introduced by a reporting verb (said) and this is also shown in the Spanish translation with the verb "dijo". The quote is used to demonstrate how Lady Worcester cheated on the Queen by telling her brother about the latter's love affairs at court. In addition, keeping the quotation in the target text works perfectly for two reasons: first, it renders the prosodic-illocutionary force of the original; and second, because in this scene of the documentary we see two actors (Lady Worcester and her brother) performing that very moment, which synchronises totally with the quote.

b) Partial rendition

We will start the analysis of the examples which show a partial rendition of the original meaning with those which only render the semantic meaning, and finish with those which focus on the prosodic-illocutionary force.

b1) Semantic meaning only

Q. Ex. 5 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:11:44	00:11:46		He thinks that, as King,	Cree que como rey,
00:11:46	00:11:50	Hilary Mantel	he has a direct relationship with God.	tiene una vinculación directa con Dios.
00:11:50	00:11:55		"So why? Am I still not on the right side of God?	Y se pregunta por qué no está a su lado,

243

00:11:55	00:11:57		What does God want of me now?"	qué será lo que quiere de él.

This first example of partial rendition (semantic only) has been taken from the documentary *The last days of Anne Boleyn*, and in the original English text we can see how Hilary Mantel, once again, quotes King Henry VIII talking to himself, wondering what is God's next move against him after not allowing him to father a son with Catherine of Aragon during his first marriage, followed by Anne Boleyn's miscarriage (of a boy) years later during his second marriage, and now Anne's love betrayal. Hilary Mantel, again, demonstrates what Henry could have said by quoting, and uses a tone of voice which truly depicts the King's fatigue and disappointment with God. However, if we analyse the Spanish translation for this scene, we can see that the quote has been paraphrased, which makes it work as a description but not as a demonstration, and therefore, the prosodic-illocutionary force of the original text is not fully rendered. In addition, we can notice that the sound bites at the start and at the end of the quote do not match literally.

	Sound bite START	Sound bite END
English source text	So why?	Want of me now
Spanish target text	Y se pregunta por qué	Lo que quiere de él.

In the start sound bite "so why" we get a close-up frame of Hilary Mantel, in which she moves her hands outwards to reinforce her/Henry's complaint towards God. The speaker is characterising king Henry VIII both with her voice and with her body language, which have a significant pragmatic load in this case. Unfortunately, with the paraphrasis in the target text, this is not achieved, so we only get partial rendition of the original message.

			Q. Ex. 6 Teen Mom (clip 2)	
In	Out	Character	Source text	Target text
00:15	00:52	Male	She so wanted to go. She just said, "I don't wanna fake this anymore." She so was having a hard time. Because I told her, "I'm gonna be honest with you, honey." I said, "I can't do this for the rest of my life. That's all I'm telling you right now." And she told me, she's like, "No, I don't either." I said, "Because this whole first year of marriage, this whole first year of being a parent, it is so opposite of what I thought." I think I started noticing things when Nova was like around 3 months old. I started noticing like her... "But I don't know. You're a new mom. I don't know what you're doing." I eventually went and took Nova with me to wake her up by noon. I am like, "OK, you know what, dude? Get your ass up!"	Quería ir. Dijo que ya no quería sentirse así más. Lo está pasando mal porque le dije que siendo honesto yo no podía seguir así el resto de mi vida. Y me dijo que ella tampoco. Lo dije porque después de un año casado y un año siendo padres, no es como pensaba. Creo que empecé a notar cosas cuando Nova tenía unos 3 meses. Empecé a notar que ella, ya sabes… Pero no lo sabía. Era primeriza y no sabía lo que hacer. Al final yo me quedaba con Nova hasta que la despertaba a mediodía. Tenía que insistirle para que se despertara.

This example has been retrieved from *Teen Mom* (clip 2 in our corpus) and illustrates, in our view, partial rendition of the original message in the quotation. As we can see, the extract is a long one, showing the male character talking to a woman about his relationship with his wife. Both (the boy and his wife) are teenagers and parents, and he is complaining about his wife's behaviour after their daughter was born. In his speech, the boy quotes, thus demonstrates, the conversation they had about being parents. He uses his own words and also his wife's to depict the scene with the actual language they used. The Spanish target text, on the other hand, uses reported speech for this whole scene, and the realism and pragmatic value of the original text is not fully

rendered. The following table shows how the conversation was quoted in the source text and reported in the target text, in more detail.

	English source text	Spanish target text
1	She just said, "I don't wanna fake this anymore."	Dijo que ya no quería sentirse así más.
2	I told her, "I'm gonna be honest with you, honey." I said, "I can't do this for the rest of my life. That's all I'm telling you right now."	le dije que siendo honesto yo no podía seguir así el resto de mi vida.
3	And she told me, she's like, "No, I don't either."	Y me dijo que ella tampoco.
4	I said, "Because this whole first year of marriage, this whole first year of being a parent, it is so opposite of what I thought."	Lo dije porque después de un año casado y un año siendo padres, no es como pensaba.
5	I started noticing like her... "But I don't know. You're a new mom. I don't know what you're doing."	Empecé a notar que ella, ya sabes… Pero no lo sabía. Era primeriza y no sabía lo que hacer.
6	I am like, "OK, you know what, dude? Get your ass up!"	Tenía que insistirle para que se despertara.

As we can see, there are 6 quotations in this passage. In the first one we get a truly hard and direct statement ("I don't wanna fake this anymore") with a heavy load of illocutionary force. The wife sounds angry and tired of her marriage and the whole situation of being a mother, as if this is the really the last straw. However, in the Spanish version we get a soft version with a considerably more neutral pragmatic meaning "dijo que ya no quería sentirse así más".

The second passage shows the husband's answer, which, like the wife's, is really direct and tough: we get the word "honey" to refer to the wife and "that's all I'm telling you right now" to close the statement, which shows the boy is putting his foot down. This is not reflected in the target version, which depicts a conversation instead of a row.

In the fifth quote the Spanish version is a bit confusing, in our view, because of the repetition of the verb "saber/to know", which appears three times. In addition, in the final "y no sabía lo que hacer" we do not know whether the subject of the verb is the husband or the wife. From the context of the Spanish sentence we assume it is the wife but, in the original text, the subject of that part is the husband ("I don't know").

Finally, the sixth quote is, again, very direct and hard in the source text. There is some kind of tiredness, irony and anger implicit in this statement, which has been completely softened in the target text. In addition, the word "despertara" (awoke) does not work very well here because the wife is already awake (she's talking to her husband). Probably "levantara" would have worked better, because what the husband actually wants is his wife to get up and out of the bed so she can look after her daughter. All these pragmatic and illocutionary nuances are not rendered in the target text, and what we get is just a partial, basic and softened rendition of the original message.

The example we have just analysed is rather unusual, as there are six quotations in a very short time frame. If the translator had decided to render these in the target text by means of quotations rather than reporting them, the result could have been confusing to the target audience. However, the fact that all the quotations of the source text have been omitted in the target text, together with the softening of the tone in the reported speech sentences, hinder the full rendition of the pragmatic meaning of the original English text.

Q. Ex. 7 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:31:59	00:32:03	Greg Walker	So he takes him back to his house, and questions him.	Lo lleva a su casa y lo interroga.
00:32:03	00:32:05		We're not sure whether torture was used.	No sabemos si lo torturó.
00:32:05	00:32:08		Some people say there was torture, other people say there wasn't,	Algunos dicen que sí y otros que no.

| 00:32:08 | 00:32:11 | | but he remarkably confesses. | Pero él confiesa. |
| 00:32:11 | 00:32:14 | | He says, "I had sex with the Queen on three occasions." | Asegura haber mantenido relaciones sexuales con la Reina en tres ocasiones. |

This example, again, shows a quote that is paraphrased in the target text, and therefore, the original prosodic-illocutionary force is not fully rendered in it. We can see how Greg Walker speaks for Mark Smeaton, who was being accused of having an affair with the Queen. In the quote itself we get the demonstration of Mark's confession, and in the documentary, we actually see two actors performing the scene while the confession takes place; one playing Mark Smeaton and another Thomas Cromwell, the accusing party. However, in the target text the quote is paraphrased, so the demonstration is lost. Example 4 above included a similar case, with some actors playing the scene that is being described and demonstrated by the interviewees in the documentary. In that example, the quote was kept in the target text, so is a bit intriguing why this has not been the case this time. Nevertheless, despite the fact that the prosodic-illocutionary force has not been rendered, the basic semantic meaning of the original has been transferred, so we can say there is partial rendition of the source message.

Q. Ex. 8 *Cupcake Wars*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:16:34	00:16:56	Candace Nelson	What a very ambitious cupcake here! When you were listing off all your ingredients, I thought: "I wonder how all those flavors are gonna work". And you know what? They did.	Es un cupcake muy ambicioso. Cuando mencionaste todos los ingredientes, me pregunté si quedarían bien juntos, y la verdad es que sí funcionaron.	00:08:54	00:09:13

This last example of partial rendition (semantic only) shows a scene from *Cupcake Wars* in which one of the judges is giving a contestant this verdict. The judge, Candace Nelson, introduces the quote with a reporting verb ("I thought") and then pauses slightly before quoting herself. This is not shown in the Spanish target text, since the quote is paraphrased, and the demonstration is rendered as a description.

b2) Prosodic-illocutionary only

There are three examples showing only prosodic-illocutionary rendition (no meaning): two from the documentary *The last days of Anne Boleyn*, and one from *Teen Mom* (clip1).

Q. Ex. 9 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:16:45	00:16:48		She kisses the letter.	Jane besó la carta
00:16:48	00:16:51		Jane hands it back to the messenger,	y se la devolvió al mensajero
00:16:51	00:16:55		saying, "It would be quite improper for me to take this,	diciéndole que sería inapropiado aceptarla.
00:16:55	00:16:57	Hilary Mantel	"but please tell the King	"Dígale al rey
00:16:57	00:17:01		"that he should send it again	que me puede volver a enviar obsequios
00:17:01	00:17:05		"when I should happen to make a good marriage."	cuando esté honorablemente casada".

This example shows a quote translated as a quote in the target text, which means that the prosodic-illocutionary force is rendered. However, the semantic meaning of the

original is not successfully transferred, in our view: Hilary Mantel is commenting on the seduction game between king Henry VIII and Jane Seymour (Henry's wife after Anne Boleyn); Mantel's words are accompanied, once again, with images of actors performing the scene between Jane Seymour and the member of the king's court who delivers the letter. The first problem in the target text appears when "obsequios" is used to refer to the letter Jane has just received ("tell the King that he should send it (the letter) again...". The Spanish word used here implies that there is more than one thing since it is in the plural, and gives the impression that Henry has given Jane some presents apart from the letter that the interviewee is talking about. The second problem is in the very last line. The translation sounds confusing since, if Jane were married, there would be no point in sending her love letters then; while, if she is married and her heart belongs to another man, she would not be allowed to get letters from anyone since that would be considered a sin. What we understand from the original text means is that Henry should send the letter again once he has got rid of his current wife (Anne Boleyn) and is ready to marry her (Jane). The English grammatical construction "happen to + infinitive", is commonly used to refer to things which are unlikely or not too likely to happen (Swan 2014: 261). This is how Jane teases and flirts with Henry VIII, by rejecting the letter and telling him she will probably not be his wife. But this is not reflected in the Spanish version of the documentary, which in fact conveys a rather different meaning.

Q. Ex. 10 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:30:50	00:30:54	Suzannah Lipscomb	One of Anne's ladies, Lady Worcester,	A una de estas damas, Lady Worcester,
00:30:54	00:30:59		was being told off by her brother for her loose living.	su hermano le estaba censurando su comportamiento libertino
00:30:59	00:31:02	Hilary Mantel	She says, "Huh! Don't blame me!"	Y ella le contestó: "Bah,
00:31:02	00:31:06		"It's nothing to what the Queen gets up to,"	eso no es nada comparado con lo que hace la reina."

00:31:06	00:31:08		or words to that effect.	O algo por el estilo.

In this example we can see again Hilary Mantel quoting someone else: Lady Worcester, one of Anne Boleyn's ladies at court, who was reprimanded by her brother for her loose living. At the beginning of the quote, we see Mrs. Mantel moving her hands towards her chest when she says "Don't blame me!". This sentence within the quote works as Lady Worceseter's self-defense statement (for Hilary Mantel at least), before accusing Anne Boleyn of showing a worse moral behaviour than hers. This sentence works as an IP and shows narrow focus although the nucleus falls on the last item of the IP, "me", which, as a pronoun, is only accented in English for contrasting purposes. The image below shows a graphic illustration of the pitch contour for this IP (in which we can also observe a stepping head, which is the emphatic variant of typical high heads as we saw in section 2.2.1).
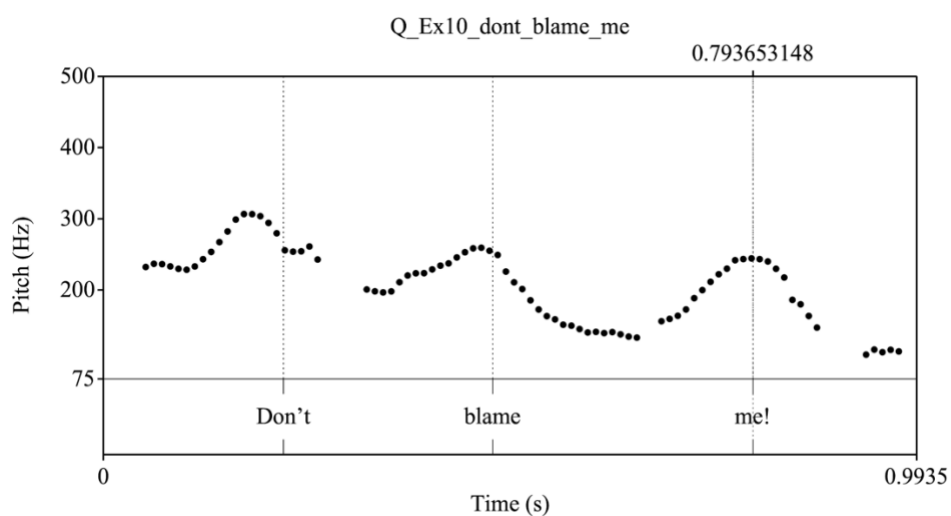

Image 30 Pitch contour in Q.Ex10

This line, which is significant from an illocutionary and kinetic perspective, has not been translated in the Spanish target text. Therefore, this piece of information and its contrastive pragmatic load is lost. In addition, the kinetics in this scene do not work, since the hand gesture is left on its own without any words that refer to the subject of the sentence.

Q. Ex. 11 *Teen Mom (clip 1)*				
In	Out	Character	Source text	Target text
00:19	00:34	Interviewer, female	I'm sorry. Back up for a second. You were thinking, "I'd like to design a ring with my daughter that is supposed to be symbolic of my connection and my commitment with Simon. I'm not going to involve him in the design of this ring…"	Espera, retrocede un momento. Pensaste: "voy a diseñar un anillo con mi hija que simbolice mi conexión y compromiso con Simon pero no pudo participar del diseño del anillo."

This last example partial rendition (prosodic-illocutionary only) has been taken from the reality TV show *Teen Mom* (clip 1). It shows how the interviewer quotes Farrah, the protagonist, explaining her decision to design an engagement ring. In the Spanish translation we can see that the quote is kept and the meaning is basically the same at the beginning. However, the last part of the quote is conveyed through reported speech, which makes a strange mixture with the first part of the sentence:

English source text	Spanish target text
I'm not going to involve him in the design of this ring…"	(...) pero no pudo participar del diseño del anillo.

As we can see, in the paraphrased bite the subject has been changed. Instead of reporting the subject "I" in the original quote, "Simon/he" is implicit in the morphology of the Spanish verb "pudo" (third person singular), which changes the meaning of the sentence and the quote. And we do not know whether this part is part of the quote or not since there is no pause between the start of the quotation and this final part in which we have indirect speech. Therefore, we can say that the original speaker's intention, which is shown in the demonstration of the quote in the source text, is kept but the semantic meaning is not, since its transfer is hidered in the paraphrased part at the end of this scene.

c) No rendition

There is only example in our corpus in which there is neither rendition of the original message nor of the prosodic-illocutionary force. It has been retrieved from the reality show *Teen Mom* (clip 2).

Q. Ex. 12 *Teen Mom (clip 2)*				
In	Out	Character	Source text	Target text
00:54	01:15	Male	Noon! I opened the door and there she is in a blanket watching Netflix in our room. And I was like, "What are you doing?" I said, "You know, it's OK that maybe you weren't ready to be a mom." And she just kept saying, "I felt so bad because she kept crying so why am I so (beep - fucked up -)? And look at me. It is affecting you, and it's affecting my mom, and it's affecting..." And she said, "I just (beep -fuck-) everybody up."	Mediodía. Abría la puerta y me la encontraba en la cama viendo la tele de la habitación. Y no entendía qué estaba pasando. Le dije que no pasaba nada por admitir que no estaba preparada para la maternidad. Y ella decía que a mí me molestaba que llorara tanto. Veía que nos estaba afectando y que le afectaba a su madre. Sentía que molestaba a todo el mundo.

This example shows a quote in the original English text which has been indirectly reported in the Spanish translation, so the original speaker's intention of demonstrating is lost. Consequently, the prosodic-illocutionary force is not fully rendered in the translated version. In terms of transfer of semantic meaning, we can see that the last part of the Spanish text is a bit confusing because we do not really know if the verbs "veía", "sentía" and "molestaba" refer to the male speaker (I/yo) or to the girl (she/ella) that is being quoted in the source text. This is due to the morphology of Spanish verbs, which do not require an explicit subject, as it is implicit in their morphology. The problem here is that the third person singular has the same morphology as the first person singular, so it is not clear who is talking in this part, the boy or the girl.

In fact, there is one personal pronoun in the whole text working as a subject in "Y **ella** decía que a mí me molestaba que llorara tanto", which clearly marks it is the girl (she) who performs the action (decía). The other verbs do not have an explicit subject, this being indicated with their morphology. The context makes it clear that the subject of the first verbs ("abría, encontraba, entendía") is "I/yo". However, after the personal pronoun "ella" appears in the following line, the subsequent verbs ("veía, sentía, molestaba") do not tell us exactly who is talking; the girl (as in the original text) or the boy. We might get the impression that it is the boy who is speaking in the last two lines: "(yo) Veía que nos estaba afectando y que le afectaba a su madre. (yo) Sentía que (yo) molestaba a todo el mundo". But this is the opposite of what the original text says, since the subject in the English quotation is "she", the girl, who complains about the whole situation of being a mom. Therefore, the semantic meaning of this quotation is not fully transferred or at least it becomes very confusing in the target text, and the result is no rendition of the original message: either in terms of prosodic-illocutionary force or of semantic meaning.

5.2.2.2.1 Concluding remarks on the rendition of quotations

After analysing all the examples that show a quotation in the source English text and how these have been rendered in Spanish, two options have become clear: translating the quote as a quote in the target text, or paraphrasing it. However, we have noticed that there is no rule of thumb to identify when translators prefer one or the other, since both strategies seem to be used instinctively. In any case, keeping the quote in the target text keeps the prosodic-illocutionary force of the original given the fact that quotations are considered demonstrations, a function which would be in the target text if the quotation appears as such in the translation. On the other hand, paraphrasing the quote modifies that original communicative function turning it into a description, which is not the original speaker's intention.

5.2.3 Tonality

This section is devoted to analysing all the variables related to tonality, or the division of spoken material into chunks or IPs, as stated in section 2.2.2.1 and 2.5.1, which are the following: subject vs predicate division, defining/non-defining relative clauses, topics, cleft sentences, and tag questions. Of these initial five variables, we did no find any relevant example of defining/non-defining relative clauses in our copus. Therefore, the other four variables will be examined here.

5.2.3.1 Subject - predicate division

In our corpus, have been identified 6 sentences in which the speakers resort to this tonality strategy in order to highlight the subject of their statements. If we recall the grammatical function of intonation (see section 2.2.2.1), English native speakers separate the subject from the predicate using two IPs, so that the subject gets the hearer's attention.

The pie chart below shows the total number of examples for this tonality variable and their rendition in the Spanish target text.
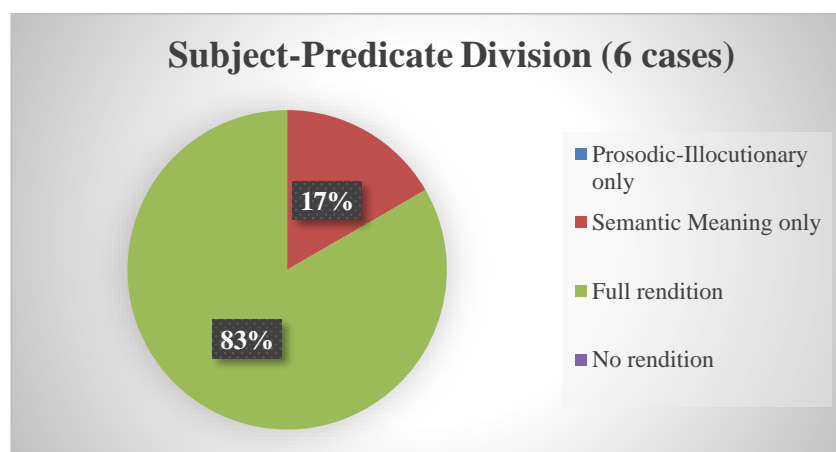

Figure 29 Rendition of subject-predicate examples

As can be seen in the chart, 83% of the examples render the original message (both in semantic meaning and prosodic-illocutionary force) in full; while 17% do so just partially (semantic meaning only).

Regarding the strategies used for this tonality function in English in the target texts, two can be identified, as shown in the bar chart below: division of subject-predicate as in original, and no division.
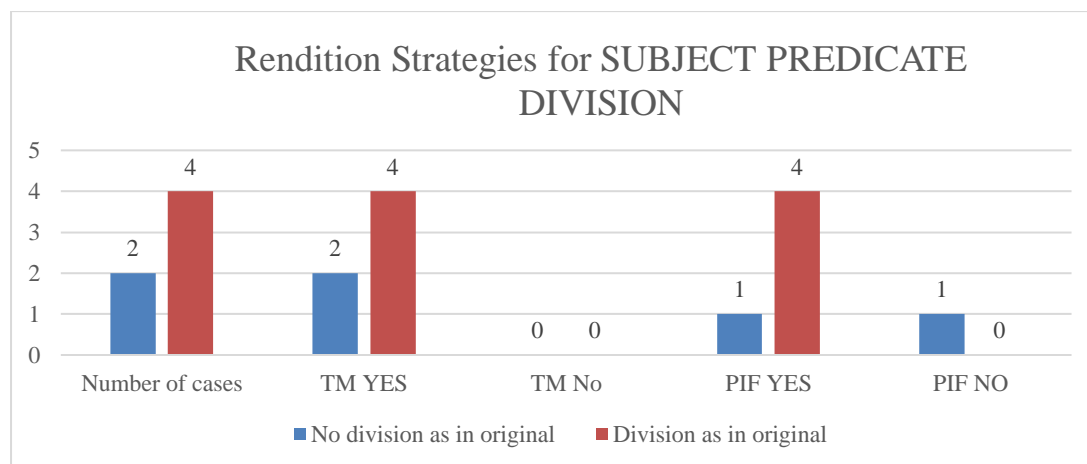

Figure 30 Rendition Strategies for Subject Predicate Division

Of the 6 examples analysed, 4 show division of subject and predicate in two IPs as in the original source text, which fully renders the semantic meaning and the prosodic-illocutionary force. On the other hand, there are 2 examples which use one IP instead of two to render this intonational feature in the target version. These last two examples render the basic semantic meaning of the original, but just one does so in terms of prosodic-illocutionary value.

The preferred strategies for this tonality feature according to each genre in our corpus is as follows.

DOCUMENTARIES			
Category	Number of Cases	Strategy	Number of cases
Subject VS Predicate Division	3	No división as in original	0
		Division as in original	3
INTERVIEWS VIDEOS			

Category	Number of Cases	Strategy	Number of cases
Subject VS Predicate Division	0		0
REALITY SHOWS			
Category	Number of Cases	Strategy	Number of cases
Subject VS Predicate Division	0		0
WEB VIDEOS			
Category	Number of Cases	Strategy	Number of cases
Subject VS Predicate Division3	3	No división as in original	2
		Division as in original	1

Table 27 Rendition strategies in Spanish for Subject VS Predicate division in English

The examples shown for this category include a vertical bar (|) to split the subject and the predicate both in the source and target texts. We will start our analysis of the 6 examples, starting with those that show full rendition of the original message, and finishing with the only example that shows partial rendition.

a) **Full rendition**

All the examples within the full rendition category show the same feature, which is that the translated lines have been delivered making the same pause as in the original source text. Therefore, there is subject-predicate division in the Spanish texts, which is a good choice to reflect the prosodic-illocutionary force of the source text, since the pause gives relevance to the subject of the phrase and also suspense to what comes next.

According to the RAE[72], the pause between subject and predicate in Spanish is usual and recommended when subjects are long, and this oral pause should not be marked with a comma on the text, as in the example below retrieved from the RAE website.

Example 72

Los alumnos que no hayan entregado el trabajo antes de la fecha fijada por el profesor ||
suspenderán la asignatura.

All the examples below show a short pause between subject and predicate, even though the subjects are not too long. These pauses can be seen on the tables displayed after each example, which show the pitch contour in both texts (source and target).

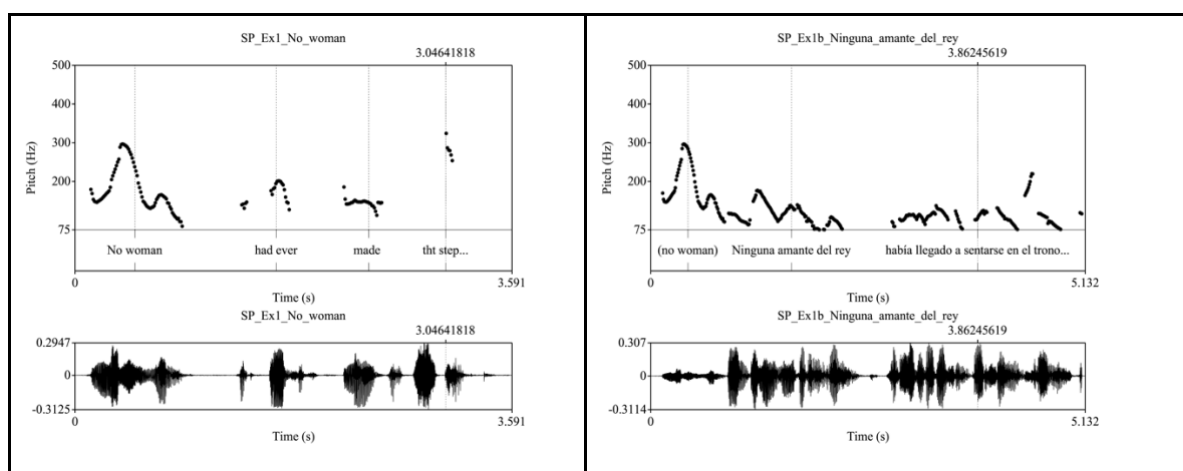| S.P.D.Ex 1[73] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:08:33	00:08:40	David Starkey	No woman	had ever made that step from royal mistress to the throne,	Ninguna amante del rey	había llegado a sentarse en el trono,

Image 31 Pitch contour in SPD. Ex1

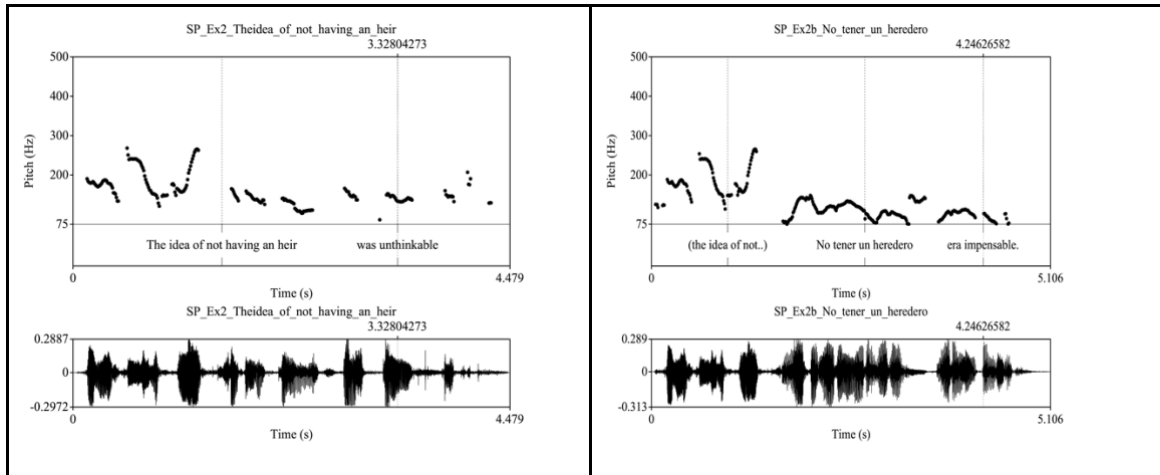| S.P.D.Ex 2 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:10:53	00:10:58	David Starkey	The idea of not having an heir | was unthinkable.	No tener un heredero | era impensable.

Image 32 Pitch contour in SP. Ex2

S.P.D.Ex 3 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:18:16	00:18:19	Hilary Mantel	But Anne | continued to be Anne.	Sin embargo Ana | era mucha Ana
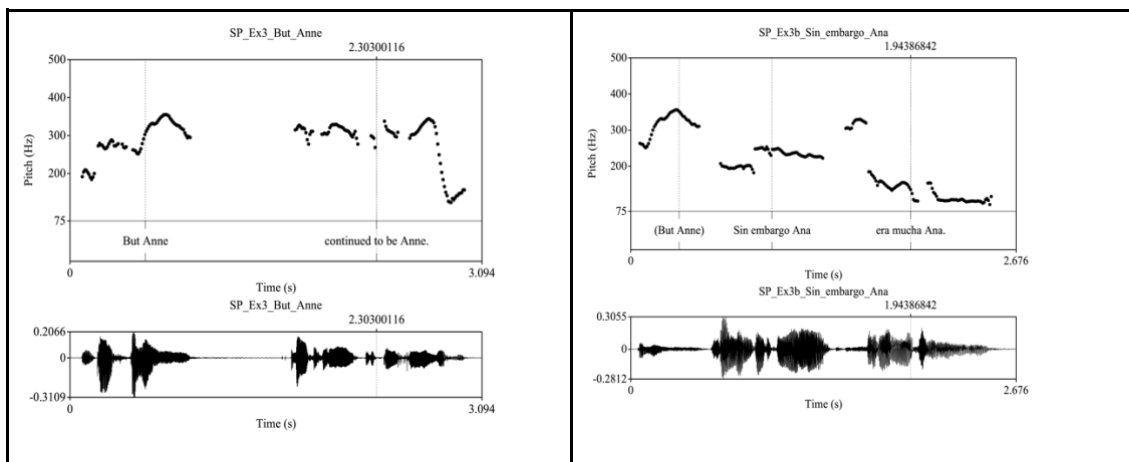
Image 33 Pitch contour in SP. Ex3

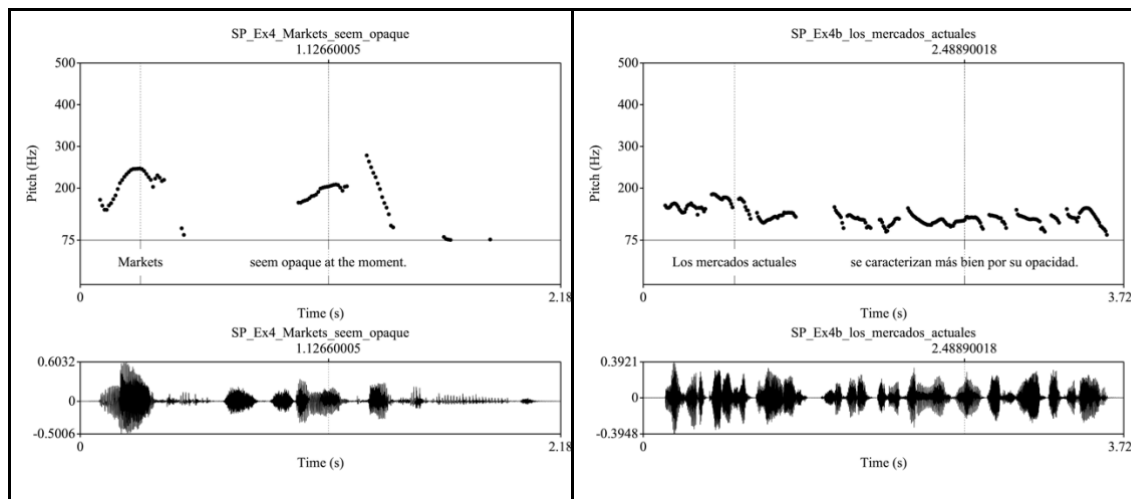| S.P.D.Ex 4 *TN Europe video Jan 17*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:12:00	00:18:00	Francis Ellison	Markets | seem opaque at the moment. They're posing questions when answers are pretty difficult to find.	Los mercados actuales | se caracterizan más bien por su opacidad, lo que nos plantea preguntas difíciles de contestar.	00:17:00	00:23:00

Image 34 Pitch contour in SP. Ex4

Considering the subjects of the Spanish texts above are not exceedingly long, the pauses shown could be removed and the meaning would be exactly the same. However, it is true that in order to highlight the subject and give it more emphasis and contrast as in example SP.Ex.3 "Sin embargo Ana | era mucha Ana", even suspense or drama, as in S.P.Ex2 "No tener un heredero | era impensable" the pause works perfectly. In addition, the literal translation used in these examples serves to keep literal synchrony and give the target text more credibility and realism.

b) Partial Rendition (semantic meaning only)

S.P.D.Ex 5 *TN Europe video Jan 17*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:49:00	00:57:00	Francis Ellison	Our job | is firstly to recognise that, and secondly to focus more on questions we can answer more easily and better.	Resulta esencial que reconozcamos primero este hecho y, en segundo lugar, debemos centrarnos en las cuestiones que podemos responder de forma más sencilla y contundente.	01:03 :00	01:12: 00

This example is the only example that shows partial rendition. As we can see, in the English source text, "our job" is the subject of the sentence "Our job | is firstly to recognise that". However, in the Spanish version, the subject has been removed and the text is a reformulation of the original. The translator has resorted to modulation in order to render the original meaning, and this is a very valid option. We have "resulta esencial", which works as an adverbial phrase to introduce the sentence. Here we do not have literal synchrony as a consequence of this modulation. In addition, there is neither pause after "resulta esencial" nor insistence stress in its delivery, which would have highlighted it as it is in the original text. Consequently, we consider that the rendition of the prosodic-illocutionary force is not fully achieved in the target text.

5.2.3.1.1 Concluding remarks on subject-predicate rendition

As we have seen, Subject-Predicate division is a feature of English tonality which can also be found in Spanish. Although in Spanish this feature is supposed to break the IP into two when the subject is too long, we have found that in voice-over translation it

appears also when subjects are short. In fact, it is used in most of the cases in which it happens in the English original text, which, in our view, is probably due to linguistic interference or synching reasons (to keep the isochrony) in order to imitate the original pace and structure.

If instead of having this subject-predicate break, we only had one IP for the whole sentence, the meaning would be exactly the same. The emphatic force of the subject would probably be lost, or not as highlighted as in the original (with a loss of the prosodic-illocutionary force) but in terms of transfer of basic meaning, it would be the same.

5.2.3.2 **Topicalisation**

As stated in section 2.2.2.1, topicalisation refers to placing an element other than the grammatical subject at the beginning of the sentence for emphatic purposes; this is, in addition, usually signalled with an intonation break (Wells: 2006: 199). There are three types of constructions that resort to topicalisation in English: topics, cleft sentences, and pseudo-cleft sentences, which will all be covered in this section.

5.2.3.2.1 **Topics**

There is only one example of topics in our corpus, belonging to the documentary *The last days of Anne Boleyn*. It clearly shows two IPs (as suggested by Wells), which can be seen in the image below of the pitch contour for this phrase. The first IP represents the topicalised element (*the survival of the dynasty*), which is thus emphasised; while the second element is the nominal relative clause characteristic of *wh-type* cleft sentences, which, under normal circumstances, would be placed at the beginning of the sentence since clefting is, per se, a way of emphasising in English. Nonetheless, since this is an example of topicalisation, we have a topicalised *wh-type* cleft sentence or pseudo-cleft.

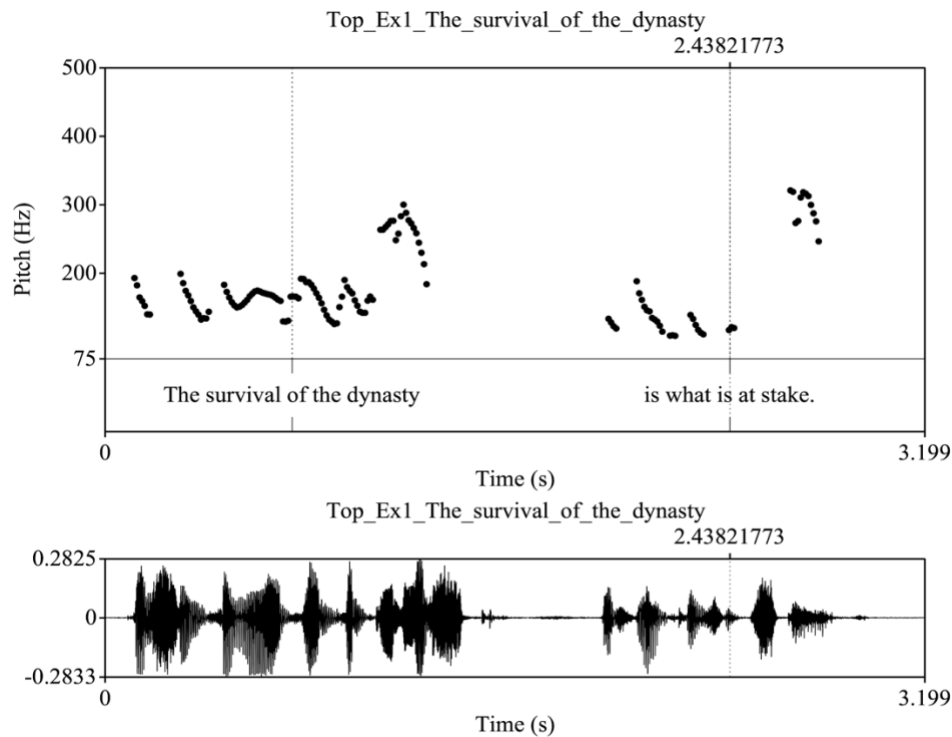| Top. Ex 1[74] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:11:11	00:11:14	David Starkey	<u>The survival of the dynasty</u> is what is at stake.	La supervivencia de la dinastía está en juego.

Image 35 Pitch contour and audio wave of Top.Ex1

The rendition of this sentence in the target Spanish text is almost a literal word by word translation. There is no topicalisation or use of a relative clause (cleft) to highlight the relevant information as in the source text. What we have instead, is a neutral sentence with a subject "La supervivencia de la dinastía", a verb "está" and an attribute "en juego". Although the subject in the target text coincides syntactically with the topicalised rheme in the source text ("the survival of the dynasty"), the fact that there is no relative clause in the Spanish version makes this sentence flat and neutral, thus not achieving the same pragmatic force as the English one.

[74] Top. Ex. Stands for Topic Example.

In the light of this information, we can say that this example of topicalisation in the source text has been rendered only partially in the target text. There is transfer of the basic semantic meaning and it makes sense. However, the prosodic-illocutionary force characteristic of this contrastive focus construction has not been rendered in the target text, since, as we have seen, there is neither topicalisation nor use of a relative construction to highlight the same information as the source text

5.2.3.2.2 Cleft-sentences

We have found 11 examples of cleft-sentences (including the *it-type* and the *wh-type* or pseudo-cleft) in our corpus, which show two different ways of rendition in Spanish, which are: using equational structures, as suggested by Braga & Rica Peromingo (2015), and literal translation without a relative clause (not following Braga & Rica Peromingo's suggestion). We should now recall that Braga & Rica Peromingo's strategy to translate emphasis in cleft-sentences can be found in section 2.4 of this dissertation, where the following example was provided to show their way of translating these constructions into Spanish.

Example 73

It is horror movies that I love: Las películas de miedo son las que me encantan / Son las películas de miedo las que me encantan.

The translation of this sentence shows two alternatives, the first of which is the one that Braga & Rica Peromingo consider the best of the two, which should be accompanied with some insistence stress.

The following charts show an overview of how the 11 examples of cleft-sentences in our corpus have been translated, what type of rendition the target texts show and what strategy has been used in each case.
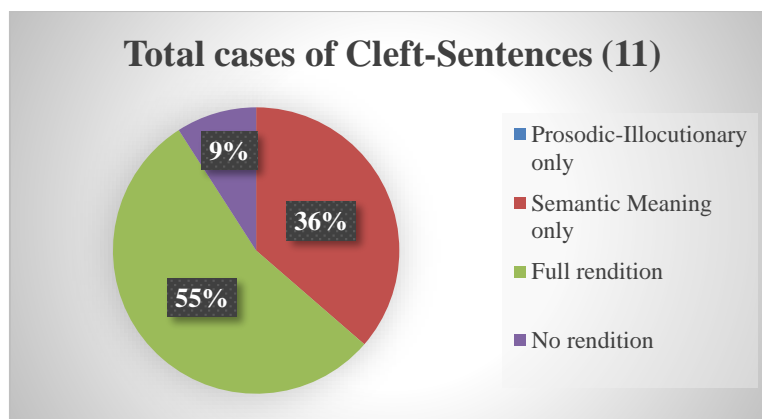
Figure 31 Pie chart showing the rendition possibilities for cleft-sentences.
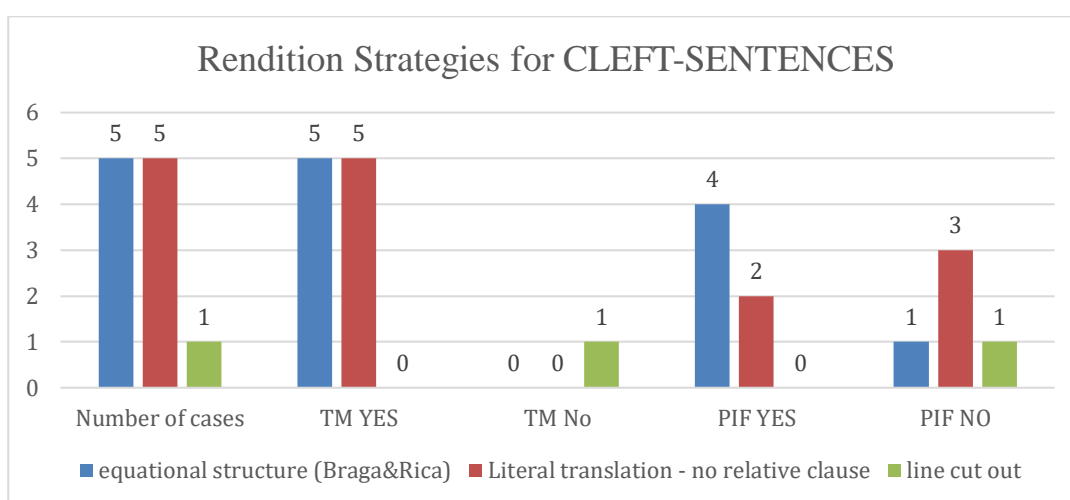

Figure 32 Bar chart showing the rendition strategies for cleft-sentences.

As we can see, there are eleven examples of cleft-sentences in our corpus with 55% showing full rendition, 36% partial rendition (semantic meaning only), and just 9% which show no rendition at all. Five of the clefts have been translated literally without including a relative clause resulting in all of them showing transfer of semantic meaning and two rendering the prosodic-illocutionary force. Five examples have been rendered using equational structures and all but one show full rendition of the original message (both semantic and prosodic-illocutionary force). Finally, there is just one example which has been cut out in the target version; the video part for that line has been edited and cut out, so there is no translation for this cleft in the Spanish video.

The following table show how the different types of programmes have rendered cleft-sentences.

DOCUMENTARIES			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Cleft Sentences	8	Equational structure (Braga & Rica Peromingo)	5
		Literal translation – no relative clause	3
INTERVIEW VIDEOS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Cleft Sentences	0	Equational structure (Braga & Rica Peromingo)	0
		Literal translation – no relative clause	0
REALITY TV			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Cleft Sentences	1	Equational structure (Braga & Rica Peromingo)	0
		Literal translation – no relative clause	0
		Line cut out	1
WEB VIDEOS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Cleft Sentences	2	Equational structure (Braga & Rica Peromingo)	0
		Equational structure (Braga & Rica Peromingo)	2

Table 28 Cleft sentence rendition strategies in our corpus

Our analysis of the examples found in the corpus will start, as usual, with those showing full rendition.

a) Full rendition

C.S.Ex 1[75] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:39:50	00:39:55	Hilary Mantel	What he could do was put people under a bit of pressure,	Lo que sí pudo hacer, fue presionar a determinadas personas,

This example of cleft sentence (pseudo cleft *wh-type*) from the documentary *The last days of Anne Boleyn* shows Hilary Mantel talking about what Cromwell could have done in order to accelerate the end of Anne Boleyn. In the previous lines she remarks that she does not think that Cromwell thought up a step by step process which would result in Anne's death, so she uses this cleft sentence to contrast with a comment on what he could have done instead. We can see that the image of the pitch contour below shows, how she stresses the word "could", which is rendered in the target text using the emphatic particle "sí" together with the main verb ("lo que sí pudo hacer"). This emphatic particle helps render the prosodic-illocutionary force, the emphatic load of the stress in "could", and the contrast with the previous line.
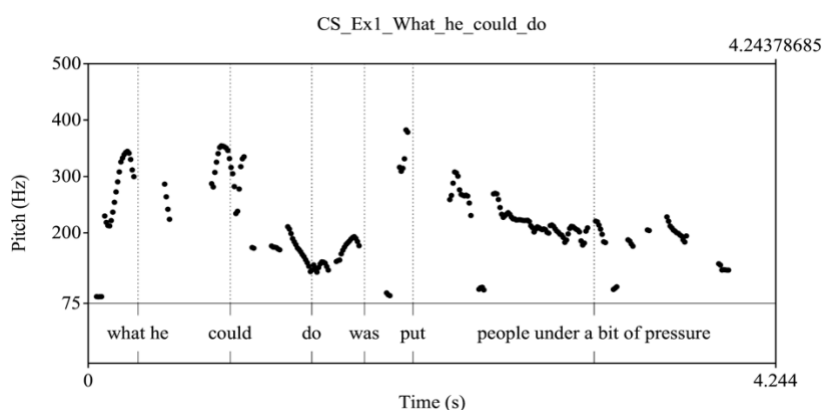

Image 36 Pitch contour in C.S.Ex1

---

[75] CS. EX stands for Cleft Sentence Example

The translation of the cleft is done following the structure of an equational construction, but placing the relative clause at the beginning and the nucleus at the end (end-focus), as in the source text. This structure works really well because it shows contrast and emphasis with the particle "sí" in the relative clause, thus giving more relevance to the information placed at the end of the sentence.

C.S.Ex 2 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:18:06	00:18:12	Hilary Mantel	What Henry looked for in a wife was one just like Catherine, please.	Lo que Enrique buscaba en una esposa, era justo lo que Catalina le ofrecía.

This example shows, again, Hilary Mantel commenting on Anne Boleyn's misfortune. She resorts to a pseudo cleft (*wh-type*) to discuss Henry's preference when looking for a wife. As can be seen in the picture below with the pitch contour, Hilary's phrase is divided into three IPs (What Henry looked for in a wife | was one just like Catherine | please). The informative load falls on the second IP, as is characteristic of pseudo clefts, which is preceded by a rising intonation in the first IP signaling more information is to come.
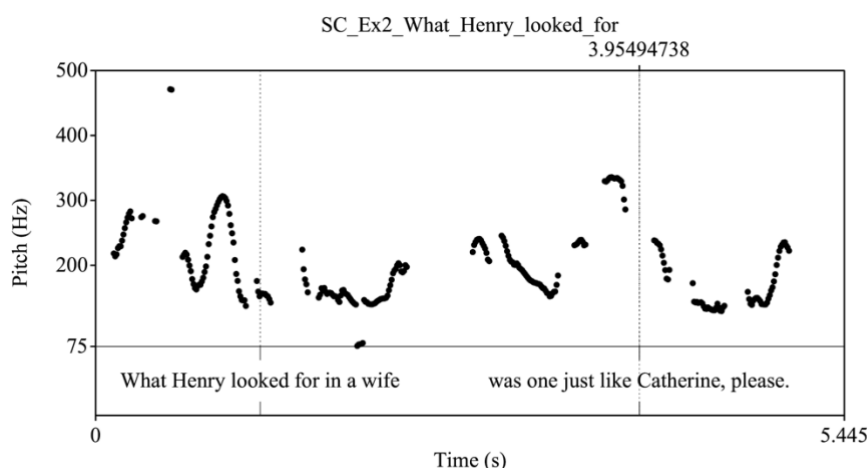

Image 37 Pitch contour in the English version of C.S.Ex2

The Spanish translation of this line shows two IPs, with the first one ending with rising intonation, which marks the boundary between the information that is known (the first IP) and new (the second IP) (RAE 2011: 483). This implies that the

prosodic-illocutionary function has been successfully delivered by the voice actor. The image below shows a graphic illustration of the pitch contour in the Spanish text.
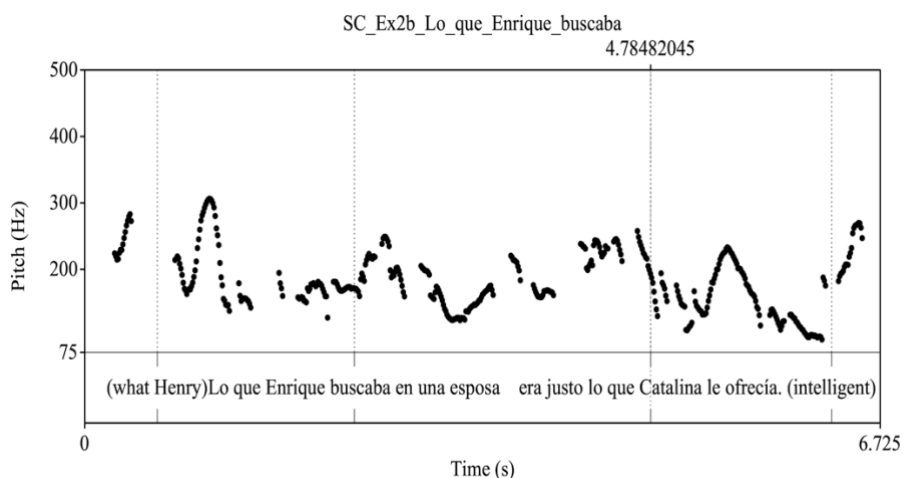

Image 38 Pitch contour in the Spanish version of C.SEx2

As we can see, the line starts with a sound bite "what Henry" followed by the first Spanish IP, which ends with rising intonation on "esposa". The second IP has a falling intonation until "ofrecía", after which we get another sound bite with a rising intonation "intelligent" that belongs to the sentence coming just after.

The translation of this example is, as in the previous example, an end-focus equational structure, which fully renders the semantic meaning and the prosodic-illocutionary force of the original.

C.S.Ex 3 *TN Philosophy*				
In	Out	Character	Source text	Target text
00:09:31	00:09:35	Collin Moore	but ultimately what we're really trying to achieve is the goals of our clients	No obstante, nuestro objetivo principal es cumplir los objetivos de nuestros clientes.
00:09:35	00:09:40	Collin Moore	it is their success that is our top priority and we need to reflect that	Su éxito es nuestra mayor prioridad, y así debemos reflejarlo.

This example has been found in the web video *TN Philosophy* and shows two cleft sentences: a pseudo-cleft (*wh-type*) first, and an *it-type* cleft last. The first of them shows full rendition of the original message, while the second shows only partial rendition (semantic only). In the first example, we can notice that the translator did not use the expression "lo que", which is the syntactic way to reproduce that type of emphasis in Spanish and has been used in the previous pseudo-cleft examples. Instead, we get a translation without a relative clause but with an adverbial phrase at the beginning ("no obstante") which works as an adversative coordinator, thus showing contrast between the subject ("nuestro objetivo principal") and the predicate ("es cumplir los objetivos de nuestros clientes").

Although the target text shows no subject-predicate division, as in the source text (see image below), probably due to the fact that the subject is not long, the illocutionary force of the original text is indeed rendered in the Spanish text thanks to the contrastive adversative conjunction.
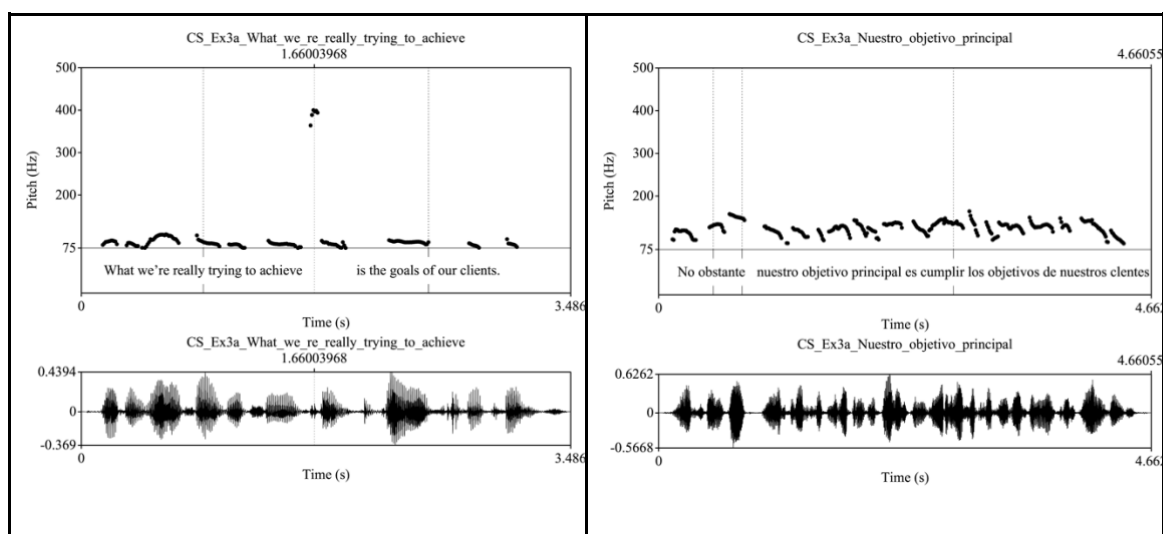

Image 39 Pitch contour in CS.Ex3 pseudo cleft.

The second cleft sentence in this example is an *it-type* cleft, of which, surprisingly, there are only two in our corpus. We can notice that the original cleft is divided into two IPs "it is their success | that is our top priority". Although there is a slight pause for breathing after "it is", we understand it does not make an IP on its own and functions as the pre-head in the first constituent of the cleft. The translation of this

cleft sentence has been done without including the relative construction that characterises these types of sentences or an adverbial phrase or connector to mark contrast or emphasis. The table below shows both options how this would have been done following Braga & Rica Peromingo's model to render cleft-sentences in Spanish.

Original Cleft	It is their success that is our top priority and we need to reflect that
Spanish Translation	Su éxito es nuestra mayor prioridad, y así debemos reflejarlo.
Spanish Translation equational structure - Braga & Rica Peromingo	Su éxito es lo que constituye nuestra mayor prioridad, y así debemos reflejarlo.

The Spanish translation shows two IPs just like the original one, showing, once more, subject-predicate division when subjects are short. This has probably been done to highlight the subject as in the source English text, since it is the topicalised constituent that gets the attention of the hearer. There is a subtle rising emphatic tone on the tonic syllable of "**é**xito" which falls on the subsequent syllables, just as stated in section 5.2.3.3a, and by the RAE (2011: 484). The picture below shows the pitch contour for this sentence.
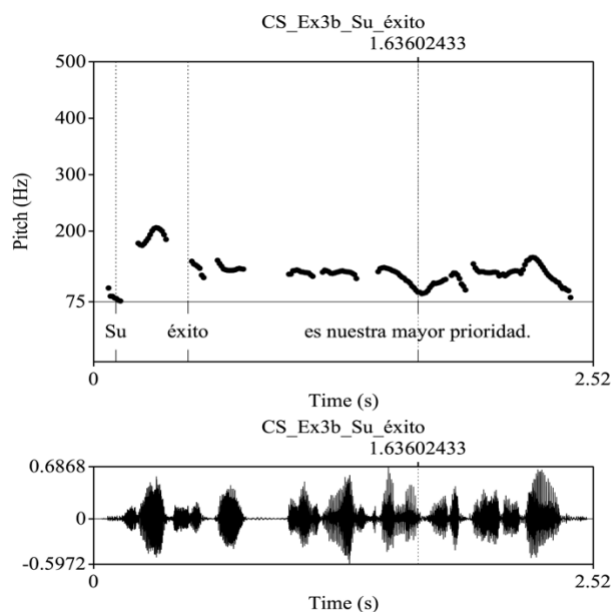

Image 40 Pitch contour for example C.S. Ex3 *it-type* cleft Spanish version

C.S.Ex 4 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:44:33	00:44:36	Suzannah Lipscomb	<u>What I think it is, is a game of courtly love gone wrong</u>.	<u>En mi opinión</u>, fue una historia de amor cortesano que salió mal.

This is the last example of full rendition within the clef-sentences section and has been found in the documentary *The last days of Anne Boleyn.* As can be seen in the source text, the speaker utters a pseudo-cleft (*wh-type*) in which the relevant part, the one in focus, in its second constituent "a game of courtly love gone wrong", while the first clause ("what I think it is") works as an introductory and contrastive clause to what comes next. If we now observe the Spanish translation for this line, we will notice that the translator has not used an equational structure, as suggested by Braga & Rica Peromingo (2015), but a literal translation without relative construction and modulation in the first clause ("En mi opinion"). In our view, this strategy works really well because, although there is no presence of the relative clause that characterises these contrastive constructions in Spanish, the idiomatic element "En mi opinion" functions as a contrastive introductory clause, which tells us that the relevant information is yet to come in the next clause ("fue una historia de amor cortesano que salió mal"). In addition, this idiomatic expression sounds more natural than a relative construction such as "lo que pienso es que" or "lo que creo es que", which would have been a bit long and alliterative. Therefore, this translation fully renders the semantic meaning and the prosodic-illocutionary force of the original text.

b) Partial rendition (Semantic Meaning Only)

C.S.Ex 5 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:36:15	00:36:22	David Starkey	It's precisely the magnitude of the charges that makes them convincing.	Es precisamente la magnitud de los cargos, lo que los hace más convincentes.

This is the second example of *it-type* cleft sentence in our corpus together with C.S.Ex3. It has been retrieved from the documentary *The last days of Anne Boleyn*, and shows David Starkey talking about Henry VIII's plan to get rid of his wife. The original line is split in three IPs "it's precisely | the magnitude of the charges | that makes them convincing". The first two IPs are separated by a short pause in which the speaker probably thinks the words he is about to say next, which, in addition, gives some suspense to his speech. We can see this in the image below showing the pitch contour for this example:
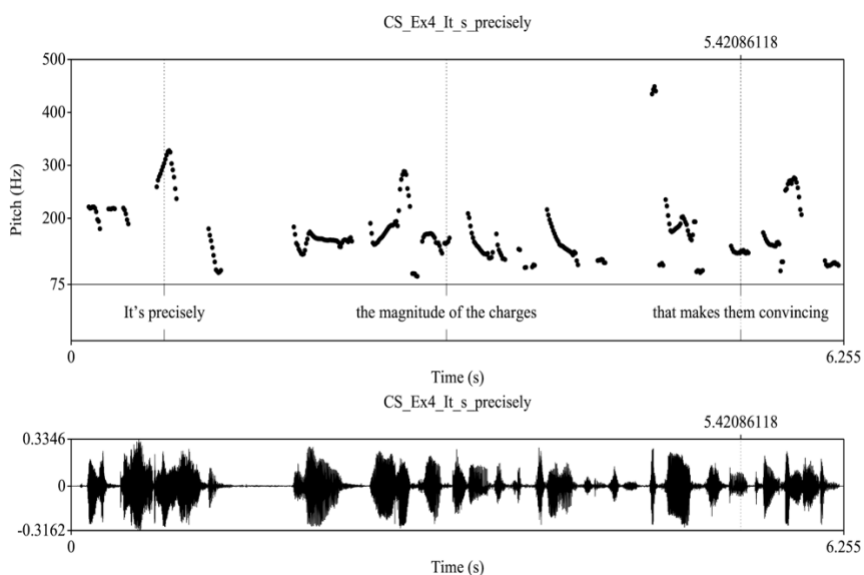

Image 41 Pitch contour and sound wave in C.S.Ex5

The Spanish translation of this cleft sentence has been done using an equational structure, as suggested by Braga & Rica Peromingo (2015). However, this time, the translation starts with the verb "es" rather than with the subject ("la magnitud de los cargos"). Although this sentence is grammatically correct, the best way to render the emphatic purpose of the cleft would have been with the subject in theme position, as seen in section 2.4. Therefore, we can say that the prosodic-illocutionary force has not been rendered with enough emphasis this time. A possible reason for this can be the sound bite we hear at the beginning ("precisely") and the need or will to keep literal synchrony at the beginning of the line. However, to keep literal sync and keep the emphasis following Braga & Rica Peromingo's model, we could say "Precisamente, la magnitud de los cargos es lo que los hace más convincentes".

C.S.Ex 6 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:40:13	00:40:19	Hilary Mantel	I don't think there was a pre-arranged,	No creo que hubiera una conspiración
00:40:19	00:40:23		highly intricate conspiracy.	extremadamente compleja detrás.
00:40:23	00:40:27		<u>What I **think** happened was</u> a series of events,	<u>Lo que pienso es que se</u> produjeron una serie de acontecimientos
00:40:27	00:40:33		which spiralled out of control, took everyone by surprise.	que fueron desarrollándose de tal forma que pillaron a todos por sorpresa.

This last example of partial rendition, taken from the documentary *The last days of Anne Boleyn*, shows a cleft sentence with narrow focus in which the prosodic-illocutionary force has not been successfully rendered, in our view, in the target Spanish text. In it, Hilary Mantel reflects on the circumstances that led to Anne Boleyn being eventually executed. The example shows a pseudo-cleft (wh-type) with three IPs as can be seen in the image below.
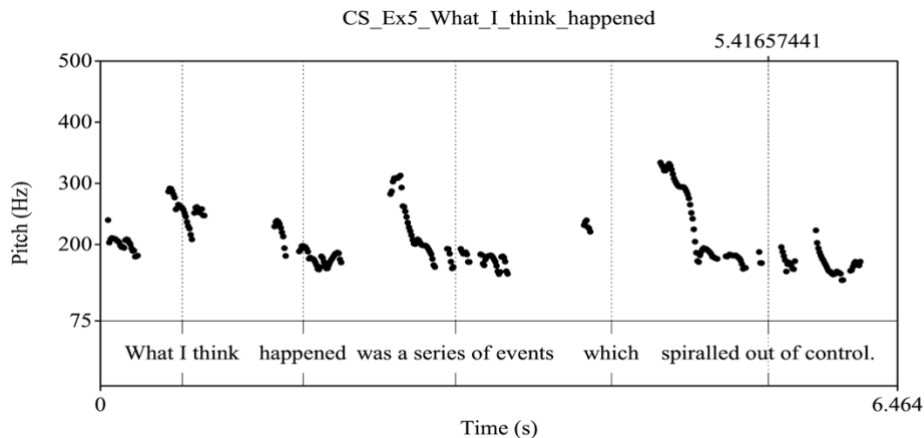
Image 42 Pitch contour in C.S.Ex6

The IP distribution is the following "What I think happened | was a series of events | which spiralled out of control" and we can clearly see in the image that the nuclei for each IP are "think", "series", and "spiralled". There is narrow focus in the first it IP, "What I **think** happened", due to the fact that in the previous line Hilary Mantel gives us her thoughts on what she does *not* think that took place (concerning Anne being killed); and she now gives us a pseudo-cleft with narrow focus on the first constituent to emphasise what she thinks indeed happened for Henry to get rid of his second wife. Again, we have two intonation variables combined in one example: narrow focus and a cleft-sentence.

If we examine the Spanish translation for this passage, we can notice that there is no emphasis in the delivery of the script by the voice actor to mark this prosodic-illocutionary load that Hilary Mantel is giving the audience in the source text. In addition, there is no syntactic particle in the text particle to mark the emphasis, as could have been "lo que sí pienso es que…", which would have been very emphatic (probably more than the original), or a construction such as "en mi opinion" (in my opinion), which sounds more natural in Spanish than "lo que pienso". Therefore, despite the fact that the transfer of semantic meaning in this passage is suitable and the equational structure renders the contrast of the cleft, the rendition of the prosodic-illocutionary force implicit in the narrow focus has not been satisfactorily rendered, and thus, we consider this passage has been rendered only partially (semantic meaning only).

c) No rendition

C.S.Ex 7 *Kitchen Nightmares*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
0:42:52	0:42:58	Gordon Ramsay	<u>What is going on upstairs is</u> actually brilliant. It's up to you now not to allow it to go back to where it was.	Ahora depende de ti no permitir que vuelva a donde estaba.	0:42:12	0:42:15

This is the only example of no rendition of a cleft sentence in our corpus, and has been found in the reality show *Kitchen Nightmares.* The reason for it not being rendered is that the target version of the programme has been edited and this very line has been removed. This is remarkable, because the cleft removed is significant in the scene, so it should have been included. But we know that other parts of the programme in the target Spanish version have also been cut out, since, if we compare the timings in the original and target versions, they do not match.

The cleft in this example serves as an introduction to the information that comes next, which is Gordon Ramsey encouraging the lady, the owner of the restaurant, to keep things as they are now so that the restaurant can be viable and successful. The cleft itself, "What is going on upstairs is actually brilliant" gives us the contrast between what happens in the kitchen now ("brilliant"), and what happened there in the past. Then we get Ramsey's advice. This is why we understand that leaving the cleft out in the translation does not make sense, because opening the scene with the advice and no context makes it hard to understand. In the Spanish version, we go from a scene in the kitchen, in which Ramsey is paying compliments to the chef for his great work, to a straight piece of advice to the owner of the restaurant ("Ahora depende de ti no permitir que vuelva donde estaba"). This advice comes out of the blue, out of context.

We do not know whether the video was edited before or after it was transcribed and translated, but we understand that this line should have been kept for

contextualization. This is an example of the *funnel effect* mentioned in section 3.1.3 by which translations can be altered due to the several processes they undergo (editing of video, transcription, translation, adaptation, etc.).

5.2.3.2.3 Concluding remarks on topicalisation rendition

Topicalisation and clefting are tonality resources which help the speaker highlight information in different ways. Topics and *it-type* cleft sentences place the relevant information at the beginning of the sentence, while pseudo-clefts (*wh-type*) do so the other way around, with the focused material at the end. All three resources are used to show contrast and are pragmatically and illocutionary pertinent, therefore relevant for translation. We found that there are two main strategies to translate these types of constructions; one is using an equational structure following Braga & Rica Peromingo's (2015) suggestion, the other is by means of a literal (almost word by word) translation without a relative clause. In addition to these strategies, the use of emphatic or insistence stress, or other particles, should be present to mark the contrast intended in the original construction. As we could see, almost all the examples identified in our corpus have been successfully rendered, with the exception of some that do not render the prosodic-illocutionary force due to lack of emphasis in the delivery or the absence of an emphatic particle or expression to show contrast.

5.2.4 Tone

In this section we will deal with the translation of tone in Spanish voice-over and it is divided into three subsections: falling tone, rising tone and tag questions. Tag questions can have either a falling or rising tone and, therefore, their tone direction has been included in the fall and rise tone sections for statistical purposes. However, they will have their own analysis section given the fact that they constitute a common and relevant feature of English oral discourse.

5.2.4.1 **Falling tone**

We have found 24 examples of falling tone worth commenting in our corpus. From these, 8 belong to tag questions, so these will be left aside and commented further on in our analysis. Nonetheless, the translation strategy used to render these tags is, in fact, included here. With this in mind, and as stated on the pie chart below, from the 24 examples of falling tone, 54% show full rendition, 42% partial rendition (semantic meaning only) and 4% no rendition at all.
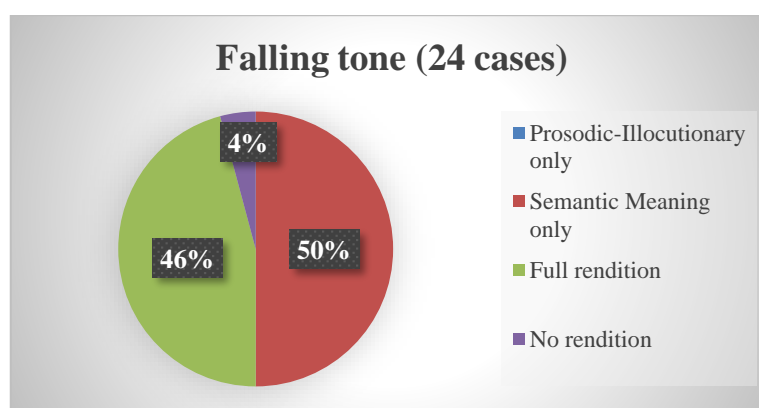

Figure 33 Pie chart with rendition results for falling tone

There are several strategies used to translate these examples of falling tone and they show different results in terms of rendition of semantic meaning and prosodic-illocutionary force, as can be seen in the bar chart below. Thus, out of the 24 examples, 5 are falling tags translated as questions, 3 are falling tags not translated as questions, 4 are declarative sentences rendered as questions, 3 are exclamative sentences translated as questions, 3 are declarative falls rendered by means of dubbitis (using a weird rising intonation in the delivery), 5 are emphatic high-falls that show flat intonation or no emphatic particle, and finally there is 1 question with falling intonation translated as a declarative sentence.
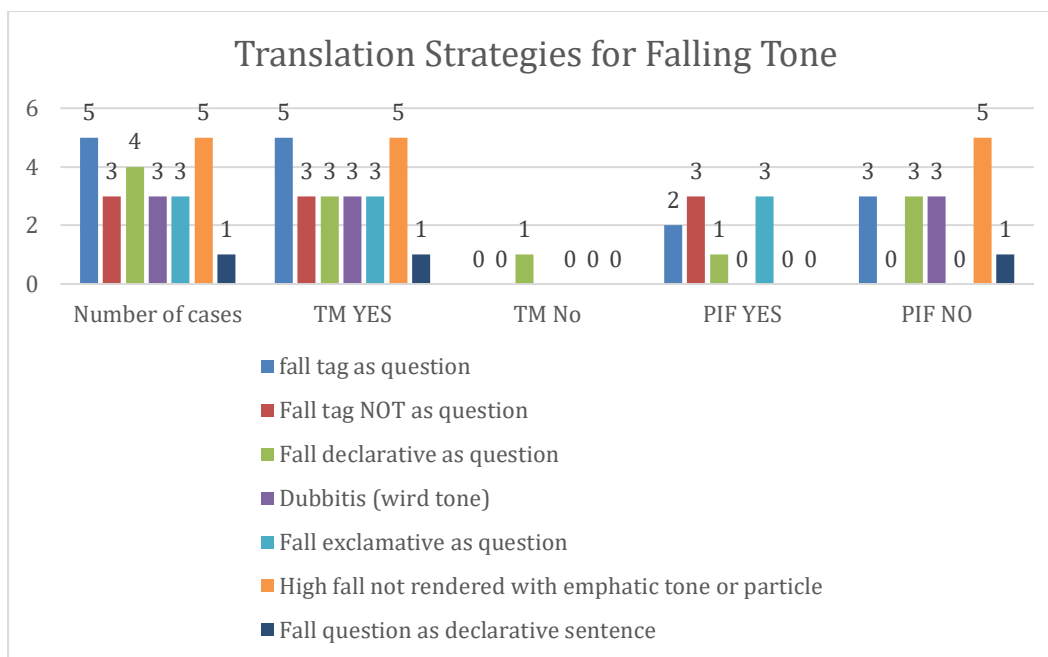
Figure 34 Rendition strategies for falling tones

The strategies are varied and show different rendition possibilities. Thus, fall declaratives translated as questions show no rendition of the prosodic-illocutionary force in 3 out of 4 examples, and surprisingly, fall exclamatives translated as questions show full rendition of the original.

In terms of genre, the number of examples and the preferred strategies used to translate falling tones are shown in the following table.

DOCUMENTARIES			
Category	Number of Cases	Strategy	Number of cases
Fall	11	fall tag as question	3
		fall tone as question	0
		Fall tone (exclamative) as question	1
		Dubbitis (weird tone)	3
		High fall not rendered with emphatic tone or particle	4

		INTERVIEWS VIDEOS	
Category	**Number of Cases**	**Strategy**	**Number of cases**
Fall	2	fall tag as question	1
		fall tone as question	0
		No tone or emphatic particle to match original	0
		High fall not rendered with emphatic tone or particle	1
		REALITY TV	
Category	**Number of Cases**	**Strategy**	**Number of cases**
Fall	10	fall tag question NOT translated as question	3
		fall tone (declarative) as question	4
		fall tone (exclamative) as question	2
		Fall tag as question	1
		WEB VIDEOS	
Category	**Number of Cases**	**Strategy**	**Number of cases**
Fall	1	Fall question translated as a declarative sentence	1

Table 29 Examples of falling intonation and translation strategies per genre

The table above shows that there are two genres that show falling intonation (exclamatives) translated as questions or using a rising intonation: documentaries and reality shows.

Let us now delve into the analysis of the most significant examples for each category: full rendition, partial rendition and no rendition.

a) Full rendition

From the 11 examples of full rendition found in our corpus, 8 belong to falling tag questions, which leaves us just 3 examples of falling tones successfully rendered in terms of semantic meaning and prosodic-illocutionary force. Moreover, these 3 full rendition samples share the same translation strategy since they show declarative and exclamatory statements rendered as interrogative sentences. And, in addition, the three belong to the same genre (reality TV); in fact, they belong to the very same programme: *Kitchen Nightmares*.

F.T.Ex 1[76] *Kitchen Nightmares*						
ST in	ST out	Character	Source Text	Target text	TT in	TT out
00:01:50	00:01: 50	Gordon Ramsay	Two customers the whole \night.	¿Dos clientes en toda la /noche?	00:01.35	00:01.35

This example shows a declarative sentence with a high fall nuclear tone on the word "night", showing broad focus, translated as an interrogative sentence in the target text. The scene shows Ramsey talking to Sue Ray, the owner of the restaurant. If we check the whole scene, we see that Ramsey and Mrs Ray are talking about the number of customers the restaurant is going to have for lunch that day and then Ramsey asks about the night before. The whole scene is transcribed below.

Gordon Ramsay	How many are you having for lunch?	¿Cuántas personas han venido a comer?
Sue Ray	Two.	Dos.
Gordon Ramsay	Two? Not many.	¿Dos? ¿En serio?

[76] FT.Ex stands for Falling Tone Example

Sue Ray	That's it.	Nada más.
Gordon Ramsay	And last night?	¿Y anoche?
Sue Ray	Last night, I think we did two again, so…	Anoche creo que vinieron también sólo dos. Nada más.
Gordon Ramsay	<u>Two customers the whole night.</u>	<u>¿Dos clientes en toda la noche?</u>
Sue Ray	Two customers the whole night, yeah. Unfortunately, it's just out of death, absolute death.	Dos clientes en toda la noche, sí, desgraciadamente. Esto está totalmente muerto.

Ramsey shows his surprise when he knows that the night before only two customers showed up at the restaurant for dinner. The pitch contour for his line and the translation are shown in the table below: there is a clear fall on "night" and a rise in "toda la noche".
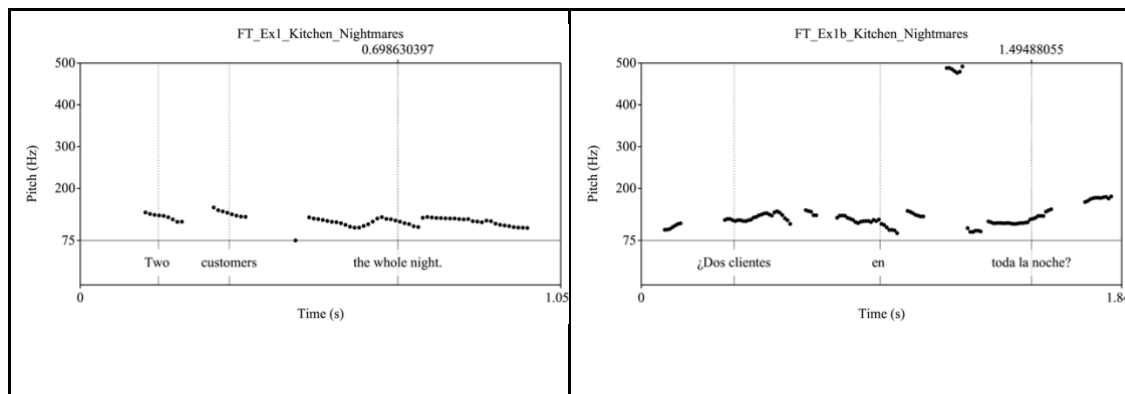

Image 43 Pitch contour in FT.Ex1

The Spanish translation for this sentence is, nevertheless, absolutely right. The translator resorts to an exclamative question or a non-neutral interrogative sentence, normally used when the goal is not to demand information but to express another pragmatic purpose (for example, surprise, as in this case). This type of construction is, for Vidal (1999: 67), used under one condition; the questioned fact has to do with the immediate factual situation and cannot be about previous information or knowledge shared by the speakers. However, the author understands that this construction is uttered with a circumflex intonation (rise-fall), while, for the RAE (2011: 469), the tone is a

rise, which coincides with the tone used in this example as illustrated in the table above. With this translation. The target text shows the same prosodic-illocutionary force as in the original text despite having changed the tone and the syntactic category of the sentence, since Ramsey is still showing his surprise and frustration about the situation of the restaurant he has to rescue.

F.T.Ex 2 *Kitchen Nightmares*						
ST in	ST out	Character	Source Text	Target text	TT in	TT out
00:06: 19	00:06: 21	Gordon Ramsay	It's burnt! You don't taste \that, man!	Están quemados. ¿/No los has probado?	00:06:0 3	00:06:0 6

This example shows a similar situation to that of the previous one; an exclamatory question with rising intonation in the translation which has been used to render an exclamatory question with falling intonation in the source text. In this scene Ramsey is telling off his pupil, Tim, who has just burnt the ravioli he intends to serve the customers. In his speech, Ramsey gives us two IPs, whose pitch contours are shown in the table below. The first IP | It's /\burnt |, shows a fall-rise tone on "burnt", which is the nucleus. The tone used here by the chef implies surprise and disapproval. The second IP | You don't taste \that man | shows a low rising head with a high fall nuclear tone on "that", and a low level tail on the vocative "man". The tone used in this IP is characteristic of exclamative sentences, particularly exclamatory questions that function as a yes-no question like this one (Leech and Svartvik 1994: 155), and serve to confirm the disapproval shown in the previous IP.
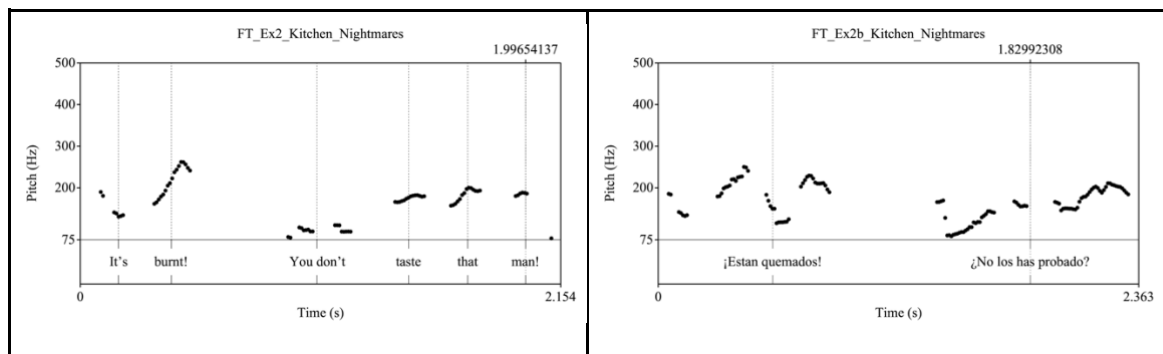
If we now analyse the Spanish text for this scene, the image above, on the right, shows that there are two IPs, as in the original English text. The first one, "| Están /\quemados |", shows a falling tone on "quemados". There is a rising on the tonic syllable "ma" and then an abrupt fall on "dos", characteristic of exclamative statements in Spanish when the speaker wants to show emotion and interest on the topic of the conversation (Alarcos 1994: 67). The second IP, "| /no los has probado |", shows rising intonation, typical of non-neutral Spanish interrogatives (as in FT.Ex1). In addition, we can see that the melodic line shows a steady rise, which for Fernández Ramíerz (1959: 257) is characteristic of reiterative and exclamative sentences. In fact, the translator has used a rhetorical question, more specifically, a *confirmative question* (Vidal 1999: 68), which has external negation[77]. From a communicative perspective, the purpose of this type of sentence is to make the receiver admit the presupposition stated in the question, therefore, in the end, to get confirmation. According to this, we may conclude that the rendering of the original exclamative high fall English tones, by means of a circumflex fall and a rising confirmative question in the Spanish target text, has managed to successfully render the semantic meaning and the prosodic-illocutionary force of the source text.

F.T.Ex 3 *Kitchen Nightmares*						
ST in	ST out	Character	Source Text	Target text	TT in	TT out
00:06: 23	00:06:27	Gordon Ramsay	\Smell it then. You \honestly can't taste that burn!	¡/\Huélelos! ¿/De verdad no otras que están quemados?	00:6:0 8	00:06:1 1

This third example, shows a similar structure to the previous one. The English source text shows two IPs with high fall exclamative tones: "| \smell it then | you \honestly can't taste that man |". The pitch contour shown below offers the tune direction in both IPs.

---

[77] External negation affects the whole content of the sentence, unlike internal negation, which is limited to certain elements within the sentence. (http://gramaticaespanola.com/publ/las_construcciones_sintacticas_fundamentales/la_negacion/negacion_externa_interna_a_distancia_y_contigua/52-1-0-863)
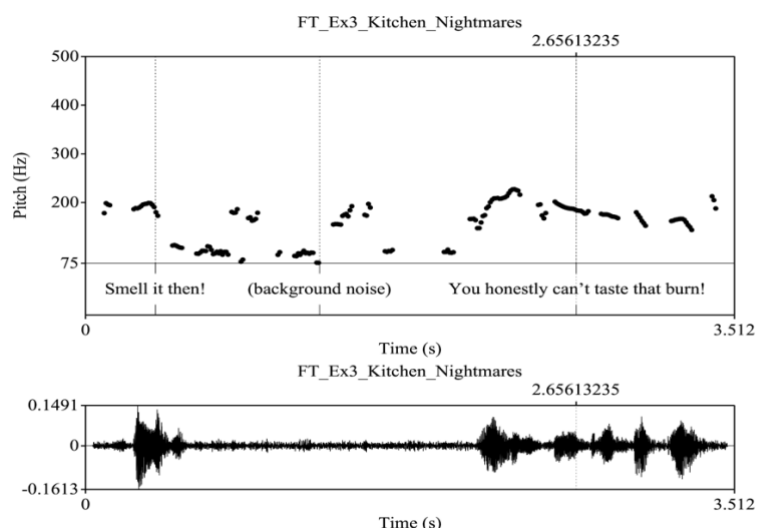
Image 45 Pitch contour in FT.Ex3

As can be seen, the second IP, another exclamatory question, has narrow focus with the nucleus on "honestly" since the rest of the information ("can't taste that burn") is considered given information, so it is deaccented as it comes from the previous lines in the scene (which belong to FT.Ex2, the example we have just seen).

The Spanish translation for this passage shows, as in example 2 above, two IPs: "| /\Huélelos | /de verdad no notas que están quemados |". The intonation used for these two IPs is exactly the same as in example 2; a falling circumflex exclamative in the first IP, and a steady rise for the second, characteristic of exclamative interrogatives.
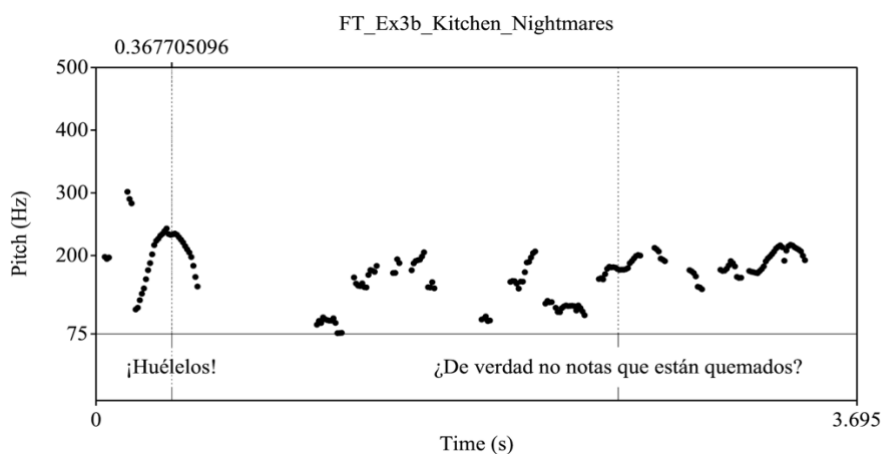

Image 46 Pitch contour in the Spanish translation of F.T.Ex3

With this in mind, we can say that the tone has been successfully rendered in

the target text, and the same happens with the prosodic-illocutionary force. As we saw, this IP shows narrow focus in the source text, being the nucleus on "honestly". Gordon Ramsey is, thus, showing surprise and enquiring whether the fact that Tim cannot taste the burn of the ravioli is either false (he knows but it does not bother him to serve burnt food to his diners) or true (as a chef he does not distinguish burnt food from non-burnt). This emphasis on the adverb, and therefore on Tim's honesty is rendered in the translation, despite the reaccenting of the old information ("no notas que están quemados") at the end of the sentence. The translator, by placing "de verdad" at the beginning of the sentence, makes it sound very idiomatic, and the illocutionary force is transferred with "de verdad" in initial position.

b) Partial rendition

F.T.Ex 4 *Tasteology: Experience*				
In	Out	Character	Source text	Target text
00:03:37	00:04:17	Christine Flynn	Soigne is the thing you hear over and over again in culinary school, and this idea of, like, 'make it nice, make it soigné' - and it just, it really did sort of epitomise the account, the idea of, again, taking these foods, like gas station **burritos** and elevating them to a soigne **plane** was very much like at the heart of what I wanted to do.	"Soigné" es un concepto que uno no deja de escuchar una y otra vez en la escuela de cocina, y esta idea de "hazlo bien, hazlo «soigné»", y la verdad es que es el ejemplo perfecto de la idea, nuevamente, de partir de estos alimentos, como /burritos de /gasolinera, y elevarlos al plano /**"soigné".** Eso era básicamente lo que quería hacer.

This example from the documentary *Tasteology: Experience* shows chef Christine Flyn, aka Jacques La Merde, commenting about her creations and a culinary concept that is nowadays a buzzword in the culinary business: *soigne*. In her speech she provides three IPs with three high fall nuclear tones, which have been rendered using

some sort of uptalk[78] in Spanish, a melodic pattern, which, in our view, does not belong to the Spanish language. The IPs identified and analysed, together with their corresponding pitch contour are shown below.

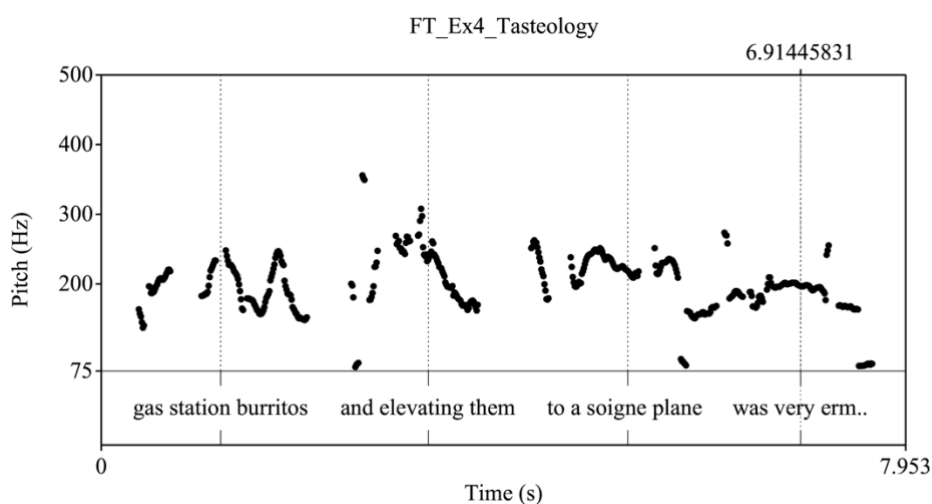English text: | gas station \burritos | and \elevating them | to a soigne \plane |

Image 47 Pitch contour in FT.Ex4

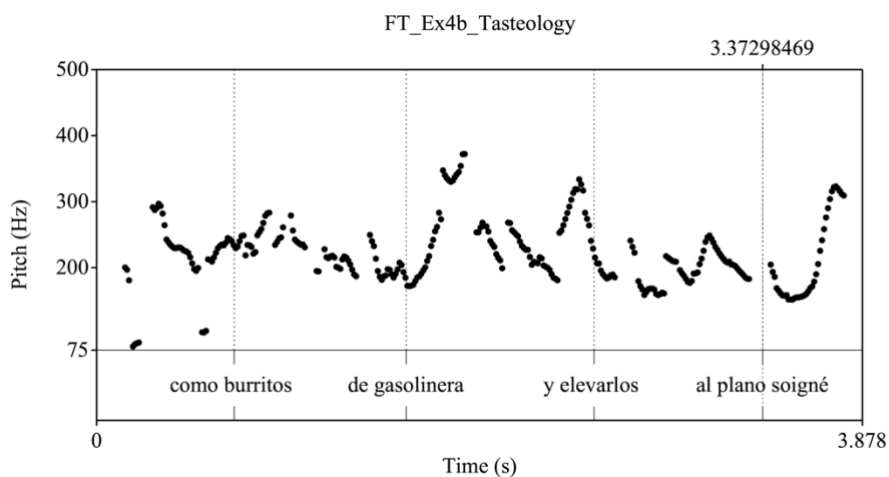Spanish text: | como /burritos de gasolinera | y elevarlos al plano /soigne |

Image 48 Pitch contour for the translated version of F.T.Ex4

As can be seen on the images above and heard in the audio and video files, the

delivery of the translated text uses a strange tone pattern in Spanish, which, moreover, does not reflect the meaning of the intonation used in the original text. In the source text Flynn is here talking about cooking burritos and she is making up what she says as she goes along. In each IP she uses a high fall tone and broad focus, so the information she gives is always new. In addition, the last IP ends with a lingering intonation followed by a new IP "was very erm…" which implies that more information is to come. This is not shown in the target text, in which, in addition to the rising intonation, the sentence structure shows as if it was incomplete. It may have been this structure that led to the intonation provided in the delivery.

Here is the whole sequence of the sentence:

"Soigné" es un concepto que uno no deja de escuchar una y otra vez en la escuela de cocina, y esta idea de "hazlo bien, hazlo «soigné»", y la verdad es que es el ejemplo perfecto de la idea, nuevamente, de partir de estos alimentos, como burritos de gasolinera, y elevarlos al plano "soigné". Eso era básicamente lo que quería hacer."

We understand that this sentence is, firstly, too long, since we have the three coordinated phrases linked by "y" and the use of juxtaposition to split the information, which can be tricky for the voice actor; secondly, one is not sure what is actually meant because the translation is too literal and it sounds as if the line was incomplete. The result of what this line says is given in the next one "Eso era básicamente lo que quería hacer", which is not an independent sentence in the original text. It is probably the splitting of this sentence at the end, that leads the voice actor to use that tone in her delivery; in our view this is not suitable and does not render the prosodic-illocutionary force intended by the original speaker.

F.T.Ex 5 *What's it like for EU migrants in Britain amid UKIP's success?*				
In	Out	Character	Source text	Target text

00:03:52	00:04:04	Man 3	I specialise in housing benefits for all my tenants. Erm, and it's an amazing situation where people just turn up and sign on. It's \wrong.	Yo estoy especializado en el tema de ayudas a la vivienda. Es una situación increíble porque la gente solo tiene que venir y firmar y eso está \mal.

This second example of partial rendition from the Euronews programme *Reporter: What's it like for EU migrants in Britain amid UKIP's success?*, has a man being interviewed about the situation of EU migrants in the UK. To conclude his answer, he utters a sentence using a high fall nuclear tone which is full of intention. This sentence is made up of one IP: "| it's \wrong |"; in addition to the high fall tone, the speaker uses his body language to show his disapproval of what some immigrants do and reinforce even more the illocutionary force implicit in his intonation. The table below shows how the speaker frowns at the time of uttering this IP, so he is providing pragmatic value both by means of visuals and acoustics.

Image 49 Pitch contour and body language in F.T.Ex5

The Spanish translation provides a literal translation with flat intonation, which, in our view, does not render the same pragmatic meaning as the original. If we see the whole sequence of this scene and the pitch contour, shown below, we notice that, in the image on the right, corresponding to the Spanish translation, there are some words highlighted by means of intonation ("venir", "firmar"), while the last part ("y esto está mal") is rather flat. This could be due to the characteristic intonation of newscasters,

which sounds sing-song (Álvarez del Valle 2016: 108) and often, emphasise the wrong words.
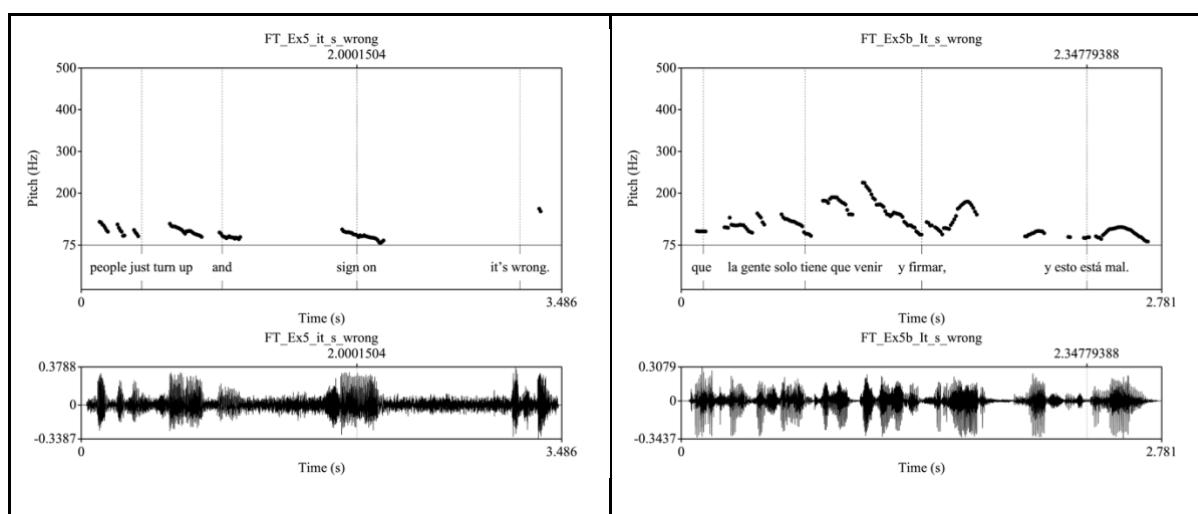

Image 50 Pitch contour in F.T.Ex5 "it's wrong"

One way in which the illocutionary force implicit in the original tone and the body language could have been rendered in the target text, would have been to include an emphatic particle in the text, such as *muy* or *realemente* ("y eso está **muy/realmente** mal"); or even an adjective with a more negative value, such as *fatal* ("y eso está **fatal**"). With this in mind, we can say that the translation of this passage renders the original message only partially, since the prosodic-illocutionary force has not been successfully rendered in the target text.

F.T.Ex 6 *Cupcake Wars*						
ST in	ST out	Character	Source Text	Target text	TT in	TT out
00:19:12	00:19:43	Janell Brown	Well, I like \Yoda. You said **Yoda.**	Y Yoda, ¿has dicho /**Yoda**?	00:11:17	00:11:45

This example from *Cupcake Wars* shows Janell Brown, one of the contestants, talking about Star Wars characters that can inspire her team in order to create a themed

cupcake for the contest. In her speech, we note she reaccents the word "Yoda" in the second sentence, which is a way of emphasising it again, focusing on it anew (Wells 2006: 178). The nuclear tone used in this second sentence, which makes one IP of its own is a high fall: "| you said \Yoda |", as illustrated in the pitch contour image shown below.
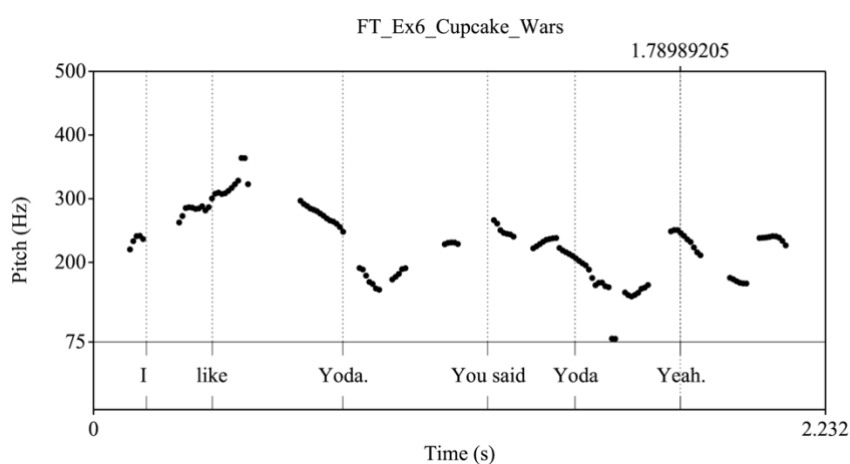

Image 51 Pitch contour in F.T.Ex6 Cupcake Wars

The second sentence ("You said Yoda") is a declarative statement, with which the speaker is confirming her first thought about using Yoda, which was suggested previously by her teammate. Thus, the speaker is affirming and supporting the idea that Yoda seems like a good option for a cupcake. In the video we can see she is talking face to face with her teammate and she is confirming her thoughts.

The Spanish translation for this passage shows a yes-no question for the second line, characterised by a steady rising intonation, as can be seen in the pitch contour below. This pitch movement is typical, as we have seen, of exclamatory questions, but here the aim is not to show surprise or give emotion, so an exclamation is not the goal of the original speaker. It is not a confirmative question either, as in example FT.Ex2, because this kind of construction demands a negative particle. What we have is a total interrogative sentence, which demand[79] a *yes* or *no* answer and is equivalent to English *yes-no* questions. In addition, these types of questions are also characterised by a steady rising intonation (Navarro 2001: 275).

---

79 Vidal (2016: 05) understands that "Toda oración interrogativa predetermina, por tanto, el tipo y la categoría del elemento que puede cerrar la proposición (...) en las totales, es *sí* o *no*."
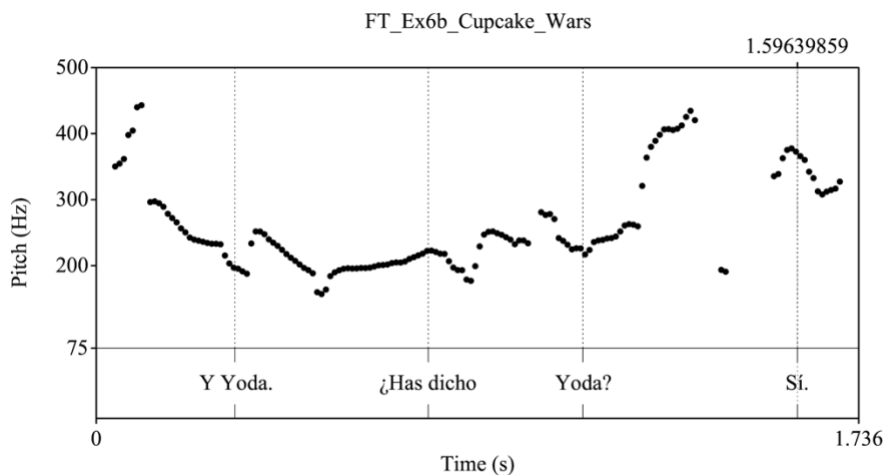
Image 52 Pitch contour in the translation of F.T.Ex6

The Spanish translation, thus, is requesting an answer, which, in our view, clashes with the speaker's original intention in the source text, which is confirming the idea of using Yoda as an inspiration for a cupcake. Consequently, the rendition of the prosodic-illocutionary force has not been successfully achieved, resulting in partial rendition of the original message.

F.T.Ex 7 *Tastelogy: Experience*				
In	Out	Character	Source text	Target text
00:12:53	00:13:34	Christine Flynn	There's that emotion and the idea of you're experiencing something... that is a little bit **magical**.	Es esa emoción y la idea de que estás experimentando algo... que es un poco /**mágico**.

The last example we will analyse in this section belongs to the documentary *Tasteology: Experience*, and shows, again, a scene in which chef Christine Flynne talks about how special the cooking and eating experience can be. In her speech, she concludes, with a mixture of emotion and passion, that the whole experience is magical, as stated in the lines of the original text shown in the example above.

The pitch contour image below shows two IPs: "| is a little bit | magical |", and

we can clearly see and hear that the first one has a level tone since the speaker has not yet finished talking, while there is a high fall tone on the second IP implying definiteness.
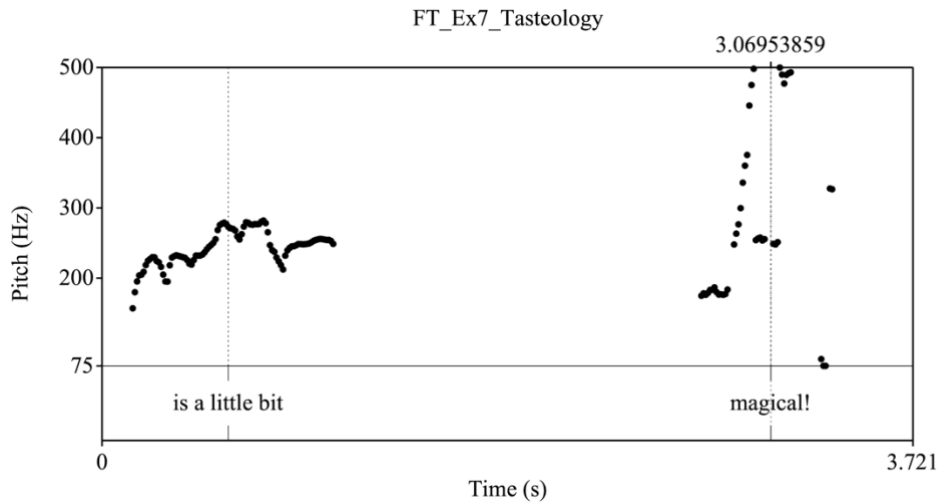

Image 53 F.T.Ex7

If we now check the Spanish translation of this line, we can see that it is divided into two IPs as the original: "| que es un poco | /mágico |", with the first IP ending with a level tone, as the speaker has not finished talking, and the second one with a rising tone, which implies, or at least gives the impression, that the speaker is not sure of what she says. The delivery of this last IP makes the character on screen sound dubious and hesitating in the target text, when the original speaker's intention in the source text is just the opposite.
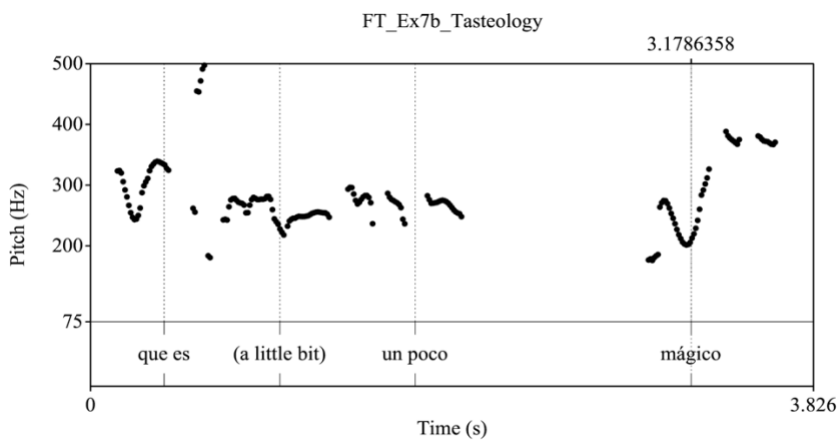

Image 54 Pitch contour in the Spanish version of F.T.Ex7

Let us see the tone used for this second IP in both texts in more detail. In the table below showing both pitch contours, we can see that in the original text (on the left) it goes as high as 500 Hz. This illustrates English speakers' use of broad pitch range in their everyday speech, as seen in section 2.5.2. This high pitch does not imply excitement, as would be the case in Spanish (in an over-excited example).
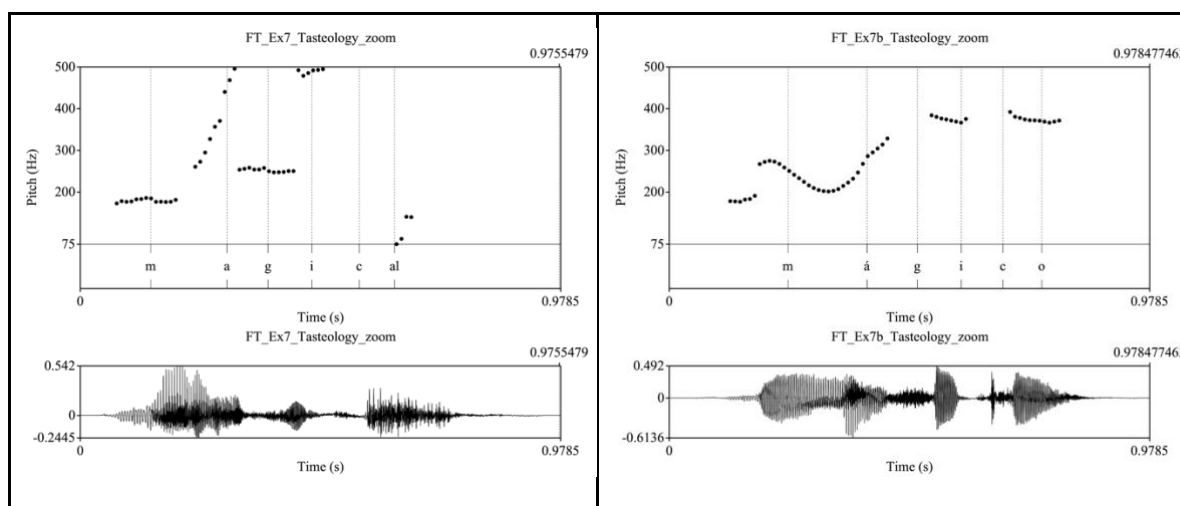

Image 55 Zoomed in pitch contour in FT.Ex7

The Spanish text, shown (on the right above) provides a similar pitch level, which is unusual in Spanish. It is probably due to the high pitch in the original text that has prompted the translator, or the voice actor, to deliver this line in such a way that it sounds both over excited and awkward, since the direction of the tone (rise) gives the impression that the speaker is doubting her own words. Therefore, we understand that the prosodic-illocutionary force for this passage has not been successfully rendered, as the pragmatic meaning of the tone in the target text is the opposite of the original.

F.T.Ex 8 *TN Philosophy*				
In	Out	Character	Source text	Target text
00:00:13	00:00:13	Colin Moore	\Do people understand that analysing what's happening in the South China Sea or Ukraine may affect investments in the US?	La gente entiende que lo que ocurre en el mar de China Meridional o en Ucrania puede afectar a las inversiones en EE. UU.

This last example within the partial rendition category that will be analysed, from the web video *TN Philosophy,* shows Colin Moore, Global Chief Investment Officer within Columbia Threadneedle, talking about the investment philosophy of the firm. At the beginning of his speech he asks a question, which is the example we are analysing here.

The sentence shows a yes-no question with a falling tone. Normally, the most typical intonation for yes-no questions in English is a rise, as we have already seen (Estebas-Vilaplana 2014b: 248; Wells 2006: 45). However, the intonation used this time is a fall, illustrated in the picture below, which is more typical of *Wh questions.* In English, when the structure of the sentence implies it is an interrogative one, with an auxiliary verb and an infinitive (as shown in the table below), the tone does not have to rise, as is the case with declaratives turned into questions (Granato 2014: 219). This sentence clearly shows a *yes-no* interrogative structure:

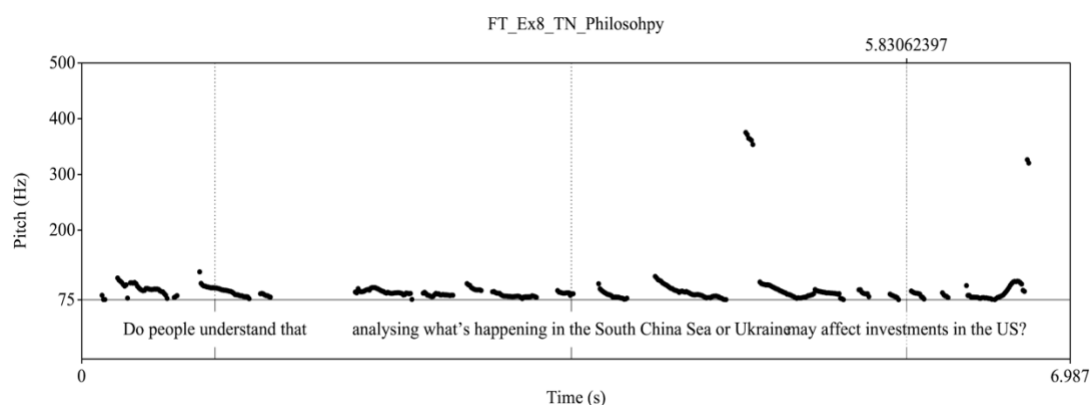| Do	people	understand	that analysing…?
Auxiliary verb	**Subject**	**Infinitive**	**Rest of clause**

Image 56 Pitch contour in FT.Ex8

Therefore, both from the structure and intonation of the sentence in the source text, it is clear that the speaker is evoking a response from the audience. If we check the Spanish translation, we can see that, instead of a *yes-no* question, it offers a

declarative statement, which has a completely different illocutionary meaning. In terms of content, the translation is literal, and there are no question marks in the script or rising intonation in the delivery to show this is also a question. In fact, it has been voiced as a definite declarative statement, so rather than evoking an answer, it provides one. The image below shows the pitch contour of the Spanish translation.
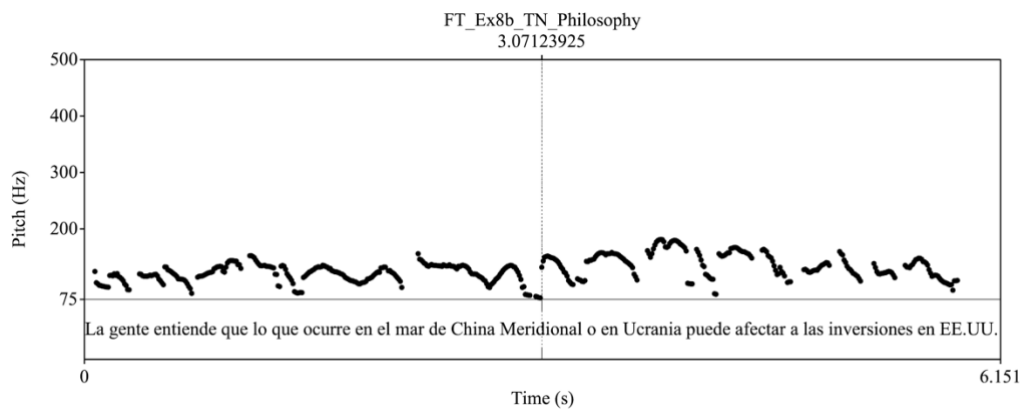

Image 57 Pitch contour in the target text of FT.Ex8

The original text, with the *yes-no* question, is probably seeking a negative answer so that the audience is convinced by the firm's investment capabilities and expertise:

- Collin Moore: Do people understand that analysing what's happening in the South China Sea or Ukraine may affect investments in the US?
- Audience answers: No they/we don't.
- Collin Moore (continues): So we must be able to think globally….

The speaker is, thus, questioning people's capability to understand how geostrategic affairs may affect investments worldwide. What Moore is not doing is affirming that they do understand how geostrategy plays a major role in world economics, which is, in fact, what we get in the Spanish version. Therefore, we can conclude that in this case, the prosodic-illocutionary meaning of the original source text has not been successfully rendered in the target text.

F.T.Ex 9 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:38:29	00:38:32	Hilary Mantel	So Henry's going quietly **mad**.	En su fuero interno se estaba volviendo **loco**.

This last example of partial rendition belongs to the documentary *The last day of Anne Boleyn* and shows Hilary Mantel talking about King Henry VIII's reaction to the rumours of his wife's infidelity. As can be seen in the image below showing the pitch contour for this line, there is one IP whose nucleus is "mad" which has a high fall nuclear tone.
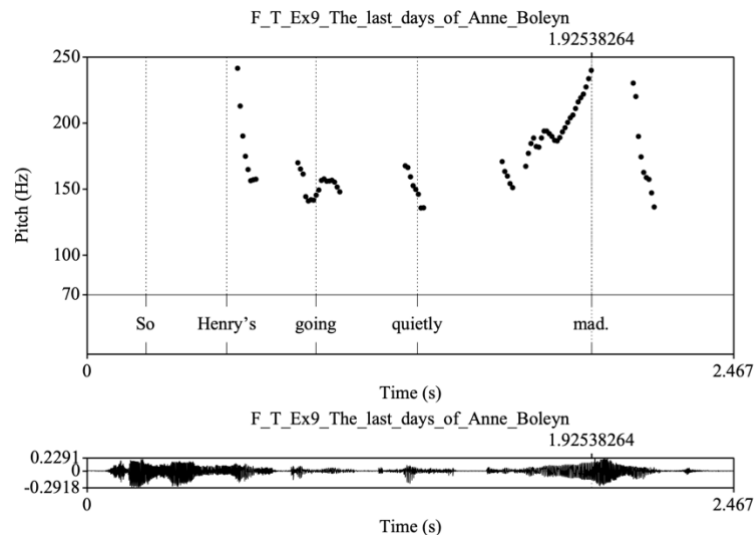

Image 58 Pitch contour in FT.Ex9

As can be seen on the pitch contour above, "mad" is uttered at a fairly high pitch, higher than the preceding words from the high head. This high fall is actually emphatic and serves to highlight that Henry was, in fact, fuming.

If we now check the Spanish translation for this passage, we can notice that "mad" has been rendered as "loco", which in our view is not the best option, since "mad" apart from referring to a mental condition or state (crazy) also means "very angry". A more suitable alternative for "mad" could be "enfadadísimo" or "furioso",

and to reflect the high fall tone we could say "realmente enfadado/furioso" or "increiblemente enfadado/furioso" to emphasise how angry he was.

The melodic pattern for the Spanish passage does not show any insistence or emphatic stress on "loco" to highlight how Henry was feeling at the time, as shown in the image below.
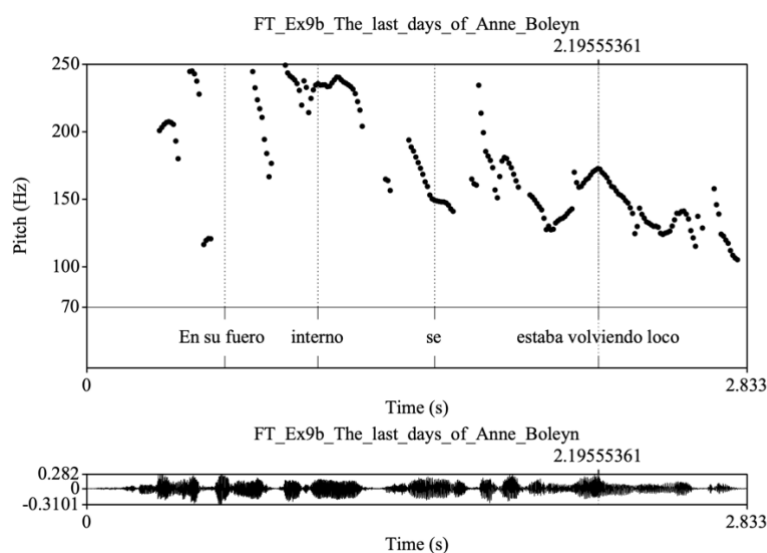

Image 59 Pitch contour for the Spanish version in FT.Ex9

The translation of this passage, with an evident emphatic high fall tone, does not successfully render the illocutionary force of the original message, and consequently a relevant part of the original message is lost.

5.2.4.1.1 **Concluding remarks on falling tone rendition**

The role that falling tone has in English is very ample, as we have seen. It can be used to state an opinion and request confirmation in tag questions, request an answer in *yes-no* questions, show excitement, and provide assertions. The translation strategies used to render the pragmatic meaning of this tone have proved to be diverse, and range from exclamative questions, showing surprise and emotion, to rhetorical confirmative questions meant to obtain the desired answer, or declarative statements stating opinions and assertions. Thus, we could see that interrogatives not always seek information, as

in examples FT.Ex1, FT.Ex2, and FT.Ex3, which confirms Vidal's difference between a question and an interrogative sentence (1999: 03).

On the other hand, we have also seen that there are times when falling tones can mislead the translator and what is meant as a declarative statement is turned into a question in the target text. Nonetheless, we shall recall that, in conversation, as explained by Leech and Svartvik (1994: 125), both statements and questions often evoke a response. When we deal with questions, this response takes the form of an answer. But declarative statements can also be responded, which does not turn them into questions. This can be seen in example F.T.Ex6. In addition to this, a falling tone combined with a high pitch (high-fall tone) can also be misleading, as illustrated by example F.T.Ex7, which shows a mixture of over excitement and hesitation in the target text, while in the source text the pragmatic meaning is completely different.

5.2.4.2 **Rising tone**

There are 10 examples of rising intonation in our corpus including those belonging to the tag question's category. Leaving these on one side, we get a total of only 5 examples of rising intonation, which will be analysed in this section. As can be seen in the pie chart below 60% of the 10 examples show full rendition of the original message, while 40% are partially rendered (semantic meaning only).
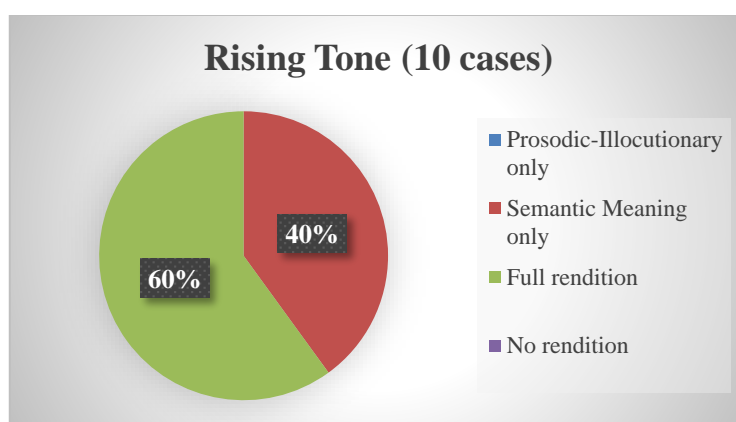

Image 60 Pie chart showing the rendition of the 10 cases of rising intonation

The various rendition strategies for these 10 samples of rising intonation in the

Spanish target text can be seen in the bar chart below. There are 6 strategies in total. From these, 2 belong to tag questions as seen in section 5.2.3.4, and the rest include: 1 example of rising question translated as a declarative/exclamative, 1 example of declarative with rising intonation (question) translated as a question, example of declarative with rising intonation (question) translated not as a question, and 2 rises indicating more information is coming which are translated as full stops.
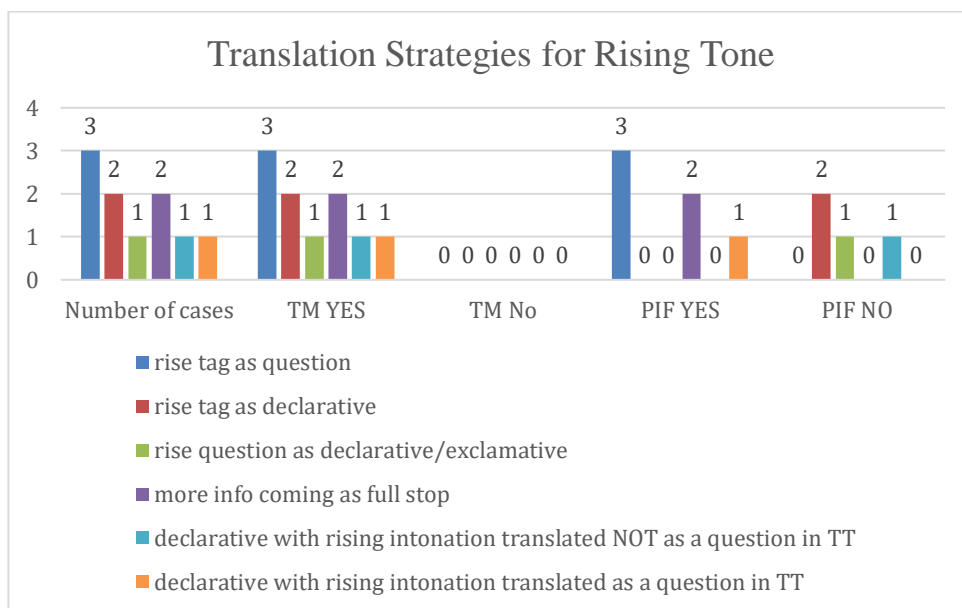
Image 61 Pie chart with the rendition strategies of rising tones

The bar chart shows that all the strategies successfully render the semantic meaning of the original message. In terms of prosodic-illocutionary force, we can see that there are 3 instances in which it is not rendered: rise tag as a declarative, rise question as a declarative/exclamative, and declarative with rising intonation translated not as a question in the TT.

The number of examples and strategies used per genre is shown in the table below.

DOCUMENTARIES			
Category	Number of Cases	Strategy	Number of cases

		rise tag as question	2
Rise	6	rise tag as declarative	1
		rise question as declarative/exclamative (1)	1
		more info coming translated using a full stop	2

INTERVIEWS VIDEOS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Rise	1	rise tag as declarative	1

REALITY TV			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Rise	3	rise tag as question in TT	1
		declarative with rising intonation translated as a question in TT	1
		declarative with rising intonation translated NOT as a question in TT	1

WEB VIDEOS			
Category	**Number of Cases**	**Strategy**	**Number of cases**
Rise	0	-	0

Table 30 Number of rising tone examples and rendition strategies in our corpus

As shown in the table above, the number of rising tags account for half the total in our corpus. We will now analyse the remaining five examples, starting with those showing full rendition.

a) **Full rendition**

R.T.Ex 1[80] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:16:19	00:16:24		She is so pale that she virtually doesn't /exist,	Tenía una tez tan pálida que era casi como si no existiera.
00:16:24	00:16:26	David Starkey	she is desperately /plain,	Y un caracter terriblemente soso.
00:16:26	00:16:31		and Henry, like a \/pendulum, swings from one to the \other.	Enrique, como un péndulo, va de un extremo a otro.

This first example of full rendition of rising tone, from the documentary *The last days of Anne Boleyn*, shows David Starkey describing Jane Seymour, Henry's wife after Anne Boleyn. In his speech, Starkey uses one sentence which is divided into six IPs, most of them showing a rising nuclear tone. The IP structure is as follows:

| she is so \/pale | that she virtually doesn't /exist | she is desperately /plain | and - Henry like a \/pendulum | swings from one to the \other |

These instances of rise and fall-rise show what Wells (2006: 69) calls *leading tones*, attributed to dependent elements within a sentence, and which imply that more material is still to come, thus indicating non-finality. The finality is shown with a fall in the independent element "swings from one to the \other". In this example, the independent structure is the last IP, which shows a falling tone. All the preceding IPs are dependent on this one and, as such, show a rising intonation. As Solé (1989: 192) puts it: "Sentences with rising tones are subordinate to groups with falling tones". The image below shows the pitch contour for the two rising tones in the IPs: | that she virtually doesn't /exist | she is desperately /plain |
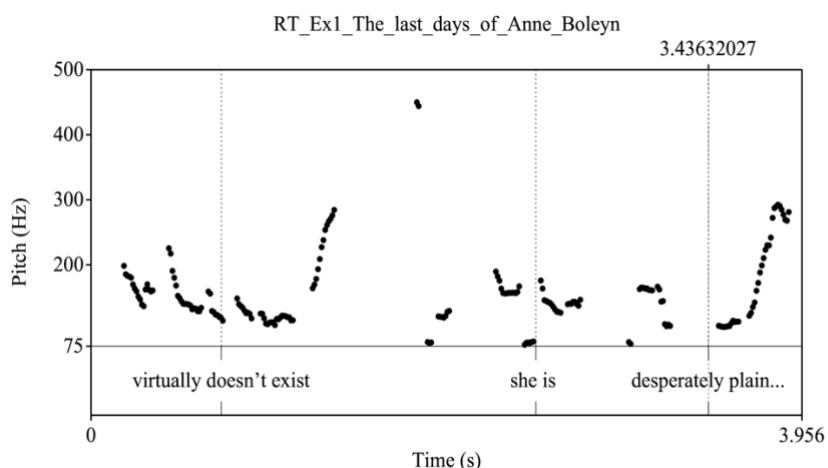
Image 62 Pitch contour in R.T.Ex1

If we now check the Spanish translation for this passage, shown in the picture below, we can see that the intonation used to voice this scene is a fall, which is the usual one in segments that are separated by commas and full stops (Álvarez del Valle 2016: 106, 107).
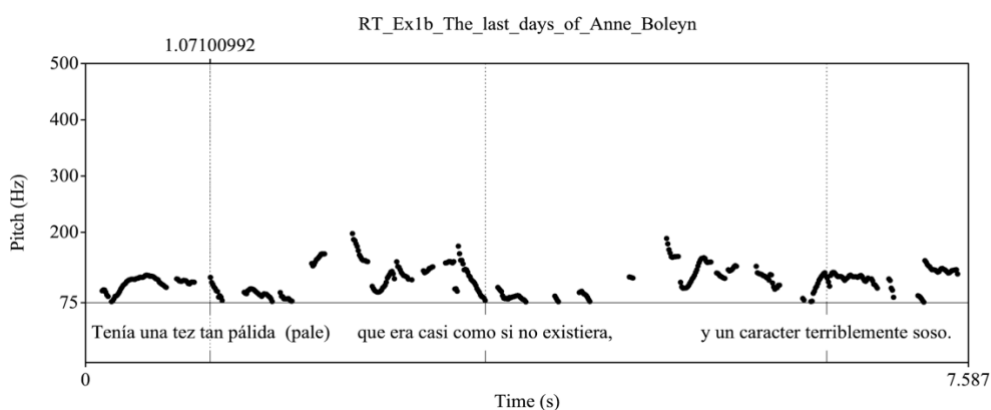

Image 63 Pitch contour in the Spanish version of R.T.Ex1

Since we have transcribed the target text on our own, and we do not have the original translated script, we are not 100% sure whether commas or full stops have been used by the audiovisual translator. Nonetheless, what we can note is that, by using this type of punctuation and intonation, the sensation that more information is to come (implicit in the original rising tones) is lost, due to the fact that commas81 split terms without relating them. However, this strategy of splitting the sentence in different independent chunks with their own definite fall is completely valid because, in the end,

81 Usage 1.2 of commas in Spanish: http://lema.rae.es/dpd/srv/search?id=V1EqcYbX4D61AWBBrd

we get the whole information.

R.T.Ex 2 *Teen Mom (clip 1)*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:01:10	00:01:11	Farrah	You were not on the phone with the /lady?	¿No hablaste con ella por /teléfono?	00:02: 31	00:02: 33

This second example of full rendition of a rising tone, belongs to the second clip of episode 6 of the reality show *Teen Mom*, and shows the USA celebrity Farrah Abraham talking to Simon, her fiancé, and a marriage counselor about an engagement ring he was supposed to buy. In her speech, Farrah asks Simon whether or not he spoke to the lady, a jeweler, in charge of making the ring. Farrah does so using a declarative sentence with rising intonation, which turns the statement into a question (Wells 2006: 36; Granato 2014: 219).

The Spanish translation for this line shows a literal translation with rising intonation, which works as a rhetorical confirmatory question, just like the declarative question in the source text. The prosodic-illocutionary meaning is therefore the same. The speaker (Farrah) is trying to get a negative confirmation of the fact that Simon did not speak to anyone at the jewelry. The melody used in both texts is shown in the picture below.
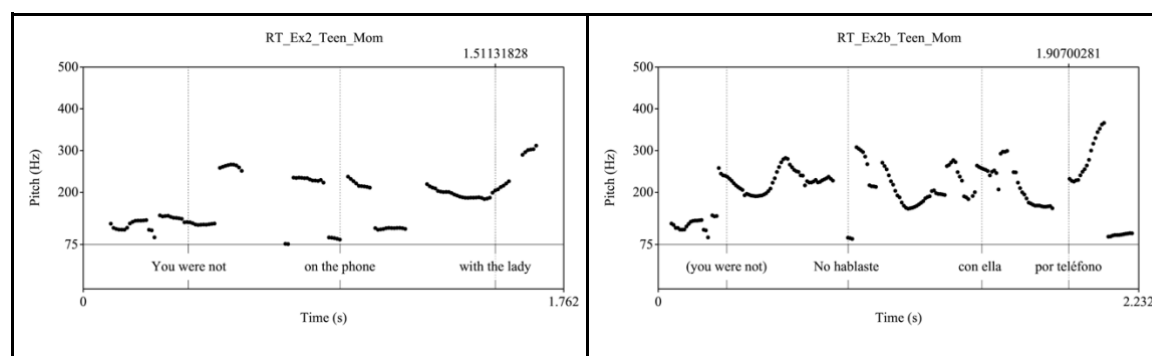

Image 64 Pitch contour in R.T.Ex2

The Spanish text shows a pragmatically equivalent construction, which even coincides in the tone used. Therefore, we can state that there is full rendition, both in terms of illocutionary force and semantic meaning, of the original source text.

b) Partial rendition

R.T.Ex 3 *Jamies comfort food*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:01:23	00:01:24	Jamie Oliver	It's all right to eat messy, /yeah?	No pasa nada por mancharse.	00:01:24	00:01:24
00:01:24	00:01:26	Christian	You're supposed to eat messy! What are you, a wuss?	Claro que no. ¿Qué eres: un cursi?	00:01:25	00:01:26

This example from the documentary *Jamie's comfort food* shows renowned chef Jamie Olivier talking to his friend Chris about burgers. After preparing a couple of juicy burgers filled with ingredients, they wrap themselves up with kitchen roll and get ready to eat. After the first bite, Jamie uses a construction with rising intonation seeking for permission to eat the burger in such a way: "| It's alright to eat \messy | /yeah |" This construction functions as a rising tag question, even though there really is no positive tag "(is it?)" at the end; the interjection "yeah", with a rising intonation, has the same pragmatic value as a checking tag (Wells: 2006: 54), so it works here as a marker of a *yes-no* question and, therefore, requests a confirmation as if it was a question tag. The image below shows the pitch contour for this construction.
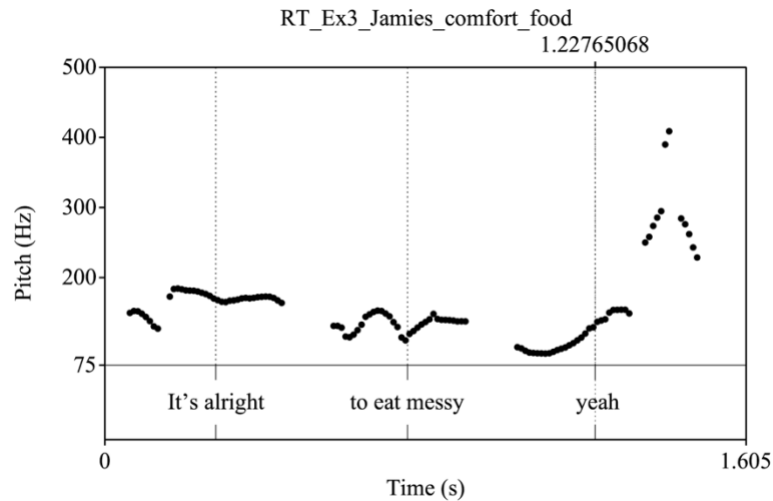
Image 65 Pitch contour in R.T.Ex3

The Spanish text provided for this line shows a declarative statement which, rather than requesting confirmation and approval, asserts it and firmly states that there is no problem about eating messy. With this type of construction, the original pragmatic value in the source text is lost in the target version. Nevertheless, the overall meaning of the line and of the whole scene is not lost since, despite the fact that we get a statement rather than a request, Chris answers Jamie in the translation too, so the original request *is* answered. The image below shows the intonation used in the Spanish text, reflecting the falling intonation typical of statements.
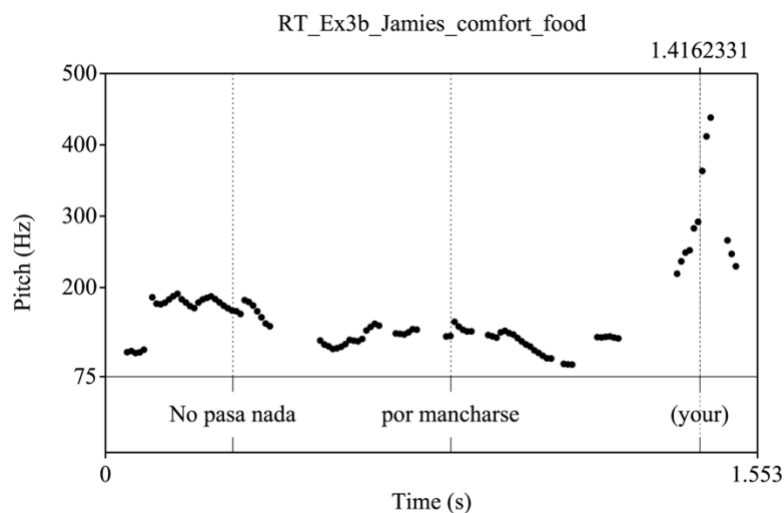

Image 66 Pitch contour in the target version of R.T.Ex3

In any case, this example would have worked perfectly with a ratified question in the target text: "No pasa nada por marcharse, ¿no?". Even an affirmative *yes-no*

question (interrogativa total) could work: "¿Pasa algo por mancharse?" In this way, we would get the request present in the original source text while its illocutionary force would not be lost in translation.

R.T.Ex 4 *Cupcake Wars*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:45:50	00:45:52	Ashley Eckstein	Now it's time to go to Star Wars /Celebration?	¡Vámonos a la fiesta de La Guerra de las Galaxias!	00:35: 22	00:35: 24

This last example of partial rendition of a rising tone example, from the reality show *Cupcake Wars*, shows judge Ashley Eckstein's final thought at the very end of the programme. This show, dedicated to the Star Wars movie saga, awarded the winning team ten thousand dollars and the chance to display their cupcakes at the Star Wars Celebration[82] VI, an exhibition dedicated to the film saga that is celebrated in a different city of the USA every year. Once the final cupcake war (the last round of the programme) takes place and the winner is named, judge Ashley Eckstein asks the people in the show (judges and contestants) a *yes-no* question in the form of a declarative question (a declarative statement with rising intonation). Her words are accompanied by a palms-up and arms-up gesture, typical of interrogative sentences (Givens 2016: 237), which reinforces the illocutionary force implicit in the rising tone. The pitch contour of this sentence and Mrs Eckstein's body language can be seen in the pictures below.

---

[82] Star Wars Celebration is a Star Wars dedicated exhibition. https://www.starwarscelebration.com/
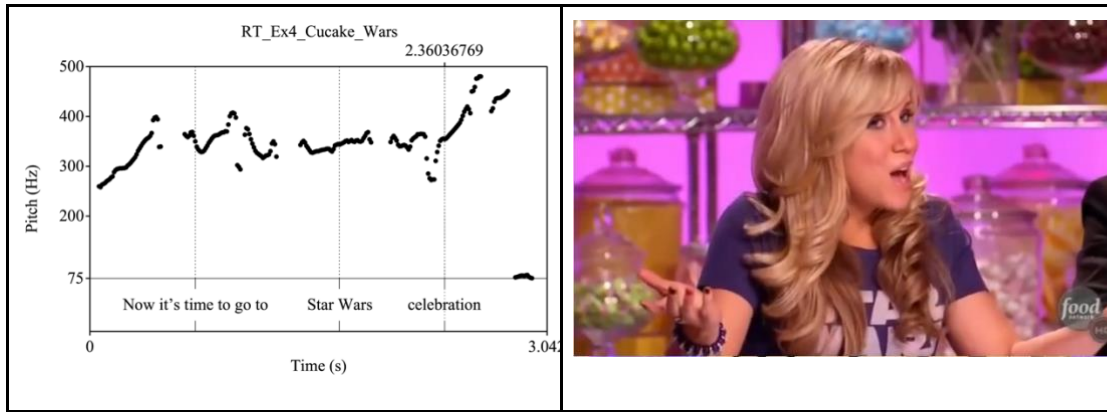
Table 31 Pitch contour and body language in RT.Ex4

The Spanish translation for this line, resorted to an exclamative sentence rather than a *yes-no* question or another construction in Spanish demanding an answer, such as rhetorical questions, exclamatory questions, and ratified questions. Thus, in our view, the illocutionary force of the original text has not been successfully rendered in the target text. In addition, the body language is out of sync with the verbal expression, since while the body language is asking, the words in the exclamation used express the speakers' attitude. The image below shows the pitch contour of the Spanish translation for this example.
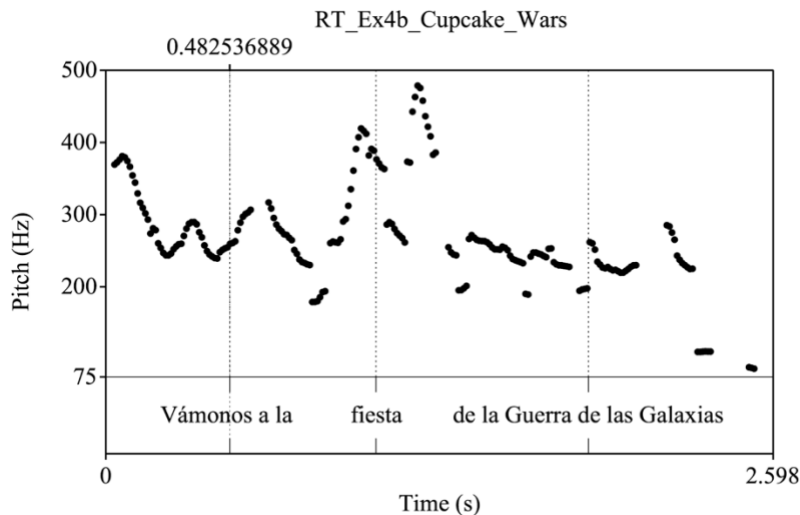

Image 67 Pitch contour in the Spanish version of R.T.Ex4

The Spanish line shows a falling intonation at the end of the line, preceded by a rise in "fiesta", which is, thus, highlighted. We are not sure why this word has been highlighted here, or why this particular intonation, rather sing-song, is used. Probably

it is a mixture of overacting and the unnaturalness of certain dubbed dialogues, voiced-over in our case, commonly known as *doblajitis* or *dubbitis* (Sánchez Mompeán 2020: 32, 148) among scholars. This fluctuations in pitch are conscious and for Sánchez Mompeán (2017: 345-346) might carry unconscious attitudinal and illocutionary divergences regarding the original source text. Therefore, we can say that in this example the rendition is only partial, since the prosodic-illocutionary force has not been successfully rendered in the target text.

5.2.4.2.1 **Concluding remarks on rising tone rendition**

Now that we have analysed all the examples, we see how versatile rising intonation is in English, and how native speakers use it in spontaneous speech. It can be used to join parts of a sentence preceding the main informative element, as in example RT.Ex1; or as a checking *yes-no* question using interjections as in example RT.Ex3, which works as an alternative to rising question tags. It can even turn a declarative statement into a question just by changing the usual falling tone for a rise, as in examples RT.Ex2 and RT.Ex4.

From these, we could say that the most relevant cases in terms of pragmatic load and rendition of illocutionary force in Spanish voice-over are: those in which rising intonation turns a declarative statement into a question; checking interjections; and as seen in 5.2.3.4, rising tag questions. It is in these types of sentences that the audiovisual translator should identify possible implicatures and illocutionary force in order to find an equivalent pragmatic construction in the target text. Therefore, if the speaker on screen utters a declarative with rising intonation and rises the arms with a palm-up gesture, the implicit meaning is that it is a question, rather than an assertion. It seems that, in American English, rising declaratives, similarly to rising yes-no questions, are more polite than falling questions (Nilsenová 2006: 133), so it is very likely that the audiovisual translator will come across these constructions in the wide range of TV shows coming from America that are translated by means of voice-over.

Solé (1989: 193) considered that the function of tone in showing the *relative*

importance of sentences did not present a major problem for the translator. We understand that this "relative importance" refers to the discourse structure signalled by tone, as in the case of rising subordinates, which, as we have seen, do not pose significant problems. Nonetheless, tone serves other purposes, in terms of signalling pragmatic function, apart from organising information, as seen in section 2.2, and this has proven to be challenging.

5.2.4.3 Tag questions

This section is devoted to the rendition of tag questions, which, as stated in section 2.2.2.4, are pragmatic resources commonly used in spontaneous oral speech in English. Tag questions can have different meanings depending on the intonation used in the tag clause at the end of the sentence. If the tag has a falling tone, then, the sentence is more like a statement than a question and no answer is expected. This type of tag is commonly used to express an opinion (Leech and Svartvik 1994: 151), and it can also appeal for agreement (Wells 2006: 49). On the other hand, if the tag has a rising tone the speaker is actually requesting an answer and the sentence has more of an interrogative value.

There are 13 examples of tag questions in our corpus, as can be seen in the image below, with 85% showing full rendition, 15% partial rendition (semantic meaning only), and no examples of no rendition.

Tag Questions (13 cases)

15%

85%

- Prosodic-Illocutionary only
- Semantic Meaning only
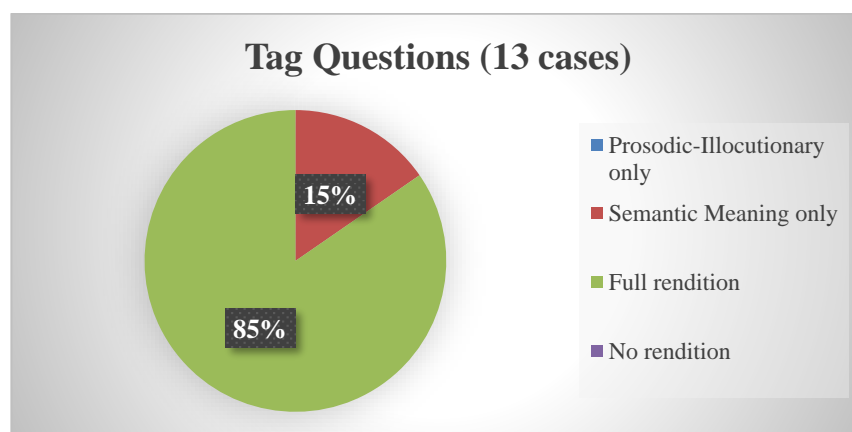- Full rendition
- No rendition

Figure 35 Number of tag questions and their rendition in Spanish

The translation possibilities for the tag questions we have found are five: rising tag translated as a question; rising tag not translated as a question; falling tag translated as a question; falling tag not translated as a question; tag question translated with ratified "¿no?/¿verdad?". As can be seen in the chart below, the strategies vary even for tags of the same kind. Thus, for the 5 rising tags identified, we get 1 translated as questions, 2 as statements, and 2 using "¿no?/¿verdad?". In addition, for the remaining 8 examples with falling tags, we have 2 translated as questions, 3 as statements and 3 using "¿no?/¿verdad?".
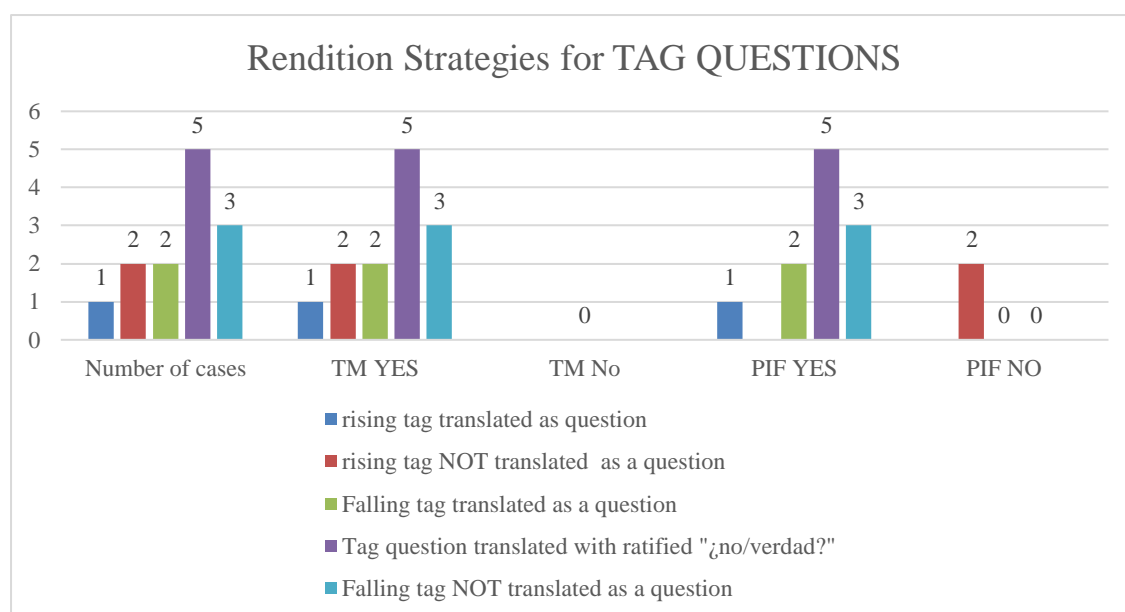
Figure 36 Rendition Strategies for Tag Questions

The examples of question tags and the translation strategies used for them in our corpus are shown in the following table according to programme genre.

DOCUMENTARIES			
Tag Questions	6	rising tag translated as question	1
		rising tag NOT translated as a question	1
		Falling tag translated as a question	1
		Tag question Translated with "¿no/verdad?"	3

		Falling tag NOT translated as a question	0
INTERVIEWS VIDEOS			
Tag Questions	2	rising tag translated as question	0
		rising tag NOT translated as a question	1
		Falling tag translated as a question	1
		Tag question Translated with "¿no/verdad?"	0
		Falling tag NOT translated as a question	0
REALITY TV			
Tag Questions	5	rising tag translated as question	0
		rising tag NOT translated as a question	0
		Falling tag translated as a question	0
		Tag question Translated with "¿no/verdad?"	2
		Falling tag NOT translated as a question	3
WEB VIDEOS			
Tag Questions	0	rising tag translated as question	0
		rising tag NOT translated as a question	0
		Falling tag translated as a question	0
		Tag question Translated with "¿no/verdad?"	0
		Falling tag NOT translated as a question	0

Table 32 Number of examples of tag questions and their translation strategies for each genre

Let us analyse now some of these examples, starting with those showing full rendition, then partial, and finishing with the one example showing no rendition.

a) **Full rendition**

T.Q.Ex 1[83]*The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:32:20	00:32:23	Greg Walker	I probably would confess to having sex with the Queen	Si a mi me torturaran probablemente habría confesado
00:32:23	00:32:25		if I was tortured and I wanted it to stop, /<u>wouldn't you?</u>	que me había acostado con la Reina para que dejaran de martirizarme, /<u>¿usted no?</u>

This first example of tag question in our analysis, from the documentary *The last days of Anne Boleyn*, shows Greg walker talking about the alleged torture Mark Smeaton suffered to make him confess he had slept with Anne Boleyn. In his speech, Walker ends with a rising tag (shown in the pitch contour image below), with which he invites the interviewer and the audience to respond. He is not confirming or making a statement as would have been the case with a falling tag; he is expecting an answer.
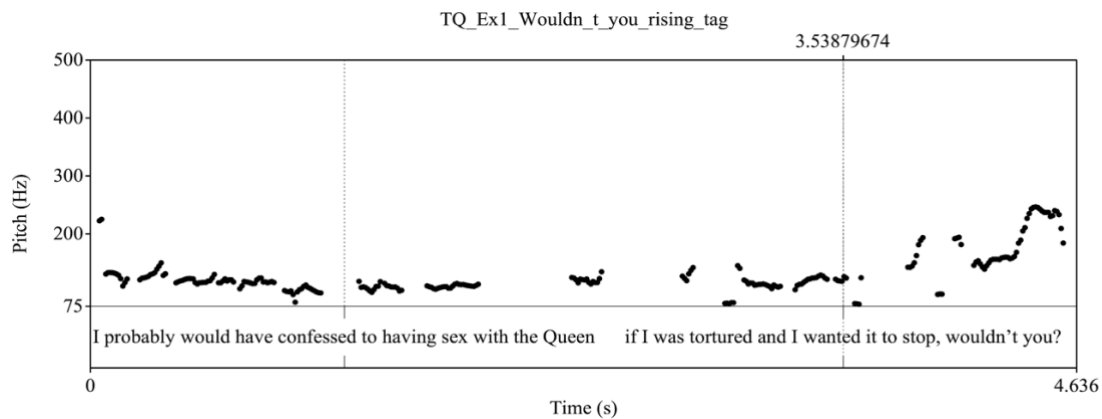

Image 68 Pitch contour in T.Q.Ex1

The translation in the target text for this tag is a question ("¿usted no?") with which the translator renders both the semantic meaning and the prosodic-illocutionary force of the original message, since, what we get is also a request, as intended by the speaker in the original text.

---

[83] TQ. Ex stands for Tag Question Example.

T.Q.Ex 2 *Kitchen Nightmares*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:18:44	00:18:49	Gordon Ramsay	Yeah, not a fucking penny. Not a single penny! But you've taken from her, /haven't you?	No, ni un penique. Ni un solo penique. Pero sí has cobrado de ella. /¿No?	00:18: 31	00:18: 34
00:18:50	00:18:50	Tim	Yeah.	Sí.	00:18: 35	00:18: 35

This second example of full rendition of a tag question, from the reality show *Kitchen Nightmares,* shows chef Gordon Ramsey telling off the cook he is set to help. He finishes his speech with a rising tag requesting an answer, which is given just after. The pitch contour for this tag is shown in the image below and the rising tune can be clearly seen at the end.
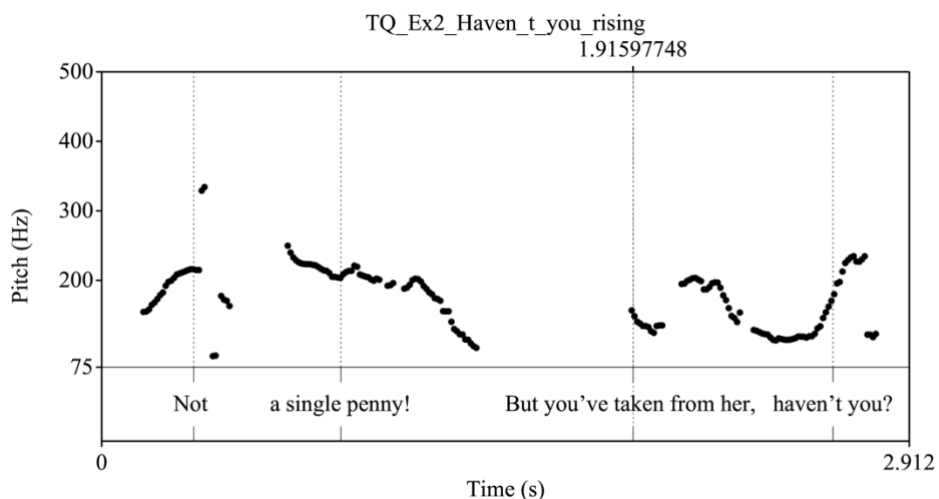

Image 69 Pitch contour in T.Q.Ex2

The ratified question "¿no?" offered as a translation for this rising tag question in the Spanish text (see section 2.5.1), keeps the prosodic-illocutionary force of the original (appeal for an answer). We can see that in the source text there is an answer for this tag question "yeah" just after, which implies that, in the target text, the tag is

expected to be translated as a question, since there is an answer to be translated afterwards. In addition to this, the emphatic high fall tone on "taken" in the source text has been rendered using the emphatic particle "sí" in the target version, which highlights the affirmative statement of the sentence, and therefore, keeps the illocutionary value of the source text.

T.Q.Ex 3 *Kitchen Nightmares*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:20:56	00:21:00	Gordon Ramsay	Thank you. Don't spill any on those new trainers, \will you?	Gracias, querida, gracias. Ese es el tuyo. <u>No te lo eches en las deportivas nuevas</u>.	00:20:41	00:20:45
00:20:00	00:21:01	Tim	I'll try not to.	No. Intentaré que no.	00:20:45	00:20:46

This third example of full rendition of a tag question shows, again, Gordon Ramsey now teasing his pupil, Tim, who has just ordered a burger in a street market. Ramsey is joking and tells him not to drop any of the ingredients of the burger on his (Tim's) new trainers using a falling tag "will you?", which is typical of weakened commands (Leech and Svartvik 1994: 167). The tag, as can be seen in the pitch contour image below, is a fall, meaning it is not a real question. In fact, the tag *will you?* is often used after imperatives or functioning as reinforcing a command.
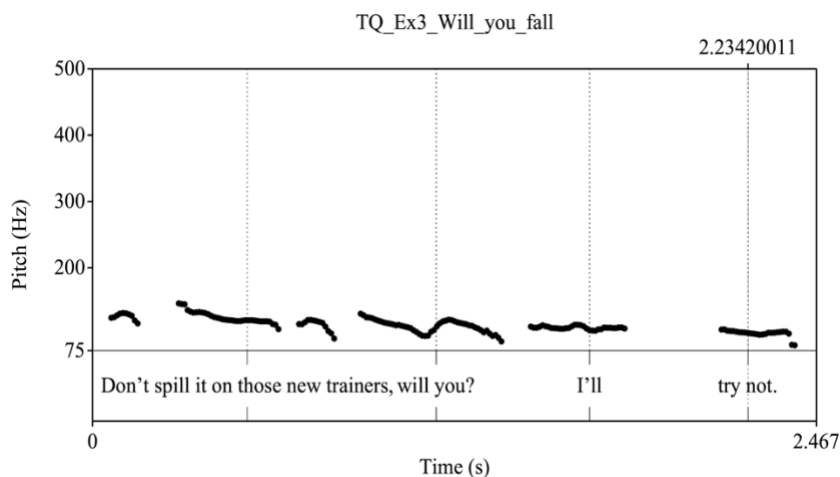

TQ_Ex3_Will_you_fall

The target text also includes a negative imperative and the tag has been elided, which makes sense, since the illocutionary force of the sentence is that of "commanding" or "giving strong advice" rather than asking. Gordon Ramsey is just teasing Tim by joking about his new trainers, which are apparently forged. All this is accompanied with an ironic tone of voice in the delivery of the target text, which fully renders the original message and the prosodic-illocutionary force.

T.Q.Ex 4 *What's it like for EU migrants in Britain amid UKIP's success?*				
In	Out	Character	Source text	Target text
00:03:17	00:03:23	Man	Seems too late really. Jobs been taken, houses been taken, \aren't they? You can't get rid of them now, \can you?	Ya es demasiado tarde. Se quedan con los puestos de trabajo, tienen casas. /¿Cree que puede deshacerse de ellos?

This is the last example of full rendition of tag questions, from the Euronews programme *Reporter: What's it like for EU migrants in Britain amid UKIP's success?,* and showing two instances of falling tags. In this scene, a man is talking about the impact of EU immigrants in the UK. As we can see in the table below, he uses two falling tags to express his opinion. We can say that the usage of this fall has exclamative force as suggested by Wells (2006: 49).
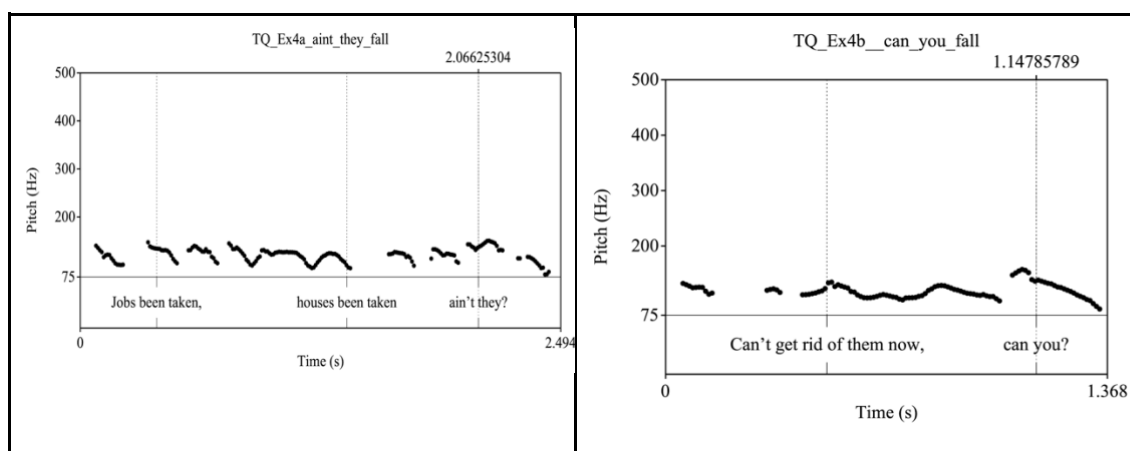
The translation has been done using a declarative statement for the first sentence with tag, which is perfectly translated. On the other hand, a question has been used for the second. As we have seen, falling tags do not work as real questions since they do not ask for information but for confirmation or are used to soften an opinion. However, this time the question works as non-neutral interrogative statement (see section 2.5.1) and the pragmatic function of the original tag is rendered in the Spanish version by means of a question whose illocutionary force is to show irony and sarcasm, rather than requesting real information.

b) Partial rendition (semantic meaning only)

T.Q.Ex 5 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Taget text
00:34:45	00:34:50	Greg Walker	She alternates between a sense that the law will save her,	En un primer momento piensa que la justicia la salvará,
00:34:50	00:34:53	Greg Walker	she's innocent, that will come out, /won't it?	porque es inocente y eso saldrá a la luz, ¿a que sí?
00:34:53	00:34:55	Greg Walker	Henry's just testing her, /isn't he?	Porque Enrique la está poniendo a prueba, solo es eso.

This first example of partial rendition of tag questions, retrieved from the documentary *The last days of Anne Boleyn,* shows two rising tags, as shown in the table below. Greg Walker comments on Anne Boleyn's reflections about her situation and how she might have felt during the accusation process while she was kept in the Tower of London. The first tag has been translated using a ratified question just after the main clause, appealing for an answer, as normally happens with rising tags in English. However, the second rising tag, which comes right after, has been rendered as a declarative statement, an assertion, while the suitable way of translating it would have been, in our view, a question, as in the previous tag.

As we can see in the table below, the two tags have the same tone pattern and function but have been translated differently.
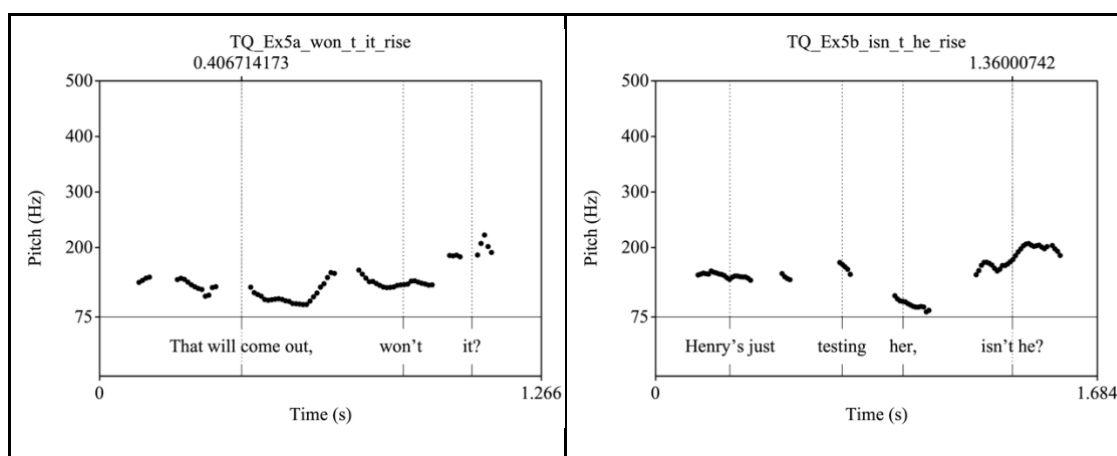

Image 72 pitch contour in T.Q.Ex5

The different translations used here are probably due to the fact that the two tag questions are said consecutively by the speaker, so the translator decided to only use a question for the first one. However, if we link the two sentences and use the ratified question at the end, both clauses would be affected by it and the illocutionary force of the second rise would have been rendered; for example:

- En un primer momento piensa que la justicia la salvará. Cree que su inocencia saldrá a la luz y que Enrique solo la está poniendo a prueba, ¿o no?

T.Q.Ex 6 *What's it like for EU migrants in Britain amid UKIP's success?*				
In	Out	Character	Source text	Target text
03:44	03:51	Woman 2	I don't think you can quantify too much or too little. I think it's a factor, and you know, it's just like it or lump it, /isn't it? hahaha.	Yo no creo que se pueda cuantificar como algo positivo o negativo. Creo que la situación es la que hay y hay que aguantarse.

This second example of partial rendition of a tag question, from the *Euronews* programme *What's it like for EU migrants in Britain amid UKIP's success?* shows a sentence with a rising tag translated as a declarative statement. As we can see in the

image below the pitch contour of the tag is rising, which implies the speaker is requesting information. However, the Spanish translation shows a declarative statement, expressing an opinion, that would have worked perfectly if the original tone was a fall.
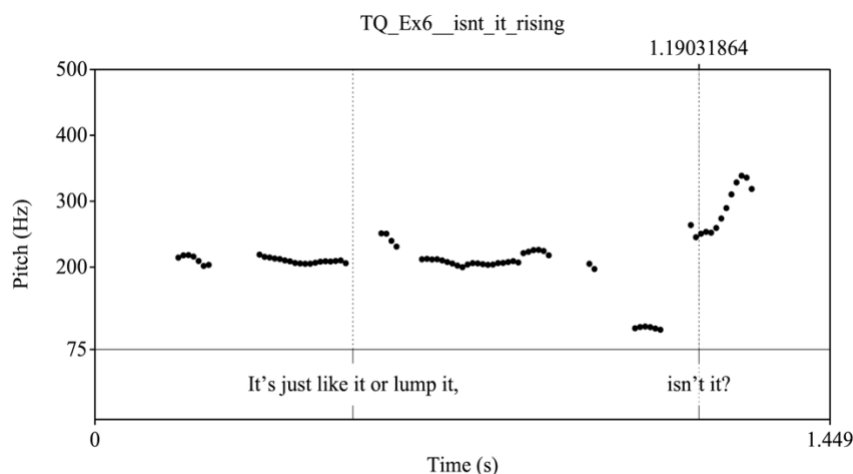

Image 73 Pitch contour in T.Q. Ex6

This time the pragmatic function of the original tag, implicit in the rising tone, has not been successfully rendered in the target text. The translation would have been fine with a ratified question asking the interviewer and the audience for confirmation. Nonetheless, the basic semantic meaning of the original sentence has been transferred, resulting in partial rendition.

5.2.4.3.1 Concluding remarks on the rendition of tag questions

As we have seen in this section, there are several ways of translating tag questions into Spanish, the commonest ones being ratified questions. The use of this strategy seems to work fine in terms of meaning transfer since, after all, it is the equivalent construction in Spanish (declarative statement + interrogative). However, we have seen that there are other options, such as simple declarative statements and absolute exclamative questions, that can work even better when the original tag does not request real information. These are perfect to render the prosodic-illocutionary force of falling tags expressing opinions, exclamations and control. Rising tags, on the other hand, when used in reversed polarity tags, genuinely ask for information, which means that they are

real questions and cannot be rendered by means of declarative statements. Reese and Asher (2007: 459) understand that the semantic contribution of a final rise in tag questions blocks the default inference of *acknowledgment* falling tags. Thus, ratified questions can work in this case, since, in the end, they are questions and request an answer from the receiver. Nonetheless, real questions can also be used if we want to make use of other types of constructions in Spanish

5.2.5 **Tonicity**

This section is devoted to the analysis of the examples showing marked tonicity in our corpus; that is when the nucleus does not fall on the last lexical item as is the case of broad focus. We have decided to focus our study on narrow focus examples since these are clear illustrations of how intonation in English works in order to highlight information within a sentence by just placing the tonic syllable on the word or information we want to emphasise), showing much greater flexibility than Spanish in this regard; therefore it will be interesting for this translation analysis of intonation. Broad focus examples have not been included in our study.

5.2.5.1 **Narrow focus**

We have found 61 examples of narrow focus in our corpus. From these, 36 are due to deaccenting (default accent) and 25 to contrastive focus. The rendition results shown in the pie chart below include those shown in the deaccenting section. However, the rendition strategies differ, since here we will only show the rendition strategies for the nucleus and not for the deaccented material. The study that follows will hopefully serve to demonstrate the relevance that tonicity plays in English communication and also in voice-over translation.

From the 61 examples found, 62% show full rendition, 2% no rendition at all, and 36% partial rendition (35-5 semantic meaning only and 1% prosodic-illocutionary only). The study that follows will hopefully serve to demonstrate the relevance that tonicity plays in English communication and also in voice-over translation.
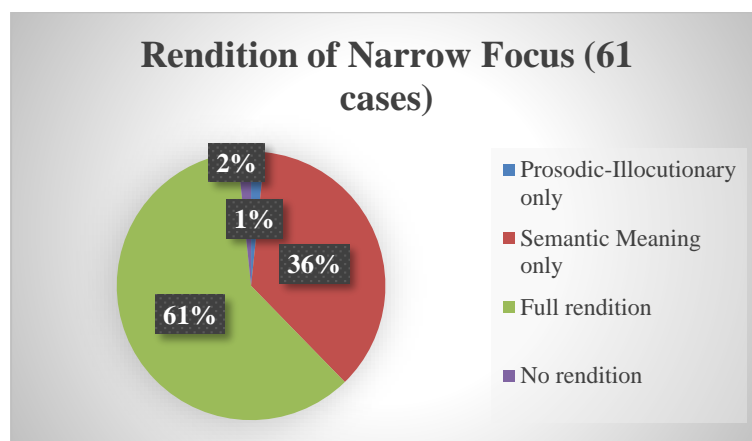
Figure 37 Pie chart with the rendition results for narrow focus examples

In total we have identified 6 translation strategies for the 61 narrow focus examples in our corpus, with different success results as can be seen in the bar chart below: literal translation with flat intonation (16 cases), literal translation with emphatic intonation (4 cases), broad focus - placing the nucleus at the end of the sentence - (20 cases), use of emphatic particle or syntax (11 cases), ellipsis or deletion of nucleus (9 cases), literal translation with wrong emphatic intonation (1 cases).

There is just one example showing no rendition, also belonging to the deaccenting section (see D.Ex10), in which the target text shows a complete reformulation and, therefore, neither the semantic meaning nor the illocutionary force are equivalent to the original text.

As can be seen in the chart below, the most successful rendition strategy to render English narrow focus in Spanish is using a literal translation with an emphatic intonation in the delivery of the text with 6 cases out of 6 showing both rendition of semantic meaning and illocutionary force. The second most successful strategy is broad focus (placing the nucleus at the end of the sentence) with 20 cases out of 20 showing transfer of basic meaning and 19 out of 20 showing rendition of prosodic-illocutionary force. The third most successful strategy is using an emphatic particle and syntax with 13 out of 13 showing transfer of semantic meaning and 12 out of 13 prosodic-illocutionary force. Finally, deleting the nucleus has worked well to render the prosodic-illocutionary force in one example, which is quite exceptional.
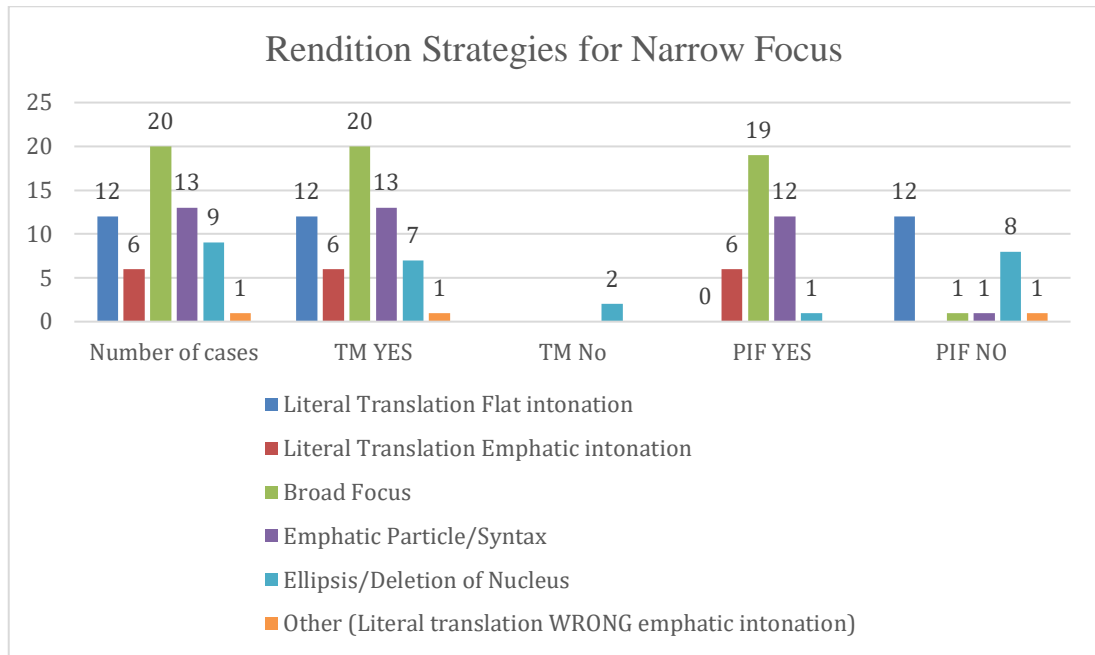
Figure 38 Rendition strategies for narrow focus examples

In terms of genre, the strategies identified for each voice-over category are as follows:

DOCUMENTARIES		
Number of Cases	**Strategy**	**Number of cases**
38	Literal Translation with Flat intonation	4
	Literal Translation with Emphatic intonation	5
	Broad Focus	17
	Emphatic Particle/Syntax	7
	Ellipsis/Deletion of Nucleus	5
INTERVIEWS VIDEOS		
Number of Cases	**Strategy**	**Number of cases**
6	Emphatic Particle/Syntax	2
	Broad focus	2
	Ellipsis/Deletion of Nucleus	1

Number of Cases	Strategy	Number of cases
	Other (Literal translation WRONG emphatic intonation)	1
REALITY TV		
Number of Cases	**Strategy**	**Number of cases**
11	Literal Translation with Flat intonation	3
	Literal Translation with Emphatic intonation	1
	Broad focus	1
	Emphatic Particle/Syntax	4
	Ellipsis/Deletion of Nucleus	2
WEB VIDEOS		
Number of Cases	**Strategy**	**Number of cases**
6	Literal Translation with Flat intonation	5
	Ellipsis/Deletion of Nucleus	1

Table 33 Rendition strategies for narrow focus

The table shows that the strategy that all the voice-over categories share is "literal translation with flat intonation", which illustrates that voice-over has traditionally been delivered using a rather flat monotonous pitch melody.

If we now focus on those narrow focus examples with contrastive purposes, leaving deaccenting aside, we observe there are two rendition possibilities, as shown in the graph below: full rendition (52%) and partial rendition (semantic meaning only) (48%).

Contrastive focus rendition (25 cases)

52% | 48%

- Prosodic-Illocutionary only
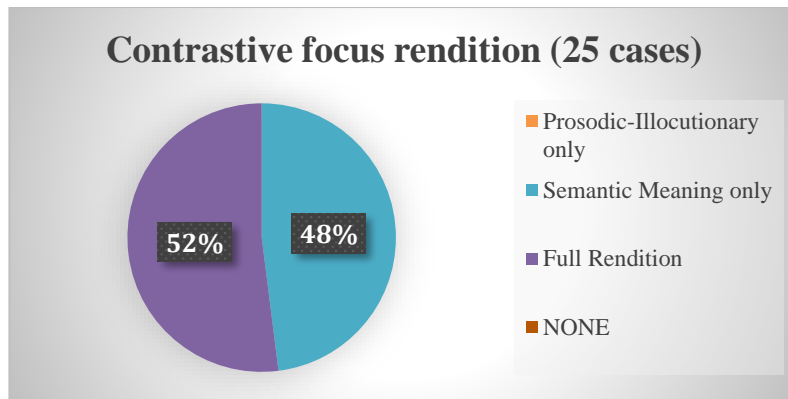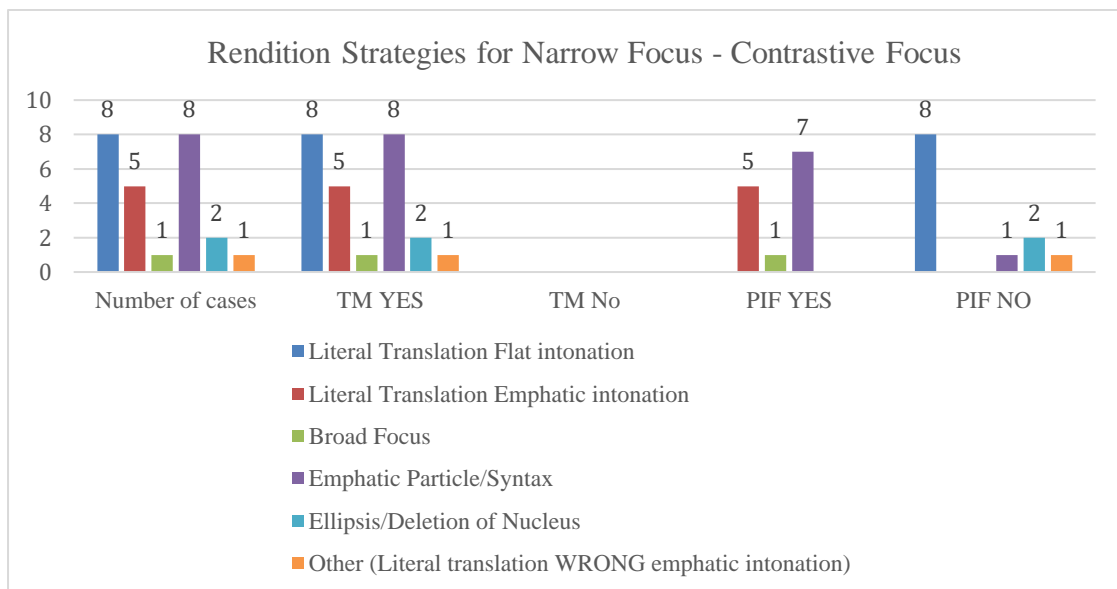- Semantic Meaning only
- Full Rendition
- NONE

Figure 39 Contrastive focus rendition success

The rendition strategies for these contrastive focus examples are the same as the ones shown in figure 39 above, being the only difference the number of cases for each strategy. Thus, we can observe that the use of syntax or emphatic particles to show contrast has been resorted to 8 times out of 25, while a literal translation of the original has been applied 13 times (5 showing emphatic intonation as in source text). In addition, broad focus appears only once, compared to the 19 times we found it in the deaccenting section.

Rendition Strategies for Narrow Focus - Contrastive Focus

- Literal Translation Flat intonation
- Literal Translation Emphatic intonation
- Broad Focus
- Emphatic Particle/Syntax
- Ellipsis/Deletion of Nucleus
- Other (Literal translation WRONG emphatic intonation)

Our analysis of the most relevant examples will start, as usual, with those showing full rendition and finish with those showing a partial type (semantic meaning only), and it will show instances of both narrow focus as a result of anaphora rule or deaccenting, and contrastive focus.

a) Full Rendition

There are some examples of narrow focus in our corpus which have already been analysed in other sections of our analysis, as is the case with CS.Ex1. This is due to the fact that, as has already been mentioned in this dissertation, intonation variables can be combined and do not need to appear in isolation; this is so in most cases of deaccenting, in which the deaccented material is placed in final position and the focus is consequently found in an earlier section of the sentence (see section 5.2.1.1).

N.F.Ex 1[84] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:54:48	00:54:53	Anne Boleyn	I swear upon the eternal damnation of my soul...	Juro por la eterna condena de mi alma,
00:54:55	00:54:56		..I have been a **true** wife...	que he sido una esposa **fiel**.

This example of narrow focus from the documentary *The last days of Anne Boleyn* shows a recreation of Anne Boleyn's words to the priest in her last confession before her execution. In her speech she insists that the has been faithful to the king, by highlighting the word "true" in her intonation, as can be seen in the pitch contour image below.

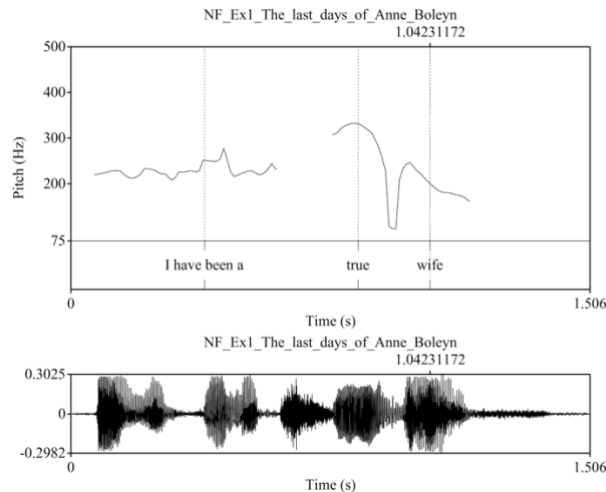[84] NF.EX stands for Narrow Focus Example.

Image 74 Pitch contour and sound wave in NF.Ex1

As we can see, there is narrow focus since the tonic syllable does not fall on the last lexical item ("wife") but on "true", and thus, Anne emphasises her truthfulness and fidelity to Henry VIII.

If we now check the Spanish translation for this passage, we can see that the strategy used has been to turn the narrow focus into broad focus, placing the nucleus at the end of the sentence ("fiel"). A literal translation in terms of word order (que he sido una fiel esposa) would not have worked as well as the syntactic reorganisation of the words shown in the target text since the focus would have fallen on "esposa/wife" rather than on "fiel/true". Therefore, this example shows full rendition of the original message (both semantic meaning and prosodic-illocutionary). However, we must say that the performance in the target text is not very faithful to the original. While the original voice sounds firm and convincing, the target text delivery sounds as if Anne Boleyn was crying.

N.F.Ex 2 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:11:04	00:11:11	David Starkey	It's... It's... This is **failure** for a king of a most terrible sort.	Es el **mayor fracaso** para un rey, de la peor especie.

This second example of narrow focus, also retrieved from the documentary T*he last days of Anne Boleyn*, shows David Sarkey commenting on Henry's difficulty to father a son after Anne's miscarriage of what would have been the King's first son. In his speech, Starkey highlights the fact that, for Henry, that was a catastrophe. And he does so by placing the nuclear accent on the first syllable of the word "failure" in the IP "| this is **failure** for a king |". The pitch contour can be seen below.
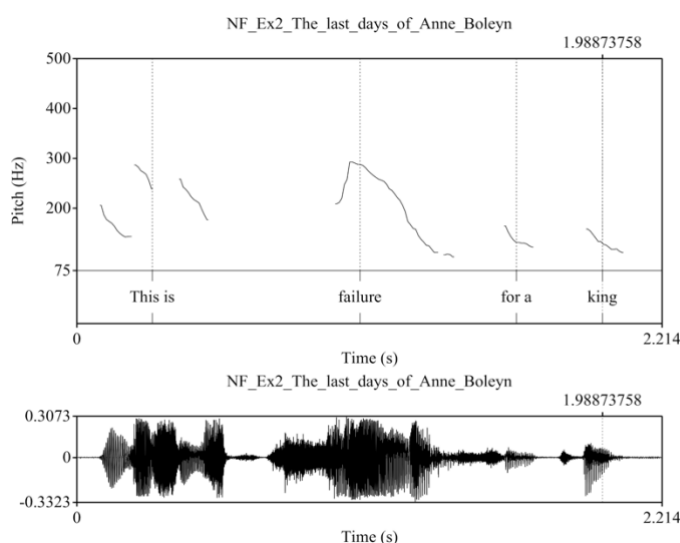

Image 75 Pitch contour and sound wave in N.F.Ex2

The image above shows how the word "failure" gets the nuclear accent with a high-fall tone on its first syllable, followed by a low-level tail that goes from its second syllable until the word "king", which is therefore deaccented. Thus, the speaker focuses our attention on "failure" rather than on "king", since "failure" is the main content of information in the sentence, while "king" is situationally given information in this case.

The Spanish translation for this passage shows the strategy we can describe as "emphatic particle /syntax" since, by including the adjective "mayor", the translator conveys through lexical means the intonational highlighting of the source text. In this way, the illocutionary force is kept in the target text and the semantic meaning of the original message is also rendered.

N.F.Ex 3 *Tasteology: Experience*				
In	Out	Character	Source text	Target text
00:11:54	00:12:34	Jozef Youssef	We don't just use **dry ice**, we **scent** the dry ice,	No solamente usamos el **hielo seco**, lo **aromatizamos**,

This example from the documentary *Tasteology: Experience* shows narrow focus with deaccenting. Chef Jozef Youssef is talking about cooking ingredients and how to use them. He comments on dry ice, which is accented in the first line, but not on the second one, where the nucleus falls on the preceding content word (scent) resulting in default accent. The pitch contour of this example can be seen below:
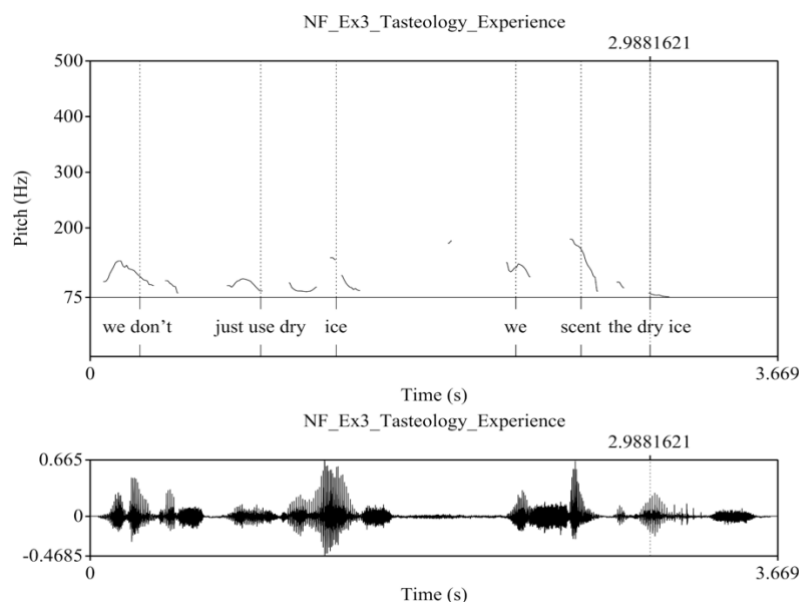
Image 76 Pitch contour in N.F.Ex3

In the Spanish text for this phrase, the original deaccented material has been deleted by means of ellipsis, so we just get the nucleus preceded by the article working as a "noun"[85]: "| lo aromatizamos |". Thus, both texts highlight the same information through different means, so they are pragmatically equivalent since the prosodic-illocutionary force is the same in both versions.

---

[85] For Alarcos (1994: 82), "Al sobreentenderse un sustantivo eliminado, el adjetivo o las otras palabras precedidas del artículo cumplen en el enunciado funciones propias del sustantivo y quedan así sustantivadas."

N.F.Ex 4 *Tastelogy Experience*				
In	Out	Character	Source text	Target text
00:09:11	00:09:36	Christine Flynn	But I **love** texture, like, I've always been a big texture girl,	Pero me **encanta** la textura. Siempre me ha apasionado la textura,

This fourth example, also found in the documentary *Tasteology: Experience*, shows chef Christine Flynn talking about her cooking preferences. In this speech, she says that she is very fond of texture, emphasising the word "love" in the IP "| but I \love texture |". Thus, "love" is in narrow focus since it receives the nuclear accent instead of "texture", which would have normally received it, being the last lexical item. The melodic pattern for this IP is as follows:
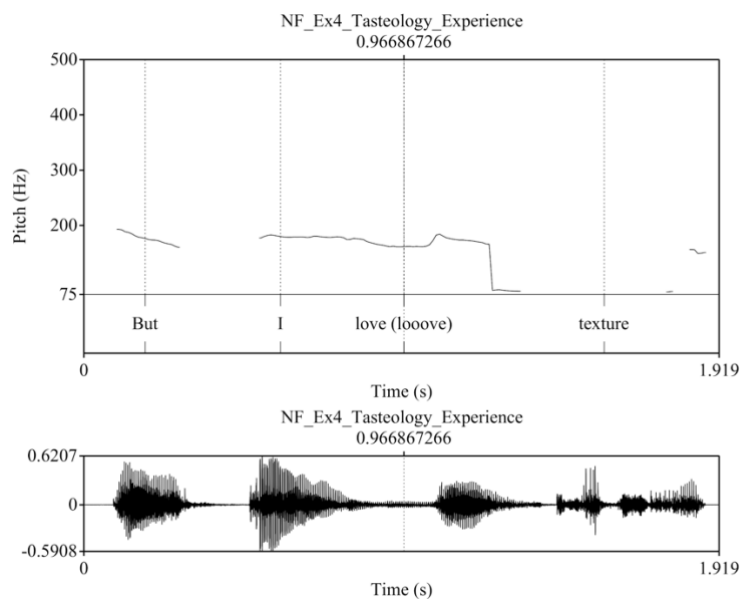
Image 77 Pitch contour in N.F.Ex4

The Spanish text for this passage has been translated literally, word by word: "pero me encanta la textura". With this structure the nucleus falls on the last word ("textura"), as usually happens in Spanish. However, the delivery of the line shows "encanta" being highlighted with an emphatic tone, therefore brought into focus, as in the original source text. The image below shows the pitch contour for this line.
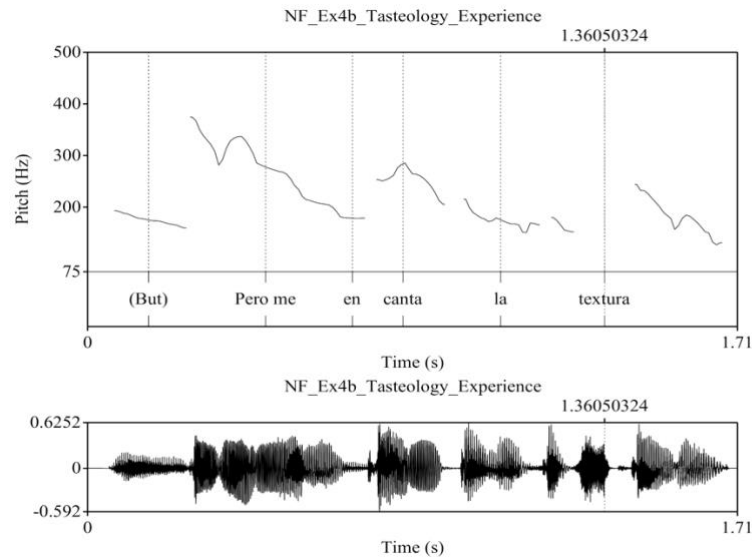
Image 78 Pitch contour in the target text of N.F.Ex4

As we can see, the melodic line in the image above fluctuates, which is a characteristic of certain voicing techniques, known as *dubbitis*, as has already been mentioned. This intonation pattern used in this phrase is not very common in colloquial or natural speech in Spanish due to its roller-coaster-like pattern and its high pitch, as stated by Estebas-Vilaplana (2014a: 182). Despite the fact that "pero" shows a higher pitch level, probably due to the initial flow of air in the delivery, the stressed syllable of "encanta" shows an emphatic peak in its melody, highlighting it more than the others in the utterance (more than "textura", which also shows a melodic rise), thus attracting the audience's attention. In any case, this is in line with RAE's (2011: 484) focalising (or integrating) function of intonation in Spanish, which allows to hierarchise the different items of information within utterances and seems to highlight the same word as in the original English text. Nonetheless, although we understand that it shows a full rendition of the original English text, we consider that it does not sound natural in Spanish. It seems a clear example of interference caused by the original audio, which resulted in an unnatural delivery in the target version.

N.F.Ex 5 *Kitchen Nightmares*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:18:55	00:19:07	Gordon Ramsay	What I want you to do, you two, is to confirm that you **do** know the basics but, at the same time, cook me a fucking omelette.	Lo que quiero que hagáis es que me confirméis que **sabéis** lo básico. Me vais a cocinar una tortilla los dos a la vez.	00:18:39	00:18:55

This first last example of full rendition, from the reality show *Kitchen Nightmares*, shows chef Gordon Ramsey telling off the chef and sous-chef of the restaurant he intends to help. In his speech, he uses the emphatic particle "do", which also works as the nucleus in the IP "| just to confirm that you \do know the basics |" in order to highlight that the two chefs know, indeed, how to cook an omelette.
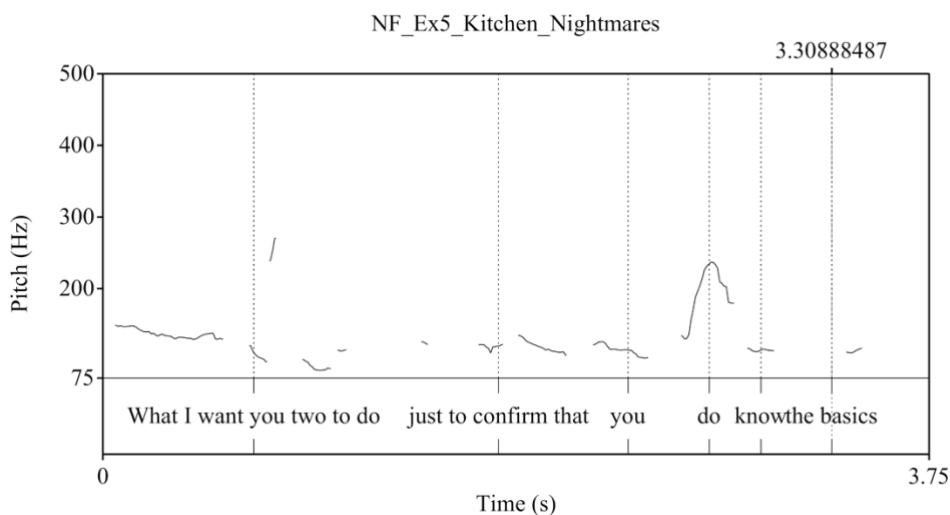

Image 79 Pitch contour in N.F.EX5

As we can see, the *dummy operator* "do" (Leech and Svartvik 1994: 328) shows a pitch rise the second time it appears ("you do know"), thus highlighting the emphatic pragmatic meaning within the sentence.
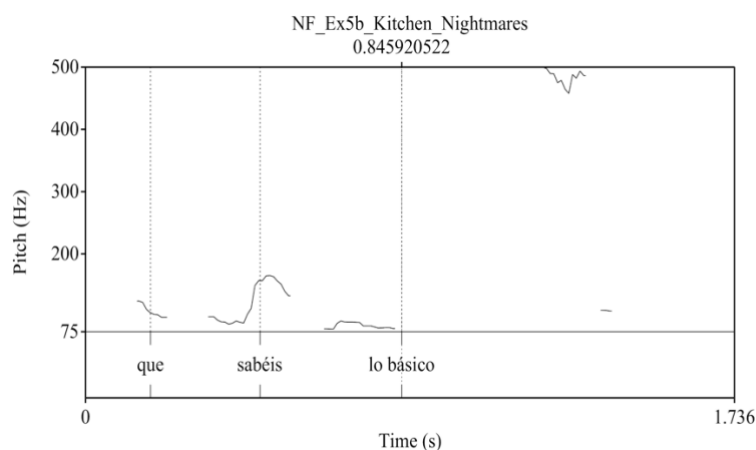
Image 80 Pitch contour in the target text of N.F.Ex5

The melodic line of the Spanish text rises on the accented syllable of "sabéis" which brings it into focus, and then falls until the end of the utterance. In addition, there is no emphatic particle or structure to highlight the meaning of the verb, as would have been the case with the particles "sí" or "claro que" ("Lo que quiero que hagáis es que me confirméis que **sí / claro que** sabéis lo básico"), or by placing the verb at the end ("Lo que quiero que hagáis es que me confirméis que lo básico (sí) lo sabéis"). Nonetheless, what we get is the intonational highlighting in the delivery of the verb "sabéis", in a similar fashion to what happened in the previous example (NF.Ex5). Again, it seems that the original audio track has interfered making the Spanish voice talent imitate a melodic pattern which is more usual and natural in English than in Spanish. Had the translator resorted to a more natural highlighting device, such as an emphatic particle or structure, probably, the voice talent would not have used that unusual intonation. Consequently, the pragmatic meaning of the original is kept, fully rendering the original message of the source text, although we consider that there are more suitable and natural options in Spanish to render this type of emphasis.

N.F.Ex 6 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:22:56	00:23:01	Hilary Mantel	He has begun to complain, way back in 1534,	En 1534, el rey empezó a decir

00:23:01	00:23:04		that Anne is turning against him.	que Ana se había puesto en su contra.
00:23:04	00:23:10		He's said that Anne's threatened him, he said Anne wants his head.	Que lo había amenazado, y que quería su cabeza.
00:23:10	00:23:16		By 1536, the conflict is ready to explode.	Dos años después, el conflicto estaba a punto de estallar.

This example from the documentary *The last days of Anne Boleyn* is the only one showing full rendition when the nucleus of the original text has been deleted or reformulated in the target text. As we can see in the sentence "By 1536", the nucleus is 6 and it is narrow focus because the previous numbers are anaphoric (we have "1534" three lines before) and therefore deaccented. This is a clear example of narrow focus and neutral tonicity, which is rather unusual. The image below shows the pitch contour for this IP.
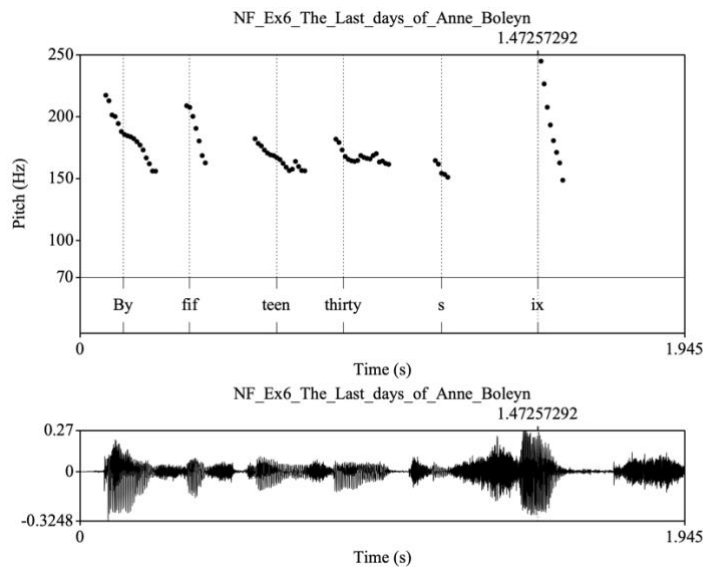

Image 81 Pitch contour in NF.Ex6

The Spanish translator has resorted to modulation to render this passage with "Dos años después" (two years later), which, in our view, is a really good option to render the semantic meaning and illocutionary force of the original since "des**pués**" is stressed on the last syllable just like "153**6**". In addition, since both tonic syllables

coincide, there are no synchronic problems, and the target text is in harmony with Mantel's head movement when she utters this passage. There is a short pause just after which highlights even more the illocutionary force of this short IP.

b) Partial Rendition

N.F.Ex 7 *Teen Mom*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:00:47	00:00:51	Farrah	OK, well, anyways, you made the price, Simon. **You** were talking to them.	Vale, pero tú negociaste el precio. **Tú** hablaste con ellos.	00:02:09	00:02:13

This first example of narrow focus with partial rendition from the reality show *Teen Mom* (clip 1), shows the starring character Farrah Abraham talking to her fiancé and a couple therapist about her engagement ring. In her speech she wants to make clear that it was Simon, her fiancé, who negotiated the price of the ring with the jeweler and not her. In order to do so, she uses narrow focus and places the nucleus in "you" in the IP "| \you made the price Simon |". The pitch contour for this IP can be seen below.
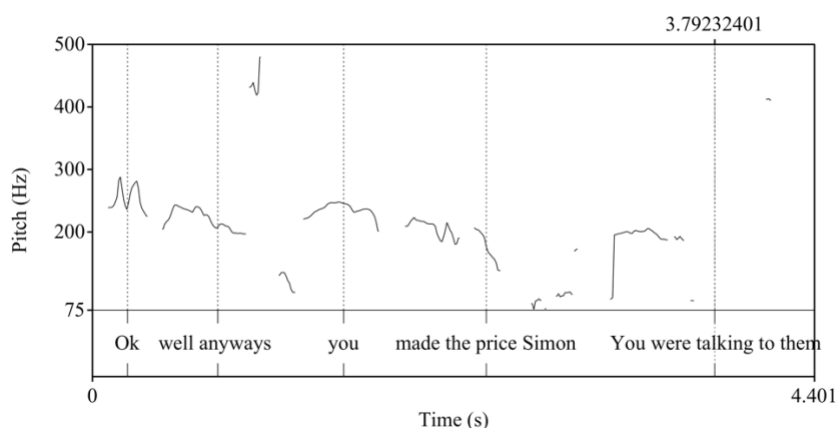

Image 82 Pitch contour in N.F.Ex7

As we can see, there is a high-fall tone in the melodic line which starts on "you" and then falls all the way until "Simon", which, as a final vocative, is deaccented and forms the tail of the unit (Wells 2006: 153). In addition to this, we can hear (and see in the pitch contour) that there is a slight pause after "you" (which is also slightly elongated in its pronunciation) in order to mark it even more.

The Spanish translation for this line shows literal translation (almost word by word) of the source text with flat intonation. This time there is no emphatic intonation to mark the nucleus ("tú"), as there was in the two previous examples analysed, but a fluctuating melodic line with no emphasis on "tú".
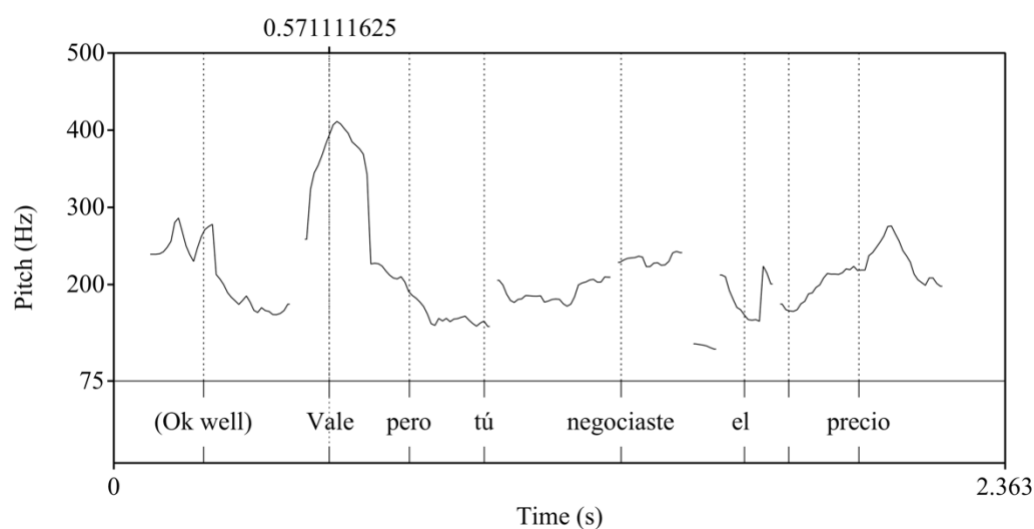
Image 83 Pitch contour in the target text of N.F.Ex7

There are, however, intonation peaks in "vale", the last syllable of "negociaste" and the last syllable of "precio", which, in our view, denotes *dubbitis* symptoms. This contrasts with the little emphasis that is given to "tú", which corresponds to the nucleus in the original text and should have received some kind of focus, therefore, in the target text too. One alternative to highlight it would have been to place it at the very end of the sentence ("Vale, pero el precio lo negociaste tú").

N.F.Ex 8 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:45:53	00:45:55	Greg Walker	To imagine the death of the King is technically a crime,	Imaginar la muerte del Rey es técnicamente un delito,
00:45:55	00:45:58		and is politically a **disastrous** thing to do.	y algo catastrófico en términos políticos.

This example from documentary *The last days of Anne Boleyn* shows Greg Walker talking about Anne Boleyn's conspiracy to end with King Henry VIII, which, in his view, was fatal for her career at court. It was, allegedly, such a bad idea, that Walker emphasises the word "disastrous" in his speech, in the IP "| and is politically a disastrous thing to do |". The pitch contour for this IP is shown below.
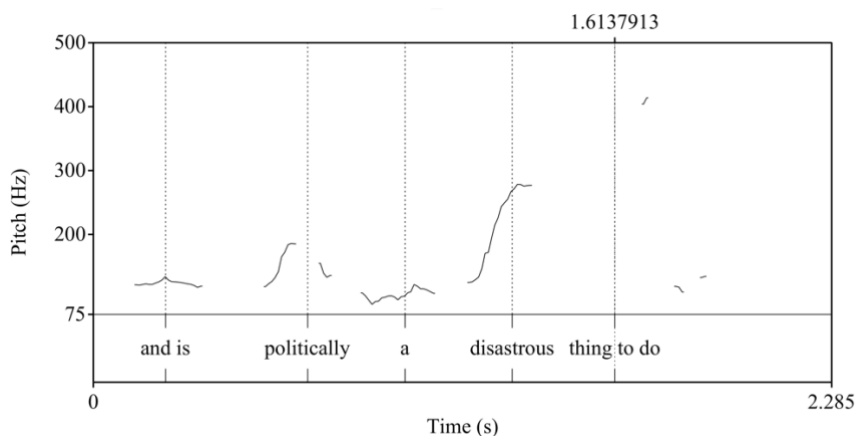
Image 84 Pitch contour in N.F.Ex8

There is a high-fall nuclear tone on "disastrous" that goes as high as 300Hz in terms of pitch. The tail "di **sa** strous thing to do" is so low that we cannot even see it on the image. This shows how much Walker wants to highlight the nucleus of the IP.

As far as the translation of this passage is concerned, the word "catastrófico" does not appear at the end of the sentence, which would bring it into focus; nor is it highlighted by means of intonation, as happened in other examples we have seen. The pitch contour for this passage is shown below.
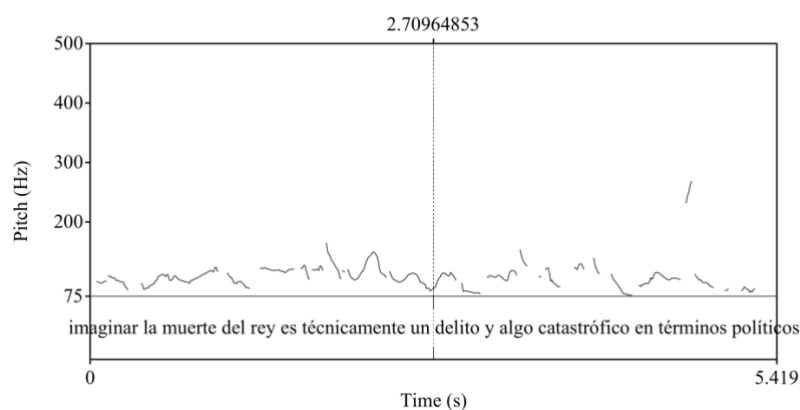
Image 85 Pitch contour in the target text of N.F.Ex8

The melodic line shown above is constant and there is no emphatic pitch movement on any word to mark it as pragmatically relevant in this passage. Consequently, we understand that the rendition of the illocutionary force from the source text has not been successfully achieved in this example, although the semantic content has.

| In	Out	Character	Source text	Target text
N.F.Ex 9 *The last days of Anne Boleyn*				
00:22:48	00:22:53	David Starkey	And to do it **in** front of the King, **in** the Chapel Royal,	Y hacerlo delante del rey, y en la **mismísima** Capilla Real,
00:22:53	00:22:56		surely, it's a declaration of war.	fue una declaración de guerra en toda regla.

This example from the documentary *The last days of Anne Boleyn* shows David Starkey talking about John Skips's sermon at the Chapel Royal on the Passion Sunday service 1536. John Skip was Anne Boleyn's personal chaplain and the sermon he gave was addressed at Cromwell, Henry's chief political counselor. This was, for Starkey, a very serious matter because Skip was criticising and condemning Cromwell, one of the most powerful men in England, in front of everyone at court. In his speech, Starkey emphasises where the scene took place and the presence of the king, by placing the nucleus on the preposition "in" in the IPs "| **in** front of the King | **in** the Chapel Royal

|" In this way, Starkey highlights that John Skip gave his sermon directly before the king himself and in his majesty's private church. The IP pitch contour is shown below.
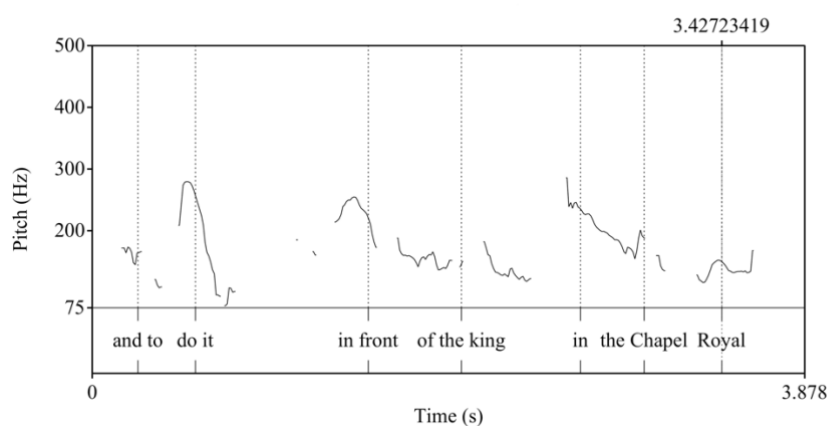
Image 86 Pitch contour in N.F.Ex9

In the Spanish translation for this passage, this emphasis has been rendered only partially. First, we can note that, in the second IP, the emphatic particle "mismísima" has been added in order to highlight the relevance of the place where the sermon was given, which is perfect to convey emphasis. However, the first IP shows no emphatic particle or intonation to mark the presence of the king before the priest, so the original emphasis is lost. The adjective "propio" would have been a good option to emphasise the presence of the king: "y hacerlo delante del **propio** rey, y en la **mismísima** Capilla Real"; and the original emphasis of both IPs would, thus, have been kept in the target text.

In addition to this, the intonation pattern used in these two IPs is confusing because, as can be seen in the image below, the first IP ends with rising intonation, showing that more information is yet to come, which is in fact the case that happens just after; but the second one shows falling intonation as if that was the end of the communicative load, when, actually, the sentence continues and the information, as we know from the original version, has not yet been delivered in full ("fue una declaración de guerra en toda regla").
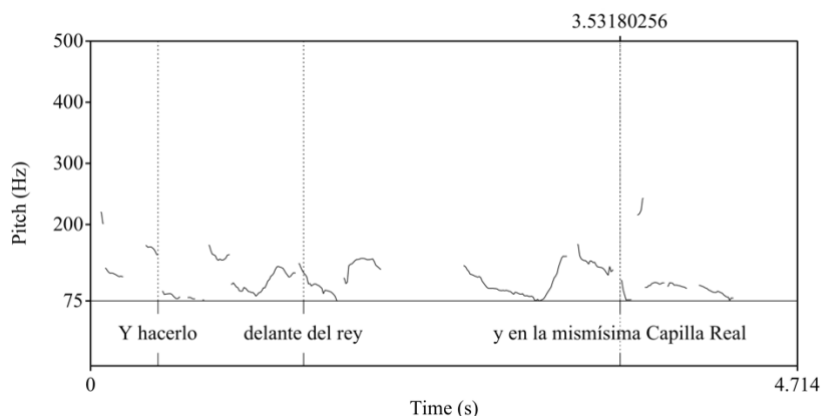
Image 87 Pitch contour in the target text of N.F.Ex9

This melodic passage sounds confusing because of the melodic pattern chosen in the delivery turns this passage confusing in the target version of the documentary. A more appropriate delivery here would have been with rising intonation in the second IP to contrast with the final fall at the end of the sentence (Alarcos 1994: 65-66), which is not shown here

N.F.Ex 10 *TN Europe Jan 17*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:02:33	00:02:38	Francis Ellison	But how does this leave our portfolios with a quality bias?	Así pues, ¿qué significa todo esto para nuestras carteras orientadas a la calidad?	00:03:26	00:03:30
00:02:39	00:02:45		That at least is a question we **can** answer –.	Esta es, al menos, una pregunta que podemos **responder**:	00:03:31	00:03:40

In this example from the Columbia Threadneedle web video *TN Europe Jan 17* showing the updates on European equities for 2017, the presenter Francis Ellison, the

Client Portfolio Manager of the firm, uses narrow focus by means of the strong form[86] of the auxiliary verb *can* in the IP "| that at least is a question we **can** answer |". In the lines coming just before the ones we are analysing, Ellison poses a series of questions concerning the economic forecast for the year, which are left unanswered. By using narrow focus in this line, he signals to the audience that his firm has, indeed, the answer to the last question. The IP pitch contour is shown below.
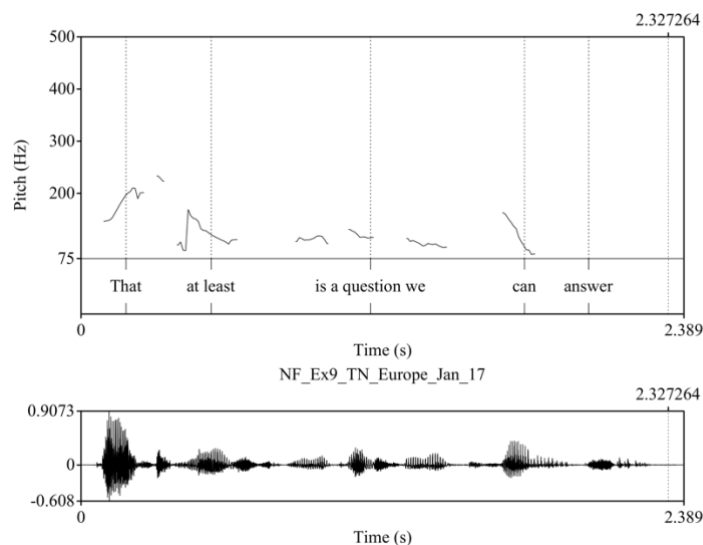
Image 88 Pitch contour and audio wave in N.F.Ex10

There is a high fall nuclear tone over "can" with a low level tail ("answer"), which is so low that can only be seen in the audio wave below the melodic line. The pitch contrast is evident, since it goes from an average 100Hz in the previous words to 200Hz on "can" to mark it as the tonic syllable in the utterance. By highlighting the modal verb using its strong form rather that the main verb ("answer"), the focus falls on the positive value of the answer to the question, rather than on the fact of answering it, as would have been the case if the weak form of the modal had been used.

In the Spanish translation there is no emphatic particle to contrast and highlight the positive meaning of the message e.g. "sí" as in: "esta es, al menos, una pregunta que sí podemos responder". Moreover, the melodic pattern used in the delivery of this

---

[86] Function words in English have two possible pronunciations: weak and strong. Weak forms are pronounced using the weak vowel schwa /ə/, as in "can" /kən/ while strong forms have strong vowels as "can" in our example /kæn/

line is rather flat and does not show any kind of intonational highlighting, as can be seen in the pitch contour image shown below:
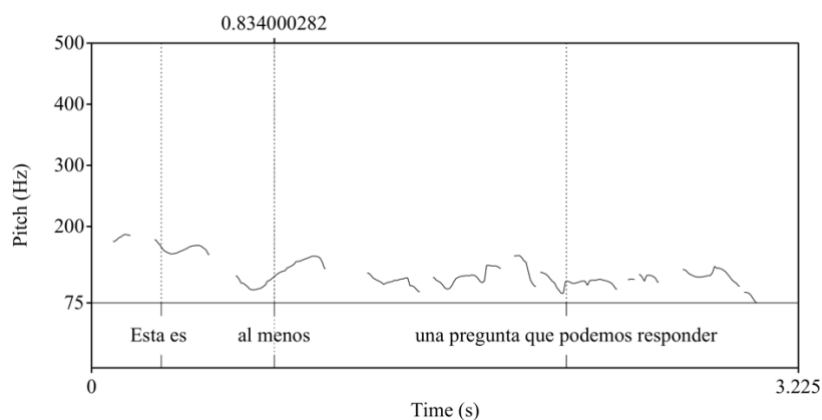
Image 89 Pitch contour in the target text of N.F.Ex10

Therefore, the illocutionary force implicit in the original message by means of narrow focus has not been successfully rendered in the target text, since there is no contrast between the unanswered questions previously asked and this one, which is, in fact, highlighted by the presenter in the source text both with the narrow focus and with the actual answer provided.

N.F.Ex 11 *TN Europe Jan 17*						
ST in	St out	Character	Source Text	Target Text	TT in	TT out
00:00:58	00:01:03	Francis Ellison	But it doesn't mean that we can **ignore** these issues, as they will have a big impact on markets and on individual stocks.	Esto no significa que podamos ignorar estos **asuntos**, pues tendrán una notable repercusión en los mercados y en los valores individuales.	00:01:13	00:01:20

This last example of partial rendition has also been found in the web video *TN Europe Jan 17.* Francis Ellison is now wondering about several matters related to Brexit and Trump's isolationism policies, which, apparently, cannot be solved yet but cannot

be obviated either since they will have consequences for the stock market in the near future. Ellison highlights his last thought by placing the nucleus on the verb "ignore" in the IP "| But it doesn't mean that we can **ignore** these issues |" rather than placing it on the last word ("issues"), thus resulting in narrow focus. The reason for highlighting the verb "ignore" and not the last lexical item "issues" is probably due to the fact that, first, the speaker wants to emphasise the action of ignoring, and second, because "issue" belongs to a group of words termed *empty words* in English phonetics and phonology, such as "things", "people", "places", etc., which tend not to receive the nucleus even though they may be the last lexical item within an IP (Wells 2006: 150).

The melodic contour for this example is shown below.
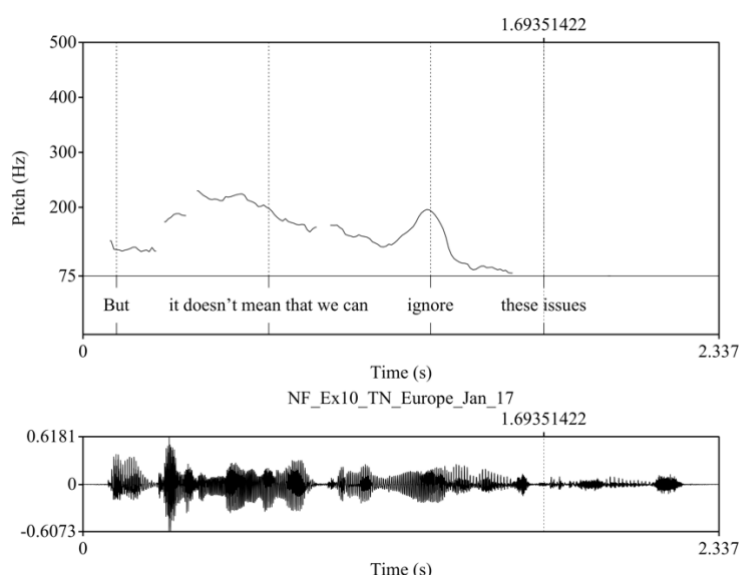
Image 90 Pitch contour in N.F.Ex11

The IP structure is clearly illustrated in the pitch contour melody, with an obvious pitch variation on the second syllable of "ignore" which receives an evident high-fall nuclear tone. Unlike the previous example, this one shows the weak form of the auxiliary verb "can", which is, therefore, unstressed. The tail of the IP ("these issues") is so low that we can hardly see it, as in NF.Ex10, but is perceivable by the audio wave graph below the melodic line.

The Spanish translation for this passage is literal, almost word by word, with the last words of the IP "estos asuntos" corresponding to the English equivalents and in

the same position "these issues". There is no insistence stress to mark the verb "ignorar", as can be seen in the pitch contour image below, which is rather stable and shows no relevant pitch variation. This, in our view, means that the focus is broad, the nucleus falling on the last word of the IP "asuntos", which does not match the original speaker's intention as there is no contrast now.
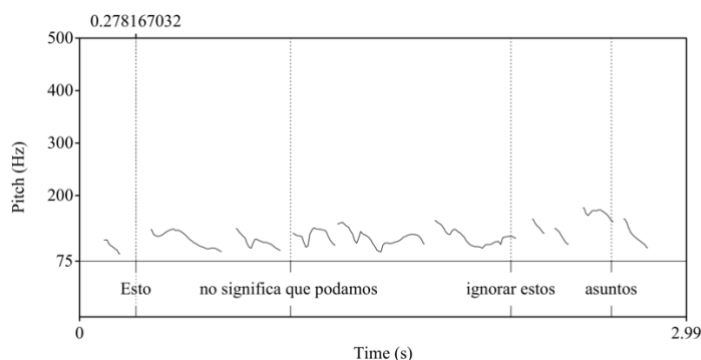
Image 91 Pitch contour in the target text of N.F.Ex11

A possible solution to match the original prosodic-illocutionary force, could have been to place the verb "ignorar" at the end, as in: "Esto no significa que estos asuntos se puedan ignorar", or "Esto no significa que estos asuntos puedan ignorarse". This way, the word receiving the nucleus would be the same in both texts (source and target) and the pragmatic meaning would have been preserved with very natural means in Spanish.

5.2.5.1.1 Concluding remarks on the rendition of narrow focus

There are several ways of rendering in Spanish the informative contrast that English offers by means of narrow focus. As we have seen in examples N.F.Ex1 and N.F.Ex3, word order can help to match the nucleus of the source text by placing the highlighted word at the end of the IP in the target text. As we know, Spanish has a strong tendency to place the new information at the end of the sentence (Gutiérrez Ordóñez 2014: 21), so this is probably the most pertinent way to achieve focus in Spanish (ibid, 29). Nonetheless, there are other ways of highlighting an item in Spanish that have also been used in our corpus to match the original prosodic-illocutionary force, such as emphatic

particles (see N.F.Ex2, N.F.Ex9 and C.S.Ex1) and emphatic intonation (see N.F.Ex4 and N.F.Ex5). This last strategy might have been the most recurrent one in voice actors' delivery, since most of the narrow focus examples in our corpus show literal translation (almost word by word) of the original text, which means that the word in focus will not be at the end of the sentence. However, this has not proved to be the case.

The intonation used by the voice actors to deliver the target text can sometimes play against the desired result of the script translation, since by using too much emphasis or fluctuating melodic patterns (aka *dubbitis*), the rendition of the original pragmatic meaning will not be fully conveyed. English and Spanish use opposite pitch levels for emphasis, as stated by Estebas-Vilaplana (2014a: 182), which means that an over the top intonation in Spanish will not be too credible by the audience of this language. In addition, a roller-coaster melodic pattern can emphasise unwanted words and bring them into focus when that is not the case in the original source text, or indicate the message is complete when it is not (as in N.F.Ex9).

All in all, there are diverse strategies to render narrow focus in Spanish voice-over, and it depends on the translator's ability to identify where the focus of the information is in the source text and to choose the best option to render it in the target text.

5.2.6 Synchrony

Now that we have finished the analysis of the linguistic features, is time to delve into the technical features that characterise voice-over, which have been described in section 1.3.3: isochrony, literal synchrony, content synchrony, kinetic synchrony, and character synchrony. In this section, we will only comment on those examples that show no synchrony, since these are the most relevant to our study.

5.2.6.1 Isochrony

In audiovisual translation, mostly in dubbing and voice-over, isochrony means that the translated voice has to match the original voice of the source text in terms of length, or as Chaume (2008: 134) puts it: "equal duration of utterances". As we saw in chapter 1 of this work when we described the main characteristics of voice-over translation, the synchrony between the two voices (original and target) may be modified due to sound bites (the time span 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) and have different durations (see section 1.3.3). We will try to shed some light on this issue and see how isochrony is applied in the different genres under analysis in this work.

5 cases have been identified in our corpus in which isochrony is out of sync; 1 belongs to documentaries and 4 to interviews videos (all in *iTalk: Is it time to scrap the CAP?*). There are no isochronic issues in reality shows and web videos in the programmes analysed here.

IS.Ex 1[87] *Tasteology: Chill*				
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…	

The line belongs to a woman who is standing on a stage and talking to the audience about food. It lasts just a few seconds and it is not really relevant to the story of the documentary. However, considering that she is the only speaker on screen, we consider that her lines should have been translated, but as we can see, they have not, so we can say that there is no sync whatsoever.

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

[87] IS.Ex stands for Isochrony Example.

IS.Ex 2 *Euronews: Is it time to scrap the CAP?*				
In	Out	Character	English Original	Spanish Translation
00:35	00:36	Daniel Hannan, male	Thank you.	

As happened in the previous case, there is no translation for this utterance, so there is no sync.

IS.Ex 3 *Euronews: Is it time to scrap the CAP?*				
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 neighboring 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.

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. There is a translation for what the interviewer says (see below) but not for Mr. Hannan's words.

IS.Ex 4 *Euronews: Is it time to scrap the CAP?*				
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?

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 2 people talking (so a third voice would probably be too much). Therefore, the translation is only provided for the second question.

IS.Ex 5 *Euronews: Is it time to scrap the CAP?*				
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?

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

Most of these examples of no isochrony are due to the fact that no target text has not been included for a particular portion of source text, so we only get the original voice in the translated version of the programme. In addition, there is just one example in which the target version lasts more than the original.

In terms of sound bite length, we have noticed that the duration suggested by some scholars in section 1.3.3, and which is shown in the table below, is not present and/or consistent all the time in the programmes of our corpus.
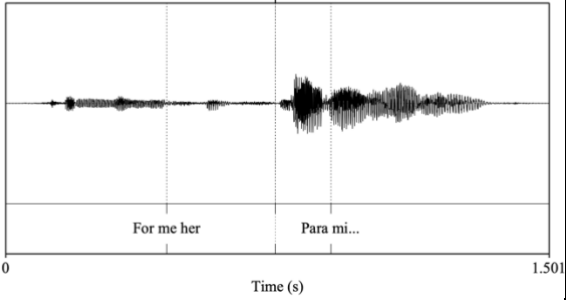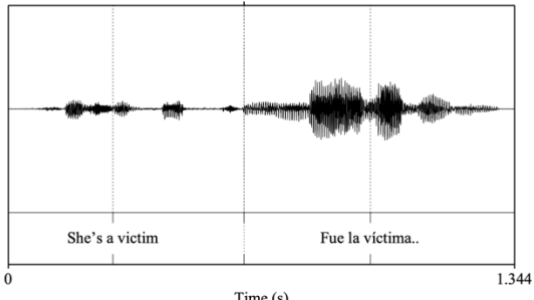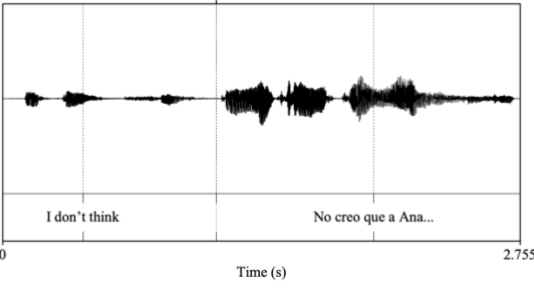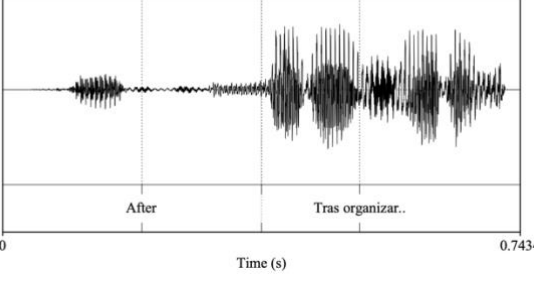
AUTHOR	SOUND BITE LENGTH
Orero (2004: 83)	"Some two seconds before and after".
Espasa (2004: 189)	"Two seconds before and after, or three or four words."
Mayoral (2001: 03)	"Three or four words."
Agost & Chaume (1999: 259)	"About three seconds at the beginning but both languages finish at the same time in the case of documentaries."
Darwish & Orero (2014: 143)	"A few words at the beginning, and usually original and translation finish at the same time in the case of voice-over in news broadcasts."
Chaume (2012: 03)	"The translation is heard a few seconds after the original voices."
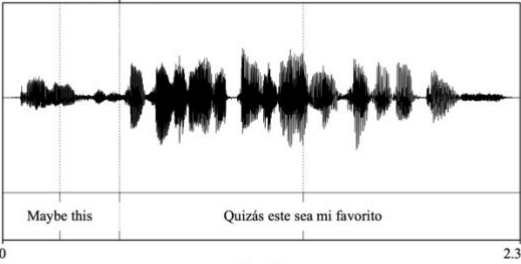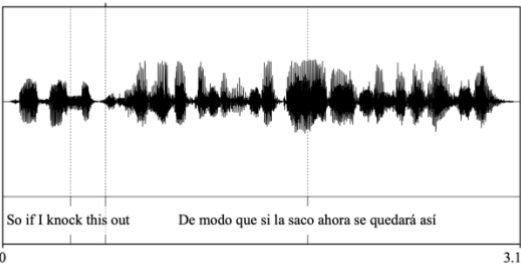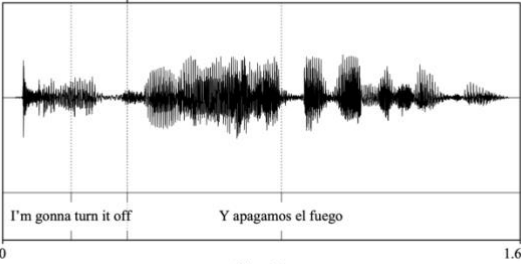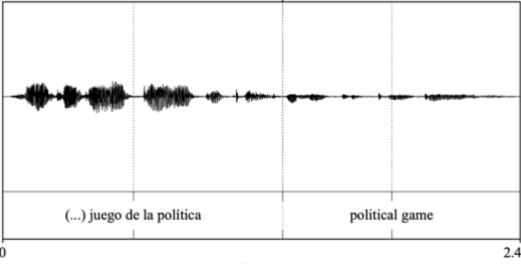
Table 34 Duration of sound bites according to different authors.
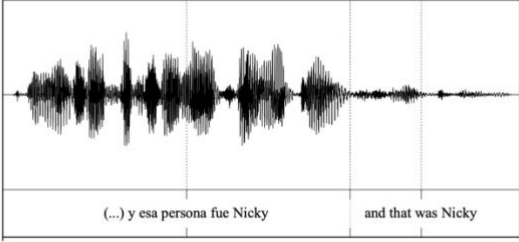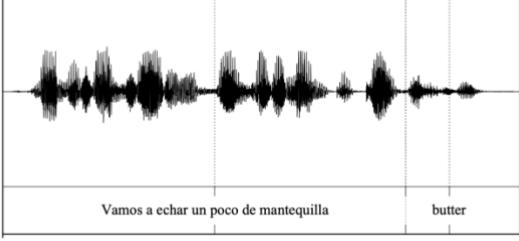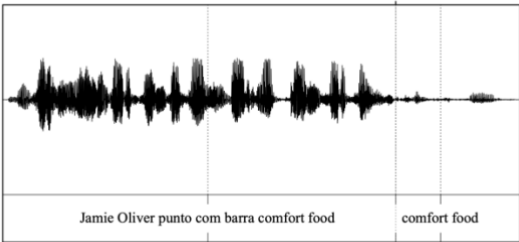
There seems to be no rule of thumb regarding the duration of the sound bites, which may depend on two things: the length of the target script to be voiced; and the pace used by the original speaker. Thus, if we take, for example, the documentary *The last days of Anne Boleyn,* we can notice that sound bite length differs almost every time depending on who speaks on camera.
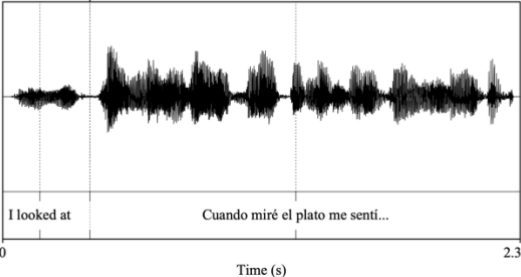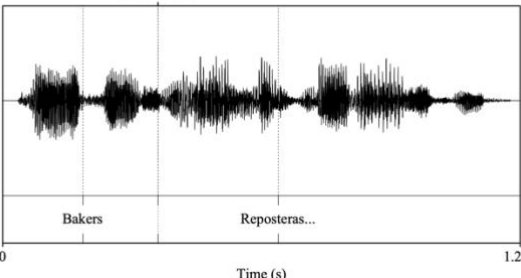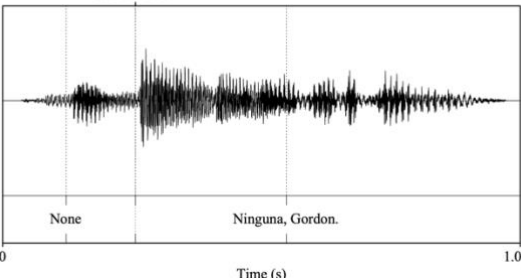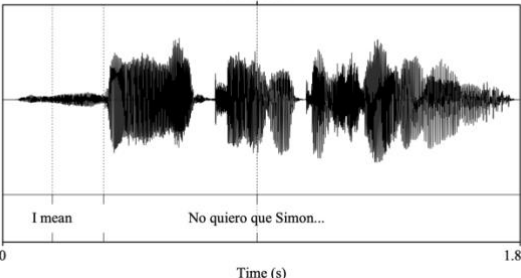
The table below shows some audio bites (at the start and at the end of certain phrases) in terms of audio wave length and the words used for each one.
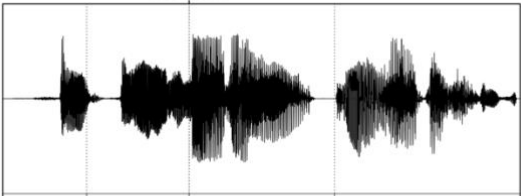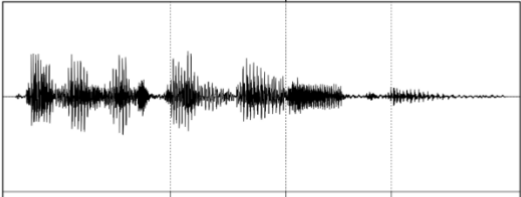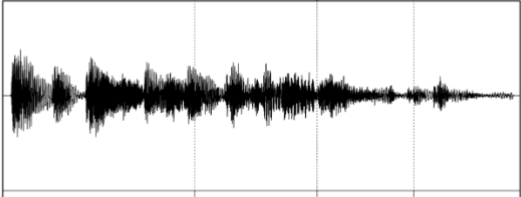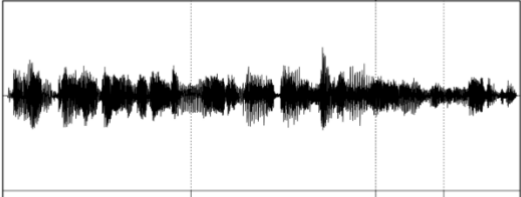
SOUND BITES		
A) DOCUMENTARIES		
STARTING SOUND BITES		
AUDIO WAVE	TIME FRAME	NUMBER OF
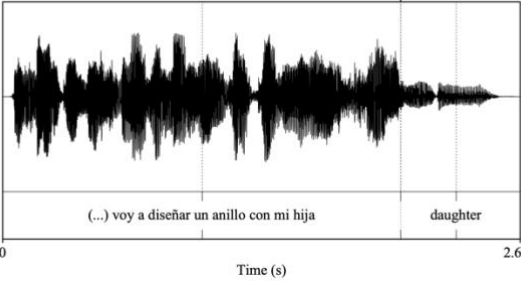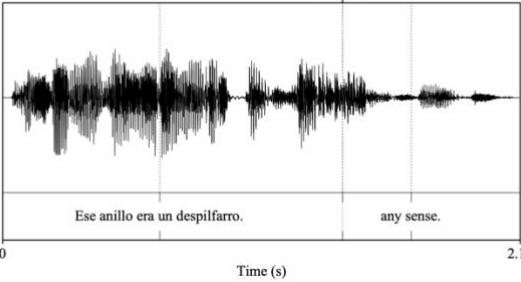
		WORDS
Sound_bite_start_1__The_last_Days_of_Anne_B 0.744463593 For me her / Para mi...	0.75 SEC.	3 WORDS
Sound_bite_start_2__The_last_days_of_Anne_B 0.62587614 She's a victim / Fue la víctima..	0.62 SEC.	4 WORDS
Sound_Bite_start_3__The_last_days_of_Anne_B 1.13728284 I don't think / No creo que a Ana...	1.13 SEC	3 WORDS
Sound_bite_start_4__Tasteology_Chill 0.371689342 After / Tras organizar..	0.37 SEC	1 WORD
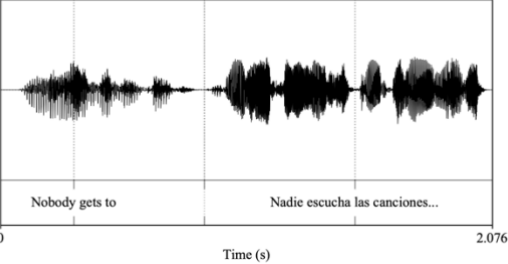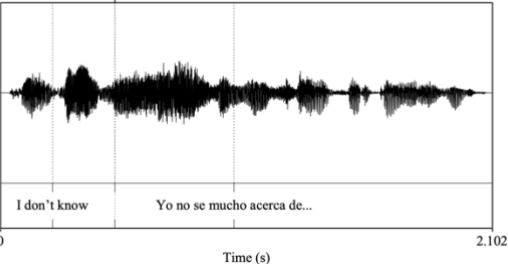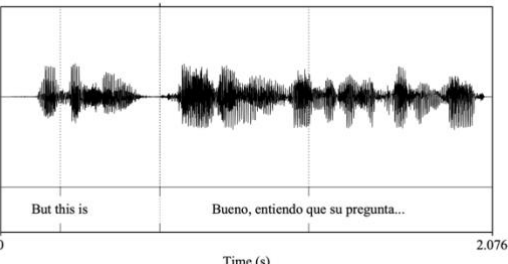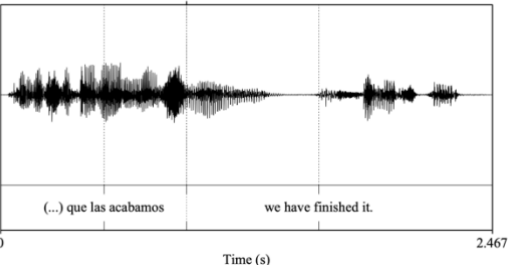
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_start_5__Tasteology_Chill 0.539247932 Maybe this — Quizás este sea mi favorito 0 — 2.389 Time (s)	0.53 SEC.	2 WORDS
Sound_bite_start_6__Jamies_Comfort_Food 0.635040187 So if I knock this out — De modo que si la saco ahora se quedará así 0 — 3.199 Time (s)	0.63 SEC.	6 WORDS
Sound_bite_start_7__Jamies_Comfort_Food 0.405331829 I'm gonna turn it off — Y apagamos el fuego 0 — 1.684 Time (s)	0.40 SEC.	5 WORDS

<div align="center">

ENDING SOUND BITES

</div>

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_end_1__The_last_days_of_Anne_B 1.34903181 (...) juego de la política — political game 0 — 2.494 Time (s)	1.14 SEC	2 WORDS
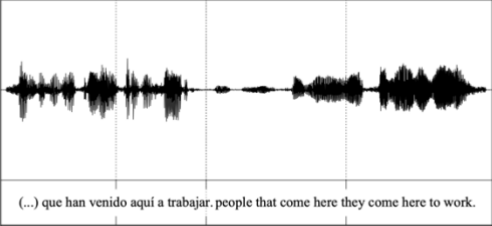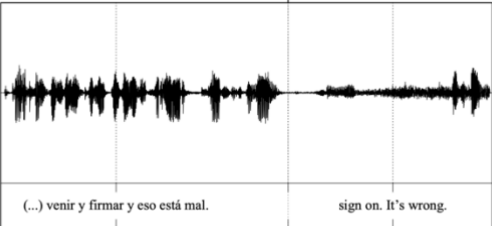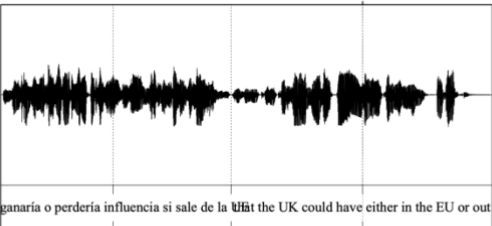
Sound_bite_end_2__Tasteology_Chill 1.25296838 (...) y esa persona fue Nicky · and that was Nicky 0 — Time (s) — 1.867	0.61 SEC.	4 WORDS
Sound_bite_end_3__Jamies_Comfort_food 1.71770562 Vamos a echar un poco de mantequilla · butter 0 — Time (s) — 2.206	0.49 SEC.	1 WORD
Sound_bite_end_4__Jamies_Comfort_Food 2.68667525 Jamie Oliver punto com barra comfort food · comfort food 0 — Time (s) — 3.538	0.85 SEC.	2 WORDS
AVERAGE START SOUND BITE IN DOCUMENTARIES		
	0.63 SEC.	3.4 WORDS
AVERAGE ENDING SOUND BITE IN DOCUMENTARIES		
	0.77 SEC.	2.2 WORDS
B) REALITY SHOWS		
STARTING SOUND BITES		
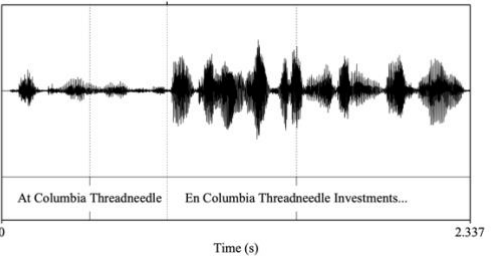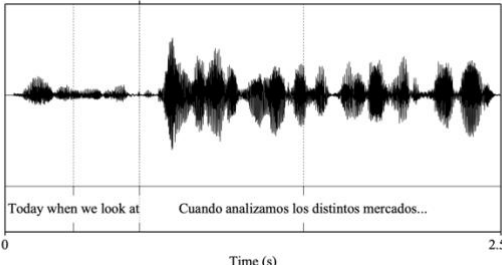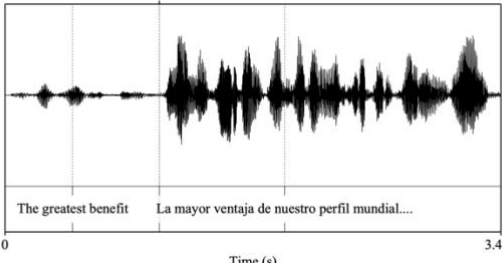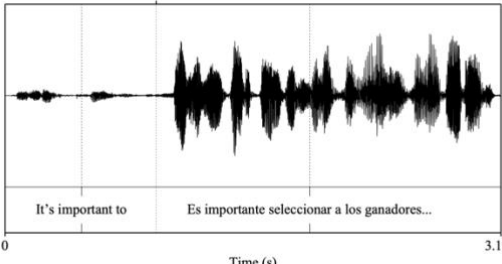
AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_start_8__Cupcake_Wars 0.402253704 I looked at　　　Cuando miré el plato me sentí... 0　　　　　　　　　　　　　　2.389 Time (s)	0.40 SEC.	3 WORDS
Sound_bite_start_9__Cupcake_Wars 0.387348101 Bakers　　　Reposteras... 0　　　　　　　　　　　　　　1.292 Time (s)	0.38 SEC.	1 WORD
Sound_bite_start_10__Kitchen_Nightmares 0.277677262 None　　　Ninguna, Gordon. 0　　　　　　　　　　　　　　1.083 Time (s)	2.27 SEC.	1 WORD
Sound_bite_start_11__Teen_Mom_clip_1 0.891883354 I mean　　　No quiero que Simon... 0　　　　　　　　　　　　　　1.814 Time (s)	0.89 SEC.	2 WORDS

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_start_12__Teen_Mom_clip_1 0.756865671 Proposed Pero nada... 0 Time (s) 2.102	0.75 SEC.	1 WORD

ENDING SOUND BITES		

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_end_5__Cupcake_Wars 0.849634715 (...) todos los tiempos all time. 0 Time (s) 1.553	0.70 SEC	2 WORDS
Sound_bite_end_6___Kitchen_Nightmares 1.48269099 (...) e invitéis a gente a cenar. locals to dinner. 0 Time (s) 2.441	0.96 SEC.	3 WORDS
Sound_Bite_end_7__Kitchen_Nightmares 2.68190903 (...) están hundidos en mierda hasta las cejas. eyeballs in shitter 0 Time (s) 3.721	1.04 SEC.	3 WORDS

Sound_bite_end_8__Teen_Mom_clip_1 2.01779882 (...) voy a diseñar un anillo con mi hija — daughter 0 — Time (s) — 2.624	0.61 SEC.	1 WORD
Sound_bite_end_9__Teen_Mom_clip_1 1.43183334 Ese anillo era un despilfarro. — any sense. 0 — Time (s) — 2.18	0.75 SEC.	2 WORDS

AVERAGE STARTING SOUND BITES IN REALITY SHOWS

	0.93 SEC.	1.6 WORDS

AVERAGE ENDING SOUND BITES IN REALITY SHOWS

	0.81 SEC.	2.2 WORDS

C) INTERVIEWS PROGRAMMES

STARTING SOUND BITES

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_Bite_start_13__Euronews_Iron_Maiden 0.86009564 Nobody gets to Nadie escucha las canciones... 0 Time (s) 2.076	0.86 SEC.	3 WORDS
Sound_bite_start_15__Euronews_What's_it_like_for_EU_migrants_in_Britain_amid_UKIP's_success 0.488203787 I don't know Yo no se mucho acerca de... 0 Time (s) 2.102	0.48 SEC.	3 WORDS
Sound_Bite_start_16__Euronews_Is_it_time_to_scrap_the_CAP 0.672367311 But this is Bueno, entiendo que su pregunta... 0 Time (s) 2.076	0.67 SEC.	3 WORDS

ENDING SOUND BITES		
AUDIO WAVE	**TIME FRAME**	**NUMBER OF WORDS**
Sound_bite_end_10__Euronews_Iron_Maiden 0.933197616 (...) que las acabamos we have finished it. 0 Time (s) 2.467	1.53 SEC.	4 WORDS

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_end_11__Euronews_What's_it_like_for_EU_migrants_in_Britain_amid_UKIP's_success 1.43539847 (...) que han venido aquí a trabajar. people that come here they come here to work. 0 — Time (s) — 3.434	2 SEC.	9 WORDS
Sound_bite_end_12__Euronews_What's_it_like_for_EU_migrants_in_Britain_amid_UKIP's_success 2.84811137 (...) venir y firmar y eso está mal.　　sign on. It's wrong. 0 — Time (s) — 4.871	2.03 SEC.	4 WORDS
Sound_bite_end_13__Euronews_Is_it_time_to_scrap_the_CAP 5.29666734 (...) ganaría o perdería influencia si sale de la UE. that the UK could have either in the EU or out 0 — Time (s) — 7.196	1.9 SEC	10 WORDS
AVERAGE STARTING SOUND BITES IN INTERVIEWS PROGRAMMES		
	0.67 SEC.	3 WORDS
AVERAGE ENDING SOUND BITES IN INTERVIEWS PROGRAMMES		
	1.86 SEC.	6.7 WORDS
D) WEB VIDEOS		
STARTING SOUND BITES		
AUDIO WAVE	**TIME FRAME**	**NUMBER OF WORDS**

AUDIO WAVE	TIME FRAME	NUMBER OF WORDS
Sound_bite_start_17__TN_Philosophy 0.824261074 At Columbia Threadneedle En Columbia Threadneedle Investments... 0 2.337 Time (s)	0.82 SEC.	3 WORDS
Sound_bite_18__TN_Philosophy 0.683230346 Today when we look at Cuando analizamos los distintos mercados... 0 2.52 Time (s)	0.68 SEC.	5 WORDS
Sound_bite_19__TN_Philosophy 1.08433381 The greatest benefit La mayor ventaja de nuestro perfil mundial.... 0 3.486 Time (s)	1.08 SEC	3 WORDS
Sound_Bite_start_20__TN_Philosophy 0.974665021 It's important to Es importante seleccionar a los ganadores... 0 3.199 Time (s)	0.97 SEC.	3 WORDS
ENDING SOUND BITES		
AUDIO WAVE	**TIME FRAME**	**NUMBER OF WORDS**
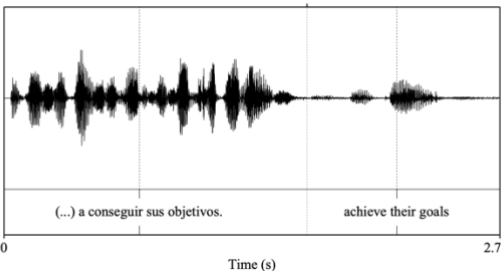
Sound_bite_end_14__TN_Philosophy 1.66729156 (...) a conseguir sus objetivos. achieve their goals 0 Time (s) 2.729	1.06 SEC.	3 WORDS
AVERAGE STARTING SOUND BITES IN WEB VIDEOS		
	0.88 SEC.	3.5 WORDS
AVERAGE ENDING SOUND BITES IN WEB VIDEOS		
	1.06 SEC	3 WORDS
AVERAGE GENERAL STARTING SOUND BITE		
TOTAL	**0.77 SEC**	**2.8 WORDS**
AVERAGE GENERAL ENDING SOUND BITE		
TOTAL	**1.12 SEC**	**3.5 WORDS**

Table 35 Duration and wordcount of the sound bites in our corpus

With the information provided in the table above, the time frame and number of words specified by certain scholars, which can be seen in table 5 in section 1.3.3, and also above, do not work 100%. From the sound bites we have selected 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.

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, in all the videos analysed in our corpus (with the exception of one web video), there are times in which there are no sound bites because the target voice starts and finishes at the same time as the original. This coincides with Baños´ (2019: 275) analysis of sound bites in Spanish reality TV programmes, for whom 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 and documentaries as regards voice-over isochrony and literal synchrony (2018: 280). In our case, after analysing both documentaries and reality TV programmes, together with news interviews and web videos, we find that there is no difference concerning how isochrony and literal synchrony, as shown in the next section, is applied. Thus, several situations can be found in all voice-over sub genres: that both voices start and finish at the same time; that there is a sound bite at the beginning but not at the end because both versions end together; and, 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 there are almost no sound bites for the Web Video category. This is due to the fact that one of the videos (*TN Europe Jan 17*) has been edited in the target version and the 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 post production 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 pacey and there is not enough time frame to fit the target text without sounding rush, so the sound bite is not used. Other times, the target script might be too long to leave some time for sound bites that they are removed.

5.2.6.2 **Literal synchrony**

Literal synchrony consists, as seen in section 1.3.3, in 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. Since the 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, which may not even occur in situations when synchronicity is hard to keep due to timing constraints.

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

	Literal Synchrony (cases of no sync)	
	The last days of Anne Boleyn	18
	Jamie's Comfort Food	9
DOCUMENTARIES	*Tasteology: Chill*	1
	Tasteology: Experience	0
	Total Documentaries	**28**
Interviews Videos	*Euronews: What's it like for EU migrants in Britain amid UKIP's success?*	3
	Euronews: Is it time to scrap the CAP?	6

	Euronews: Iron Maiden interview	1
	Total Interviews Videos	**10**
REALITY TV	*Kitchen Nightmares*	1
	Cupcake Wars	0
	Teen Mom 1	3
	Teen Mom 2	3
	Total Reality TV	**7**
WEB VIDEOS	*TN Philosophy*	6
	TN Europe	0
	Total Web videos	**6**
	TOTAL	**51**

Table 36 List of cases of no literal-sync

If we now consider how these cases of no literal-sync relate to the transfer of semantic meaning and of prosodic illocutionary force, we can see, as illustrated in the pie chart below, that more than half of them (53%) show full rendition of the original message, while 10% show no rendition of it at all. The remaining 37% renders the original message partially, with 31% transferring the basic semantic meaning and 6% the prosodic-illocutionary force.
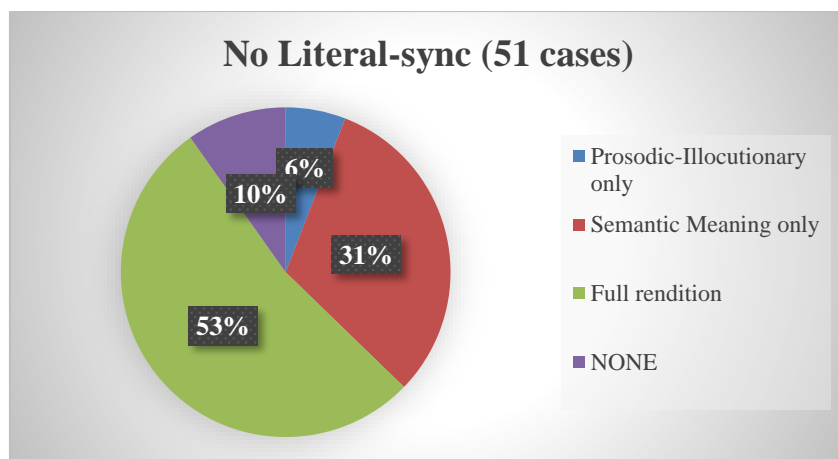
Figure 40 Rendition of No Literal-sync

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. Therefore, the presence of sound bites and the literal synchronisation of the target text with the original bites are more a question 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.

Some examples of no literal-sync are shown below.

L.S.Ex 1[88] *The last days of Anne Boleyn*				
In	Out	Character	Source Text	Target Text
00:13:47	00:13:49	Philippa Gregory	The belief was, in the medieval world,	En el mundo medieval, cuando una mujer daba a luz
00:23:10	00:23:16	Hilary Mantel	By 1536, the conflict is ready to explode.	Dos años después, el conflicto estaba a punto de estallar.
00:42:55	00:43:00	David Starkey	The final driver of everything under Henry, is Henry.	Durante su reinado, quien tomaba las decisiones en última instancia era Enrique.

[88] LS. Ex stands for Literal Synchrony Example

L.S.Ex 2 *Euronews: Is it time to scrap the CAP?*				
In	Out	Character	Source Text	Target Text
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?
05:12	05:15	Chris Burns, male	<u>So, how much are you</u> 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?

L.S.Ex 3 *Teen Mom* (Clip 2)				
In	Out	Character	Source Text	Target Text
00:03	00:04	Female	<u>For how long is Kate</u> going? What's going on? How long is she..	¿Cómo va? ¿Cuánto tiempo estará?
00:12	00:14	Female	<u>Yeah. And she went</u> 'cause she wanted to go.	Y… ¿ha ido por voluntad propia?
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.

L.S.Ex 4 *Euronews: Iron Maiden Interview*				
In	Out	Character	Source Text	Target Text

In	Out	Character	Source Text	Target Text
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. He's banned. 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,

			L.S.Ex 5 *TN Philosophy*	
In	Out	Character	Source Text	Target Text
00:00:53		Jim Cielinski	Today 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.
00:02:50		William Davies	But as 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.

5.2.6.3 **Content synchrony**

Content synchrony refers to the semantic correspondence or synchronicity of both texts (source and target) so they share the same content. This technical aspect is relevant for translation if we consider that Spanish is usually around 30% longer in word count than English, which means that there is an additional number of words to be fit in the same time frame as the original text. Besides, translators may use different procedures to render the source text and, thus, alter the target text, as can happen with expansions and

additions. Word order can also influence content synchrony. Given the fact that Spanish has a richer word order freedom compared to English, some relevant words may not match their position within the video, which can affect kinetic sync (as will be seen in the next section).

In our corpus we have identified 32 cases of no content-sync, which are shown in the table below.

	Content Synchrony (cases of no sync)	
DOCUMENTARIES	*The last days of Anne Boleyn*	15
	Jamie's Comfort Food	5
	Tasteology: Chill	1
	Tasteology: Experience	0
	Total Documentaries	**21**
Interviews Videos	*Euronews: What's it like for EU migrants in Britain amid UKIP's success?*	2
	Euronews: Is it time to scrap the CAP?	2
	Euronews: Iron Maiden interview	0
	Total Interviews Videos	**4**
REALITY TV	*Kitchen Nightmares*	1
	Cupcake Wars	0
	Teen Mom 1	1
	Teen Mom 2	3
	Total Reality TV	**5**

WEB VIDEOS	*TN Philosophy*	0
	TN Europe	2
	Total Web videos	**2**
TOTAL		**32**

Table 37 Cases of no content-sync

In terms of how these cases of no content-sync relate to the rendition of the original semantic meaning and the prosodic-illocutionary force, we observe, as shown in the pie chart below, that only 31% show full rendition of the source text, while 53% show partial rendition, and 16% no rendition at all.
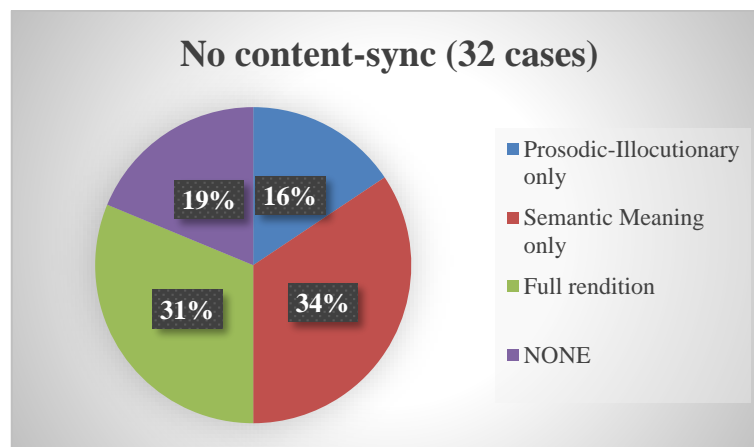
Figure 41 Rendition of no content-sync examples.

Full rendition examples are mainly cases in which the target text uses synonyms or near synonyms and reformulations to express the very same idea or meaning of the source text, as well as an equivalent illocutionary force. The examples below illustrate this situation.

Content-Sync Ex 1 *The last days of Anne Boleyn*				
In	Out	Character	Source Text	Target Text

00:08:33	00:08:40	David Starkey	No woman had ever made that step from royal mistress to the throne,	Ninguna amante del rey había llegado a sentarse en el trono,

Content-Sync Ex 2 *Jamie's Comfort Food*						
ST in	ST out	Character	Source Text	Target Text	TT in	TT out
04:25	04:26	Christian	Thanks for coming back.	Y cambiaste de opinión.	05:47	05:48
04:27	04:28	Jamie Oliver	I've come back.	Eso es.	05:48	05:49
09:34	09:35	Jamie Oliver	Have a look at that.	Vamos.	23:58	23:58

There are several examples showing partial rendition (semantic meaning only), in which the basic meaning is the same but, due to some change in the specific content, the prosodic-illocutionary force differs. Here are some examples.

Content-Sync Ex 3 *Kitchen Nightmares*						
ST in	ST out	Character	Source text	Target text	TT in	TT out
00:16:53	00:16:54	Male	I'm a pastry man.	A mí me gusta la empanada. Soy hombre de empanada	00:16:36	00:16:38

In this fragment, taken from the reality show *Kitchen Nightmares*, the translator expanded the target text, probably to reinforce the idea that the man speaking on camera prefers the pastry snack to the other one made of scallop. The target text here shows reaccenting of "empanada" (pastry) in a context in which the original speaker did not reaccent or emphasise that word, in fact, the original text is rather flat.

Content-Sync Ex 4 *The last days of Anne Boleyn*				
In	Out	Character	Source Text	Target Text
00:11:50	00:11:55	Hilary Mantel	So why? Am I still not on the right side of God?	Y se pregunta por qué no está a su lado,
00:11:55	00:11:57	Hilary Mantel	What does God want of me now?	qué será lo que quiere de él.

This example shows how the speaker is quoting somebody's words – i.e. demonstrating - in the source text, while the target text offers a reformulation of the original using modulation and reported speech. Thus, the general meaning is the same, but the illocutionary force of the original is not equalled, since describing or narrating is not the same as demonstrating.

We also have instances of partial rendition (prosodic-illocutionary only) in which the illocutionary force is the same but the semantic meaning of the text has been altered:

Content-Sync Ex 5 *The last days of Anne Boleyn*				
In	Out	Character	Source Text	Target Text
00:11:00	00:11:04	David Starkey	He's had two children, both of them daughters.	Ha tenido dos hijas. Dos mujeres.

In this example the target text has practically the same structure as the original one. In the video we can see how the speaker uses his fingers twice to refer to number 2 when saying "children" and "daughters". This is kept in the target text so the intention of highlighting number 2 is also present in it. However, the words chosen to translate "children" (hijas) and "daughters" (mujeres) are, in our view, confusing, because it might give the impression that king Henry V had two daughters and two wives. If

instead of "mujeres", which can denote "wife" apart from a female human being, the translator had used "niñas" or "chicas" (girls), it would probably have been less confusing.

Content-Sync Ex 6 *The last days of Anne Boleyn*				
In	Out	Character	Source Text	Target Text
00:16:26	00:16:31	David Starkey	And Henry, like a pendulum, swings from one to the other.	Enrique como un péndulo va de un extremo a otro.

This example shows David Starkey, again, talking about king Henry V. The structure of the target text is almost identical to the source one, as is the prosodic-illocutionary force, implying that Henry acts as a pendulum. This is reinforced by the hand gesture, which swings like that instrument. However, the words chosen in the Spanish text for "from one to the other" (de un extremo a otro) do not really render the meaning of the original, which is that Henry alternates between Anne Boleyn and Jane Seymour.

There are 6 cases of no content-sync which show no rendition of the original text at all. Three of these show extracts for which the target text has not been voiced. The other three show a completely different version of the source text in which both the original semantic meaning and the illocutionary force have been altered. Two of these cases of no-sync have already been analysed previously in this dissertation (D.Ex.10 and Q. Ex12) so we will only analyse the other instance of no content-sync.

Content-Sync Ex 7 *Teen Mom* (clip 1)						
ST in	ST out	Character	Source Text	Target Text	TT in	TT out
01:23	01:24	Simon	We were talking to…	Nunca hablé con ella.	02:45	02:46

This example of no rendition shows a completely different version in the target text. Simon, the speaker, is using the "we" to refer to Farrah (his fiancée) and himself

about something to do with the engagement ring. However, the target Spanish text shows this sentence, which has more to do with what was said before in the show that with what Simon is saying now. We can see that when Simon says this line, he is interrupted by Farrah, who swears and there is a beep which cuts the phrase. Probably, due to this interruption the translator decided to link this line with Simon's previous line, although it has nothing to do with it.

5.2.6.4 **Kinetic synchrony**

Kinetic synchrony implies that body language has to be in sync with what is conveyed orally. When we speak, we tend to gesticulate, and these gestures correlate with what we say. In audiovisual translation the gestures of the original speaker on screen have to match the translation delivered to the audience by the voice actor. If these do not match in time, the harmony between oral and body language will be broken and the sense of truthfulness may disappear from the target text.

In our corpus we have found 12 instances in which kinetic synchrony is lost. Most of the samples belong to the documentary *The last days of Anne Boleyn*, with 9, while there is one voice-over category in which we have not found any example: reality TV. The full list of examples and shows is provided in the table below.

No Kinetic Synchrony		
	Anne Boleyn	8
	Jamie's	0
DOCUMENTARIES	*Taestology chill*	0
	Taestology exp	1
	Total documentaries	**9**
Interviews Videos	*What's it like for EU migrants in Britain amid UKIP's success?*	0
	Euronews Time to Scrap the Cap	1

	Iron Maiden Euronews	0
	Total interview videos	**1**
REALITY TV	*Kitchen Nightmares*	0
	Cupcake Wars	0
	Teen Mom 1	0
	Teen Mom 2	0
	Total reality TV	**0**
WEB VIDEOS	*TN Philosophy*	0
	TN Europe	2
	Total web videos	**2**
	TOTAL NUMBER OF EXAMPLES	**12**

Table 38 List of examples of no kinetic sync

The lack of kinetic synchrony can imply some distress to the audience since words and gestures do not agree. Nonetheless, in our view, this lack of kinetic sync does not affect the translation of the source text offered to the audience in voice-over translation, either in terms of transfer of semantic meaning or of prosodic-illocutionary force. Thus, there are scenes showing kinetic synchrony which provide an unsatisfactory rendition of the basic meaning and illocutionary force of the original message, and scenes in which the opposite takes place, as can be seen in the pie chart below. From the 12 samples identified out of kinetic sync, 25% show full rendition of the original message, 8% no rendition at all, and 67% partial rendition (59% semantic meaning only and 8% prosodic-illocutionary only). This suggests that kinetic sync is, in voice-over translation, an aesthetic aspect which contributes to the illusion of faithfulness when it does occur. Needless to say, however, the more kinetic sync, the more realistic the translation will be.

No kinetic sync (12 cases)

- Prosodic-Illocutionary only
- Semantic Meaning only
- Full rendition
- NONE

8%
8%
25%
59%
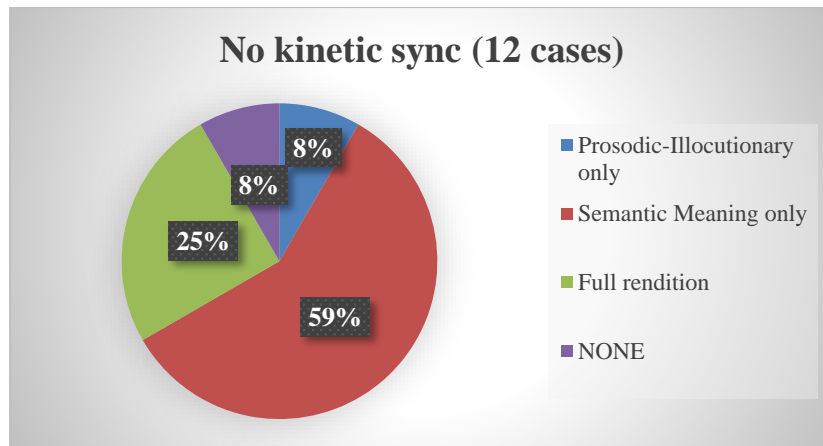
Image 92 Rendition of samples out of kinetic sync

Some of these examples are analysed in what follows, showing how they differ in their source and target versions.

K.S.Ex 1[89] *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:12:32	00:12:37	Hilary Mantel	When Henry talks about enchantments, charms, magic tricks,	Cuando Enrique habla de encantamientos, hechizos, conjuros, magia,

In this scene from the documentary *The last days of Anne Boleyn*, Hilary Mantel is commenting on some of the possible reasons Henry V had in order to annul and justify his marriage to Anne Boleyn. When she says this passage, she waves her hand

---

[89] K.S. Ex stands for Kinetic Synchrony Example

three times, each movement indicating one of the reasons ("enchantments, charms, magic tricks"). However, the Spanish version included one extra reason and, so the three original hand gestures no longer match the number of reasons we get in the target version ("encantamientos, hechizos, conjuros, magia").

K.S.Ex 2 *The last days of Anne Boleyn*				
In	Out	Character	Source text	Target text
00:22:41	00:22:43	David Starkey	absolutely full-on, square,	
00:22:43	00:22:45	David Starkey	at Cromwell himself.	ponía el ojo en el propio Cromwell.

This fragment has been reduced so much that it produces a situation in which the target audience of the documentary actually hears the original text for quite a long time. As can be seen in the picture above, David Starkey gesticulates with his hands four times, matching the words he is uttering in this passage ("absolutely, full-on, square, at Cromwell himself"). But the target audience does not get the corresponding translation of this passage at this time but some time earlier because of the reduction applied to the source text. This has led to a scene rather out of sync in which the audience may feel they are not being given the full translation of the original.

K.S.Ex 3 *The last days of Anne Boleyn*				

In	Out	Character	Source text	Target text
00:30:59	00:31:02	Hilary Mantel	She says, "Huh! Don't blame me."	Y ella le contestó: "Bah,
00:31:02	00:31:06	Hilary Mantel	"It's nothing to what the Queen gets up to,"	eso no es nada comparado con lo que hace la reina."

This passage shows Hilary Mantel, again, now commenting on Anne Boleyn's misfortune. This time she is quoting one of Anne's ladies, Lady Worcester, who was being told off by her brother for her supposed loose living. In her speech, Mrs Mantel says "don't blame me", and brings her hands towards her chest at the moment of saying "me" to emphasise and highlight this word, which, in addition, is the nucleus of the sentence (see Q.Ex 10 in section 5.2.2.2). In the Spanish version, however, the word corresponding to "me" disappears so there is no reference to the first person. The translator has decided to paraphrase and modulate the quoteusing an idiomatic expression in Spanish ("eso no es nada comparado con"), so the visual reinforcing with the hand gesture, and its meaning, are lost. This could have been avoided with a translation like: "Bah, lo que hago **yo** no es nada comparado con lo que hace la reina."

<table>
<tr><td colspan="5" align="center">K.S.Ex 4 Tasteology: Experience</td></tr>
<tr><td colspan="5" align="center"></td></tr>
<tr>
<td align="center">In</td>
<td align="center">Out</td>
<td align="center">Character</td>
<td align="center">Source text</td>
<td align="center">Target text</td>
</tr>
<tr>
<td align="center">00:05:52</td>
<td align="center">00:06:08</td>
<td align="center">Charles Spence</td>
<td align="center">People prefer an oblique line. And in particular they prefer a line that kind of sends to the right.</td>
<td align="center">La gente prefiere una línea oblicua. En concreto, prefieren una línea que tienda hacia la derecha.</td>
</tr>
</table>

This fragment from the documentary *Tasteology: Experience* shows gastrophysicist Charles Spence talking about people's food preferences. In the source text we can see and hear how his last words ("sends to the right") are accompanied by a hand gesture pointing to the right (the left of the audience). The target text has a suitable translation for this passage but we can hear that the words "hacia la derecha" are out of frame, so the sync between the arm movement and the voice delivering the translation is lost.

<table>
<tr><td colspan="7" align="center">K.S.Ex 5 TN Europe video Jan 17</td></tr>
<tr><td colspan="7" align="center"></td></tr>
<tr>
<td align="center">ST in</td>
<td align="center">ST out</td>
<td align="center">Character</td>
<td align="center">Source text</td>
<td align="center">Target text</td>
<td align="center">TT in</td>
<td align="center">TT out</td>
</tr>
</table>

00:24:00	00:26:00	Francis Ellison	Orderly? Disorderly?	¿Se llevará a cabo de forma ordenada o desordenada?	00:31:00	00:34:00
00:27:00	00:29:00	Francis Ellison	Will it hurt us? or Europe? Or Both?	¿Perjudicará al Reino Unido o a Europa, o a ambos?	00:35:00	00:38:00

This last sample of no kinetic synchrony belongs to the web video *TN Europe video January 17* and shows Francis Ellison, Client Portfolio Manager at Columbia Threadneedle Investments, commenting on the possible outcome of Brexit. During this scene he moves his hands to accompany the IPs he utters to describe the future[90] Brexit situation: "orderly", "disorderly", "will it hurt us", "or Europe", "or both". The Spanish target video, however, shows a mismatch between body language and the delivery of the target text, something which could have been avoided since the target video clip has been edited and elongated in certain parts, as the time frame differences between source and target video show: as we can see, the original source text for this scene takes place from second 24 to 29, while the Spanish target version goes from second 31 to 38.

5.2.6.5 **Character synchrony and voice quality**

The last synchronic features that will be dealt with in this dissertation are character synchrony, which, as seen in section 1.3.3 means that both the source and target voices have to agree, at least, in age and gender, and voice quality, which includes pitch and timbre or colour of the voices. Their aim is to achieve a translated product as faithful as possible to the original.

In our corpus, we have identified 6 instances in which characters are out of sync considering age and gender. These are probably the easiest characteristics to identify in a voice. However, there are others that are hidden, such as timbre, for which we need to resort to specific software. In addition, given the fact that we cannot see how the

[90] This web video was created in January 2017, two years before a final decision on Brexit was made by the British government and the EU.

voice actors that deliver the translation are in terms of their physical appearance and state of health, the use of dedicated software is paramount. As an example, we can think of the world-famous opera singing group *The Three Tenors*91: Pavarotti, Domingo and Carreras. They have the same voice range or tessitura, since they are tenors, but they sound rather differently. This is due to the timbre or colour of their voices.

This hidden characteristic of voices called timbre can be analysed by means of a harmonic spectrum, as stated in section 4.4.2. In this section we will compare similar samples of the two voices (source and target). The traditional methodology of voice identification in forensic phonetics states that these examining samples have to be the same (Rose 2002: 111) as well as short in length, such as the same word or vowel, so that we can focus on personal voice quality (Nolan 1991: 484). In our case, since the two texts do not have the same words as they are in different languages, we will use similar vowel sounds to check the harmonic properties of the two voices. There is one case, however, in which the two voices we compare do utter the same word; so we will use that word to make the comparison. In addition, considering that English and Spanish have different realisations of vowel sounds, we will try to isolate those phonemes that sound similar in the two languages.

Once we have analysed the examples of no character sync, we will briefly comment on the strategies used to render some of the most identifiable English accents found in the source texts within our corpus in the Spanish version of the different programmes.

At this point it is relevant to bear in mind that the voices from the two texts (source and target) will never be "the same". Every voice is unique and it is obviously impossible to have an identical voice delivering the target text of a revoiced product (dubbed or voiced-over). There are multiple factors that affect voice quality, such as, the physical condition of the speakers, the size of their vocal cords, their state of health, their psychological state, their breathing rhythm, etc. In addition, we should also include the recording environment of the voices (e.g. if the voices were recorded in a studio or outdoors), the quality of the microphones, if the microphone is too close to

---

91 Information on The Three Tenors: https://en.wikipedia.org/wiki/The_Three_Tenors

the speaker's mouth, etc. Therefore, the examples we are about to describe can only be, in our view, the most obvious ones in terms of voice differences.

5.2.6.5.1 Character synchrony

The first sample of no character sync has been found in the Euronews programme *Le Mag*, in which Bruce Dickinson, Iron Maiden's lead singer and frontman, is interviewed to promote their latest studio album. The original English TV show has two main voices: a narrator, who makes the introductory comment, gives some factual information in the middle and then gives a final comment before the end of the show; and Bruce Dickinson's voice, who is the man shown on camera during the show. However, the Spanish version has opted for a female voice to deliver Bruce Dickinson's part, while the narration has been done by a male voice. The fact that a male on camera is being voiced-over by a female voice can be, from a synchronic perspective, inappropriate and ill-advised, going to the detriment of the faithful and authentic translation usually expected in voice-over. Of course, the same would happen if the voices were switched the other way around. It has not been necessary to resort to the harmonic spectrum, since the difference between the male and female voice is evident.

Our next example of lack of character synch has been retrieved from the Euronews show *Reporter: What's it like for EU migrants in Britain amid UKIP's success?*. In this show, an English man, who appears to be in his seventies, is interviewed in minute 03:24. His voice is that of an elderly man, nasalised, and relatively high in pitch. On the other hand, the voice delivering the translation lacks the nasal component, sounds younger and a bit lower in pitch than the original. In order to compare the voice quality of these two speakers, we have isolated similar phonemes they uttered in their corresponding scenes. The phoneme in question is /i/ taken from "see" /si:/ in the English text and "aquí" /i/ in the Spanish one. It would have been ideal if both speakers had uttered the same phoneme at the same pitch in the video, but that is impossible. The table below shows the harmonic spectrum for the two phonemes uttered by the two speakers in different contexts.

Character sync Ex2
Original English Voice

Target Spanish Voice
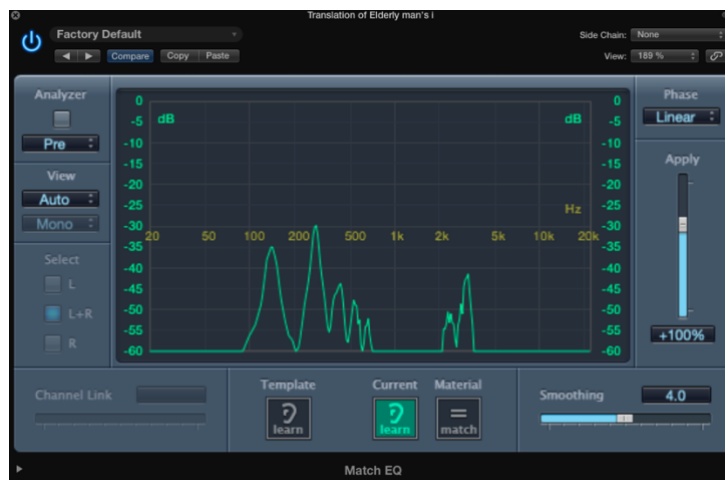
Image 93 Comparison of fundamental frequency and harmonics of two voices in *Reporter*

These two images represent the sinusoidal components of the harmonic spectrum for the two voices (English and Spanish) analysed. These sinusoidal components are called harmonics (Rose 2002: 205), which can be defined as the

overtones₉₂ of a fundamental note or frequency. These, in turn, mark the tone or pitch and the timbre of voices and musical instruments.

In the table above we can see that the two voice samples share the same fundamental frequency value (more or less), at around 150 Hz, and the subsequent harmonics have similar frequencies but different amplitudes (see the vertical axis in green). These acoustic differences clearly show, in addition to those related to the physical condition of the speakers, the status of their phonatory system and their state of health (that nasal nuance in the original English voice may signal the speaker has a cold) that the two voices belong to people of different ages; as a consequence, character synchrony does not match.

Our next example has been found in the reality show *Teen Mom OG (clip 2)*, in which we can see and hear a male teenager speaking to a lady about his couple issues. His voice sounds as that of the typical male teenager or a man in his early twenties. If we then listen to the voice used in the translated version, we can notice it is way deeper, richer in resonance (probably due to the fact that it was recorded in a studio close to a microphone), husky, and it also sounds older, as if coming from a male in his 30s or even 40s. Here is how the spectrogram depicts these two voices for phonemes / ɒ / extracted from "what", and / ʌ / from "quería". In General American English, the phoneme / ɒ / is pronounced with unrounded lips, unlike on General British English (Setter 2019: 91), which makes it sound more like an / ʌ / in Spanish, despite the former is back and the latter is central.

Character sync Ex3
Original English voice

₉₂ Overtones are resonant frequencies above the fundamental frequency (F0). More information on overtones can be found on the Encyclopedia Brittanical on-line: https://www.britannica.com/science/sound-physics/Overtones

| Target Spanish voice |

Image 94 Harmonics in *Teen Mom OG* (clip 2)

As we can see, the original English voice shows a first peak at around 110 Hz, being this the fundamental frequency. If we check the Spanish voice graph, the difference is evident; the amplitude of the first peak is way higher, and there seems to be a peak at around 70 Hz which then rises up to around 110 Hz. The area below 100 Hz is the place where the low frequencies lie, and it is evident that the target text voice is richer in low frequencies than the original. If we now check the harmonics, the differences are also evident. All these make the two voices sound different and, in this case, we consider that the voice used in the translation of the show is out of sync in

terms of character. It sounds deeper and older, not the type of voice one could expect in a show called *Teem Mom* (*dad* in this case).

The next example identified in which there is no character sync is in the web video *TN Philosophy*, showing different male characters voiced by the same voice talent in the target version. This is usually done to make the production of the target video cheaper, since hiring different voices raises the price significantly. The voice used in the translation video belongs to someone in his late 20s or early 30s, and it is neither a deep resonant voice nor one too high in pitch. In our view, this voice clashes, in terms of character sync, with the first man speaking in the video, Colin Moore, who seems to be in his 60s, and whose voice is deeper than the one used in the translation. In any case, the age gap is significant enough to signal no sync in terms of character. In order to analyse the acoustic properties of the two voices, we have isolated phoneme /e/ from the word "Threadneedle" which is uttered both in ther source and target texts.

Character sync Ex4
Original English voice
E Sound in TN Philosophy Original Male — Alternative 01 — Side Chain: None — View: 200% — Analyzer — Pre — View — Auto — Mono — Select — L — L+R — R — Phase — Linear — Apply — +100% — Channel Link — Template learn — Current learn — Material match — Smoothing 4.0 — Match EQ
Target Spanish voice

382
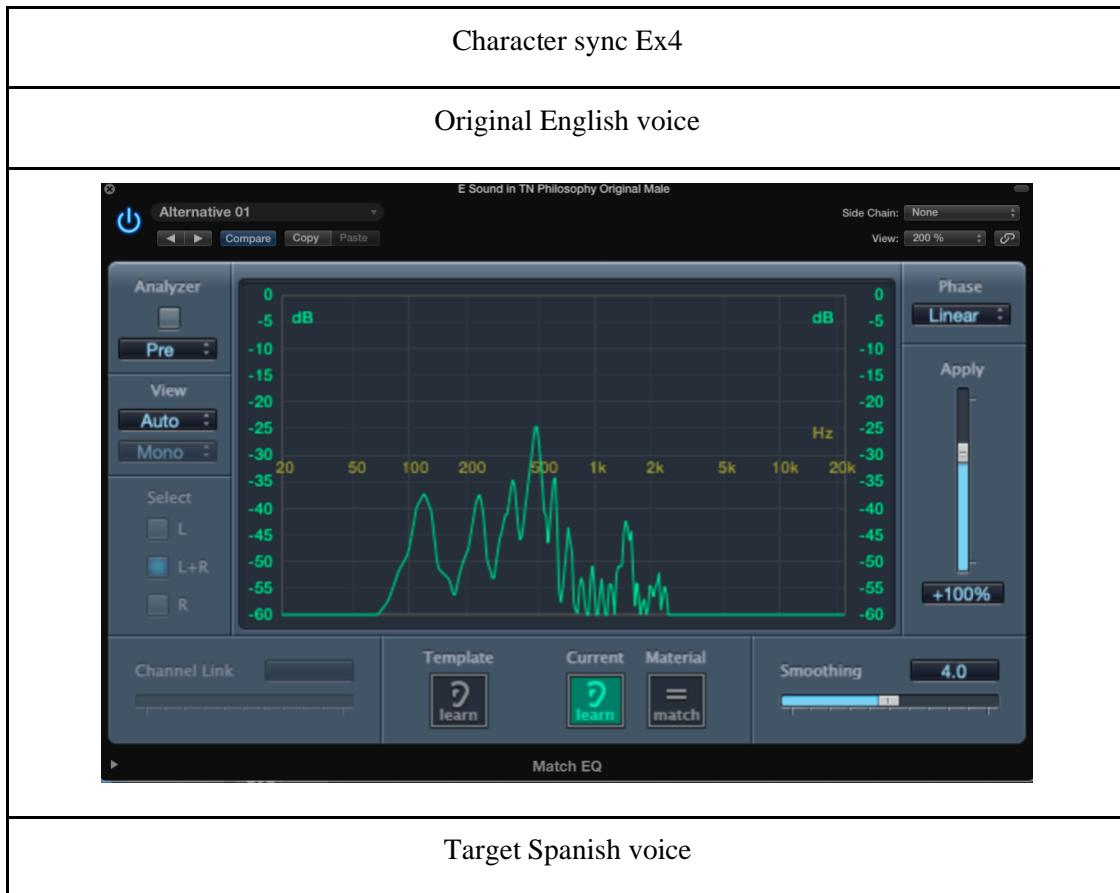
Image 95 Harmonic spectrum in character sync example 4

The graphs above depicting the two voices show several differences. First, the fundamental frequencies are different: the original English voice has a lower fundamental frequency than the Spanish voice and the area below 100 Hz is bigger, which implies that this voice sounds deeper (lower in pitch) than the one used in the Spanish translation. In addition to this, we can see that the fourth and fifth harmonics differ in terms of both amplitude (higher in the original English voice) and frequency (lower in the original English voice). All these nuances make the two voices rather different, so they may be considered out of sync in terms of character representation.

The fifth and last example of no character sync found in our corpus belongs to the web video *TN Europe video Jan 17*, in which we can see and hear Francis Ellison, Client Portfolio Manager of the firm, who seems to be in his 50s, giving his economic predictions for that year. His voice is fairly high pitched compared to the one in the translation, although the Spanish voice talent does not sound particularly deep or resonant. This time, we have isolated the word "core", which is pronounced in English in both texts (source and target) referring to "core plus" investment strategies, instead of just a single phoneme as in the previous examples. Moreover, we consider it might be interesting to use this possibility of studying one and the same word pronounced by different speakers for this purpose. The acoustic representation of this utterance is shown in the image below.

Character sync Ex5
Original English voice
Target Spanish voice

Image 96 Harmonic spectrum in character sync example 5

As can be seen from the two images shown above, the most evident difference is the fundamental frequency, which is set at almost 200 Hz for the English speaker (notice the second harmonic, almost halfway between 200 and 500 Hz - like the third harmonic in the Spanish voice) while it is between 100 and 150 Hz for the Spanish

speaker, with the second harmonic a bit over 200 Hz. This explains why the two voices sound quite different in terms of pitch and timbre, apart from the fact that they also differ in terms of age.

These examples show the most evident cases of lack of character sync in our corpus. We understand that it would be impossible to have an exactly similar voice to the original one, delivering the target text in the Spanish videos. No two voices sound the same, and by using speech analysis software we can always find differences at the acoustic level between phonemes or words pronounced by two speakers. Nonetheless, the examples highlighted in this section show clear divergences that might prove disturbing for the target audience and disrupt the expectation that voiced-over products should be reliable and authentic representations of the original.

5.2.6.5.2 Translation of accents

We have identified several accents in the source English texts of our corpus, such as David Starkey's and Francis Ellison's RP English (also called BBC English or General British English), Farrah Abraham's and Ashley Eckstein's General American English, or Gordon Ramsey's and Jamie Olvier's popular Cockney English (although some think they use Estuary English93). These oral features have been neutralised in the target texts, being reproduced as standard Castilian Spanish, as suggested by Perteghella (2002) and Tello (2011) - see section 2.3. This means that some of the social and cultural traits that are usually conveyed through accents in English (and in all languages) will, unfortunately, not reach the target audience.

5.2.7 Discourse markers

Discourse markers, sometimes called *fillers*, are words and expressions typical of spoken discourse, which, as seen in section 2.3, contribute to the coherence and

---

93 https://www.independent.co.uk/voices/columnists/susie-rushton-estuary-english-ndash-the-pride-of-kent-2293831.html

cohesion of texts (Matei 2010: 125) as well as to the pragmatic meaning of sentences uttered by speakers. In audiovisual translation, the rendition, or not, of these items has been a controversial issue of debate among scholars in recent decades. Some, like Chaume (2004) and Matamala (2008), consider that these fillers are superfluous and should be removed from the target text for the sake of brevity. Others, like Rica Peromingo (2014) and Muller (2005), understand that the use of discourse markers in film language is paramount in order to keep the interaction between the characters of a film. It is interesting, in our view, to see how such features as discourse markers are treated in our corpus.

We have found 182 examples of discourse markers in the source texts analysed, of which only 73 have been rendered, leaving 109 untranslated, as shown in the bar chart below. This means that 60%, as the pie chart on the right shows, of these oral traits which contribute to the coherence and cohesion of oral discourse, have been kept away from the target audience.
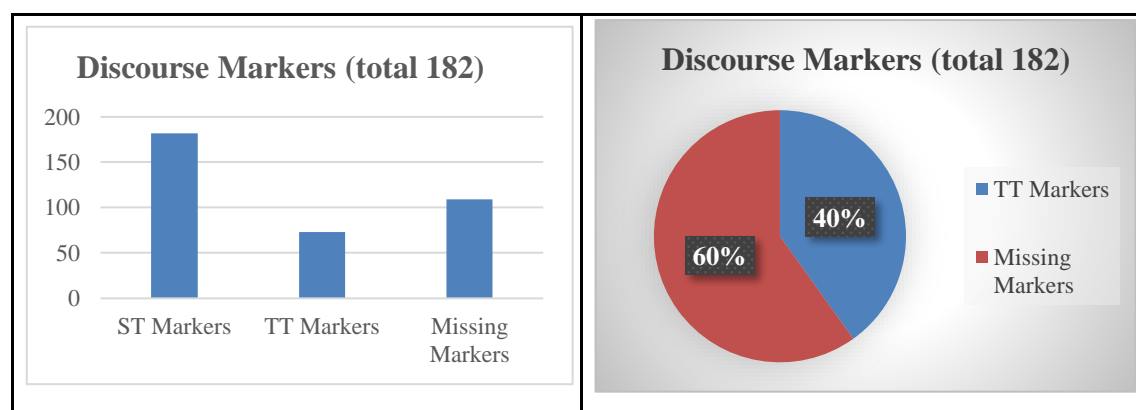
Figure 42 Number of discourse markers both in TT and ST and their percentages.

In terms of voice-over category, discourse markers are distributed according to the following table.

Category	Programme	Source Text	Kept in Target Text	Missing Markers
DOCUMENTARIES	*Anne Boleyn*	26	5	21
	Jamie's Comfort Food	18	9	9

Category	show	Source Text	Kept in Target Text	Missing Markers
	Tasteology: Chill	6	4	2
	Tasteology: Experience	14	6	8
	Total	**64**	**24**	**40**
Category	**show**	**Source Text**	**Kept in Target Text**	**Missing Markers**
Interviews Videos	*What's it like for EU migrants in Britain amid UKIP's success?*	4	0	4
	Euronews: Is it Time to Scrap the Cap?	6	6	0
	Iron Maiden Euronews	3	0	3
	Total	**13**	**6**	**7**
Category	**show**	**Source Text**	**Kept in Target Text**	**Missing Markers**
REALITY TV	*Kitchen Nightmares*	55	20	35
	Cupcake Wars	32	17	15
	Teen Mom 1	2	2	0
	Teen Mom 2	4	0	4
	Total	**93**	**39**	**54**
Category	**show**	**Source Text**	**Kept in Target Text**	**Missing Markers**
WEB VIDEOS	*TN Philosophy*	9	4	5
	TN Europe	3	0	3
	Total	**12**	**4**	**8**

		ST Markers	TT Markers	Missing Markers
TOTAL Markers		**182**	**73**	**109**

Table 39 List of discourse markers and voice-over categories both in ST and TT

The table above shows that the voice-over category with the most number of discourse markers is reality TV, followed by documentaries and interviews videos. This is understandable, considering that these three types of programme are unscripted and filmed on the go while the participants do whatever they have to in their shows, from cooking to answering questions. The category with the smallest number of markers is web videos, probably due to the fact that the reports shown on them have a written and planned script.

During the analysis of the source videos and while searching for markers, we have noticed that in the show *Kitchen Nightmares*, presenter Gordon Ramsey uses the tag or phrase "you know that" (both in the affirmative and interrogative form) as a marker of his idiolect. 19 instances of this tag have been identified in the source text. From these, 14 have been rendered in the target Spanish text, while 5 have not. The rendition of this tag in the translation has been done by means of different forms, as shown in the table below.

Source text	Target text	Number of repetitions in TT
You know that.	Ya lo sabes	3
You know that?	¿Lo sabes?	6
You know that?	¿Sabéis?	1
You know that?	¿Sabes?	2
You know that?	¿Vale?	1
You know that?	¿Lo sabéis?	1

Table 40 Gordon Ramsey's marker repetitions

The table shows that most of the characteristic markers of chef Gordon Ramsey have been kept in the target text. With the exception of one instance, translated as "¿vale?", all of the rendered tags use the same verb, "saber". This consistency in keeping the same verb for the rendition of the tag gives coherence to the target text and allows the target audience to perceive some of Ramsey's identity traits and idiosincrasy, thus getting a better picture of the character.

Another interesting part of colloquial and spontaneous language are swearwords. Like markers or fillers, these, contribute to the coordination between interlocutors and the interaction and structuring of verbal exchange (Dewaele 2004), apart from being significant markers of a speaker's personality. We have found two texts with swearwords in our corpus: the reality shows *Kitchen Nightmares* and *Teen Mom OG* (clip 1).

These words are treated differently in the source and target text. On the one hand, we have two source texts with swearwords: one coming from the USA and the other from the UK. In the USA-produced show, the swearwords are "hidden" by a beeping sound so that the audience cannot hear them. By contrast, these words are clearly heard in the UK-produced show, so that the audience can get the full picture. When we check the rendition of these words in the Spanish voice-over, we can observe that these words are kept only sometimes - most of the time not being translated. The chart below shows the swearwords found and how they have been rendered.

Kitchen Nightmares			
Source text		**Target text**	
Swearword	**Repetitions**	**Swearword**	**Repetitions**
Shit	12	Mierda	17
Bloody	1	Maldita	2
Fuck	23	Joder	1

Fucking	65	Jodido	2
Teen Mom OG (clip 1)			
Source text		**Target text**	
Swearword	**Repetitions**	**Swearword**	**Repetitions**
Unclear (beeped)	1	none	0
Unclear (beeped)	1	none	0

Table 41 List of swearwords and their translation in our corpus

This table shows that swearwords are a relevant characteristic of Gordon Ramsey's way of speaking. There are over 100 swearwords in *Kitchen Nightmares*, which, in our view, implies that this type of language is quite relevant for the full understanding of the source text since it is part of the character's personality. Gordon Ramsey is popular for being forward and for swearing very much in his shows. If we now check the target text, we find only 22 swearwords from the original 101, "mierda" ("shit") being the most frequently used with 17 repetitions. This swearword is the softest of them all, which is probably why it is used so many times in the Spanish version.

In translation of *Teen Mom OG* (clip 1), the beeped swearwords are not kept in the target text: the beep is there, but no swearword.

It is our understanding that these shows are originally produced for a certain audience and that when they are purchased by foreign broadcasting networks or TV channels, the target audience will be more or less similar. So, if swearwords are a relevant part of the original text and are not silenced or beeped in it, as is the case with *Kitchen Nightmares*, they should also play a role in the target text. And if the target culture does not accept swearwords in the same way as the source culture, "equivalent" markers should be used in order to portray these peculiar nuances, which are inherent to the source text characters' personality, since the main function of audiovisual translation is, according to Chaume (2004: 844) to produce a similar effect on the target audience as the source text produced on the source culture audience.

The data collected in this dissertation shows that there is no rule of thumb regarding the translation strategies for discourse markers. As we have seen, 59% of these are omitted, but 41% are kept. This clashes with some of the academic studies that defend that markers should be sacrificed, such as Rica Peromingo (2016: 144), Barzdevics (2012: 67), Matamala (2008: 118) and Chaume (2001: 81). Some markers, such as Gordon Ramsey's (*you know that*) in *Kitchen Nightmares*, if translated, can enhance adequately the target text, since they depict the speaker's personality. The same applies to swearwords, but that may depend on the broadcasting network's style guide. All in all, almost half of the markers in our corpus have been kept in the target text, providing it with the authenticity flavour that voice-over translation is expected to have.

6   Conclusions

The final chapter of this thesis is devoted to summarising the conclusions obtained from the results of the analysis section and the questionnaires answered by voice-over professionals (translators, actors and casting directors). As stated in chapter 4.4.2.3, three different questionnaires have been provided to voice-over translators, actors, and casting directors (one each) so as to get a deeper insight on how relevant prosody is in the process of professional voice-over translation. Thus, we have got 25 responses from casting directors, 14 from voice actors, and only 4 from translators. We would have preferred to have more answers from translators, but since answering surveys for academic purposes is a voluntary task, there was nothing we could do to get more answers after asking a large number of professionals. We may remind the reader that, as stated in section 4.4.2.3, the information provided by translators and voice actors is in Spanish, while the information given by casting directors and agents is in English, since most of them work for international localization companies outside Spain.

We will start now answering the objectives set in chapter 4; we will then check whether our initial hypothesis was fulfilled or not; and, to conclude, some final remarks on the future of voice-over translation will be presented.

6.1 Goal fulfilment

In its initial stages, this dissertation set five goals in order to analyse how English intonational and prosodic features are rendered in Spanish voice-over programmes. These were:

a) To identify the way prosodic features related to stress, pitch, together with Wells' (2006: 06) 3Ts (tonality, tonicity and tone) are translated into Spanish in the voice-over mode and whether the original prosodic-illocutionary force implicit in the source text has been rendered or lost in the final translation.

b) To identify the way discourse markers are considered in voice-over translation.

c) To check how synchronic features can affect or be affected as a result of the translation process.

d) To check how voice quality is treated in voice-over translation.

e) To check how voice-over professionals (translators, actors, adjusters, directors) conceive prosody and prosodic traits in the process of a voice-over production.

- *Answer to objective a)* [To identify the way prosodic features related to stress, pitch, together with Well's (2006: 06) 3Ts (tonality, tonicity and tone) are translated into Spanish and whether the original prosodic-illocutionary force and the communicative value implicit in the source text have been rendered or lost in the final translation.]

Chapter 5 offered the quantitative and qualitative analysis of how intonational features related to English stress, pitch, and the 3Ts (Wells 2006: 06) were treated in Spanish voice-over translation. We will now summarise the strategies and resources that translators have used to render these intonational features and their suitability for each case.

Stress: deaccenting

In terms of English stress, we analysed two features: deaccenting and reaccenting. For deaccenting, the rendition strategies found have been: anaphoric deixis, reaccenting and ellipsis. From these, the most successful ones in order to render both the semantic meaning and the prosodic-illocutionary force have been deixis and ellipsis, while reaccenting has proved to be the least successful one. This is due to the fact that, in English, the deaccenting of an element implies narrow focus elsewhere and the common reaccenting in Spanish of the deaccented English words, if placed in final or end position within the IP, can bring it into focus. However, if we omit the English deaccented word by means of ellipsis or we use a deictic pronoun in theme position to refer to it, the conveyance of the illocutionary force of the original message can easily be achieved.

Stress: reaccenting

English reaccenting has been rendered in Spanish voice-over by means of anaphoric deixis, cataphoric deixis, ellipsis and reaccenting (using the same word again). The four are valid to render the semantic meaning and prosodic-illocutionary force, since, as seen in section 5.2.1.2, the cases of partial or no rendition in the corpus were due to an inadequate delivery or a mistranslation of the original text in terms of lexical choice, as in examples R.Ex.5 and R.Ex.9. From these four strategies, the most commonly resorted to has been reaccenting, with 41 cases out of 63, doubling the second most frequently used strategy, ellipsis, with 19 cases. Considering that there are time constraints in voice-over translation, ellipsis or deixis (anaphoric and cataphoric) are, in our view, more useful in cases of reaccenting so that we can save some words and syllables if isochrony is tight.

Pitch: intonational subordination

Intonational subordination, as we have seen, signals portions with little or no information in the context of the conversation, and these intonation units are of a parenthetical or *throw-away* nature (Mott: 2011: 229), and "separate from what we are saying" (Wells 2006: 243). This implies that these IPs are not truly relevant and could be omitted in the target text if the translator is able to identify them. In our corpus, the two strategies used to render this English intonation feature have been: translation of the subordinated IP (8 cases), and deletion of IP (3 cases). Both strategies are suitable to render the semantic meaning of the original, but since these IPs are not substantial, in our view, the best option is to delete them for isochronic purposes. However, the difficulty seems to lie in identifying these IPs when there is pitch variation, since from the 11 cases of subordination, 8 have been included in the target text, which may imply their irrelevant value was probably not perceived by the translator.

Pitch: quotes

Moving now to quoting, we observe this is a very recurrent resource in oral spontaneous speech in English. The main rendition strategies found in our corpus are: keeping the quote in the target text, and reporting or paraphrasing it. There is one case in which the

quote was repeated (reaccented) in the ST, which was rendered by means of deixis in the TT, so this strategy falls on the reaccenting category rather than on quotations, although it has been quantified here. From the 69 instances of quotations in our corpus, 45 have been rendered keeping the quote and 23 paraphrased or reported. There seems to be no rule of thumb to understand when a quote should be kept and when it should not, or when translators opt for the quoting form in the voice-over too or decide to paraphrase.

From the questionnaire provided to voice-over translators, question 8 was addressed to the translation of quotations, as the following figure shows, providing the answers obtained too:
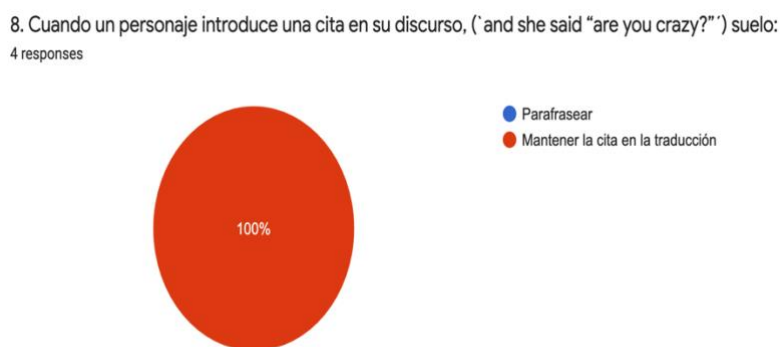
Figure 43. Answer to question 8 in the translators'' survey

As can be seen in the figure above, all four translators who answered the survey usually keep the quotation in the target text. This result explains why most of the quotes have been rendered as quotes instead of being rephrased. In addition, quotations work as demonstrators and mostly depict illocutionary acts (Clark and Gerrig 1990: 779). This means that if we want to keep the same illocutionary value in the target text, quotations are the best alternative. On the other hand, paraphrasing perfectly serves to render the semantic meaning of the original, but in terms of prosodic-illocutionary force the result is not the same as quotations. We have to bear in mind the visual part of the text when translating quotations, since most of the time the people on screen can be seen gesticulating or imitating the person being quoted. Moreover, if there are sound bites preceding or following the translation (a reporting or paraphrasing), the feel of

authenticity and realism can be lost since both texts (source and target) will not coincide.

Tonality: subject-predicate division

Oftentimes, in English, the subject of a clause is detached from the predicate with a single IP in order to spotlight it, or because the subject is long and it is preferable to have two breath groups rather than one. In our corpus we have found 6 instances of subject-predicate division, which have been rendered in Spanish voice-over in two ways: splitting subject and predicate as in English, or joining subject and predicate in one single IP. There seems to be no problem in terms of rendering the semantic meaning and the illocutionary force of the original with these two strategies. There is just one example in which the prosodic-illocutionary force has not been fully achieved since there was no break or insistence stress on the subject to highlight it as in the original.

We have noticed that most of the cases of subject-predicate division in English have been translated splitting the subject and predicate in the target text (4 out of 6), although this is not a common feature in Spanish. This could be the result of the interference of the source text during the delivery of the translation by the voice talent. Given the fact that voice actors wear headphones during the recording sessions so that they can listen to the original version and know when to start and finish their lines, this could lead them to make the same pauses as the English voices. As we know, there is no need for lip synchronisation in voice-over, but the target voice certainly needs to finish a bit before the original, or, at least, at the same time, never after, which may be the reason why we get these pauses in the Spanish version.

Tonality: Topicalisation (topics and cleft sentences)

Topicalisation is an excellent way of marking emphasis in English and one of the most resorted to strategies to do so are cleft sentences (*it-type* and *wh-type* or pseudo clefts). In our corpus, we have identified 11 instances of cleft sentences, and one of a topicalised pseudo cleft (see Top.Ex 1), for which there are two rendition strategies: using an equational structure following Braga & Rica Peromingo's (2015) model to translate English structural emphasis into Spanish (see page 121), and a literal

translation with no relative clause. There is one instance in our corpus in which the line was cut out (this is just one exceptional case in which the Spanish video has been edited and that scene was removed). From the two strategies mentioned above, the most successful one has been using equational structures, which can be formed as end-focus constructions, like English pseudo-clefts, ("What...is….”/Lo que…es..."), or placing the highlighted material in theme position, followed by the verb "ser" and the relative clause with the contrasting sequence, as in "la magnitud de los cargos es precisamente lo que los hace más convincentes".

The other strategy that has been used to render cleft-sentences in Spanish voice-over with a less successful result has been using a literal translation without a relative clause. The issue here is that without the relative clause that characterises English clefts and equational structures in Spanish, the contrastive element can be lost and, consequently, the illocutionary force cannot be fully achieved in translation. However, we have found two instances (C.S.EX 3 and C.S.Ex 4) in which the lack of a relative clause has been replaced by an adverbial locution ("sin embargo") and a prepositional one ("en mi opinión"), which introduce the contrastive element and provide end-focus. Thus, "sin embargo", as an adversative coordinator, bridges the semantic meaning and pragmatic information of the sentence, activating a series of instructions on how the sentence has to be interpreted (Maričić & Đurić 2011: 93). In addition, "en mi opinión" provides end-focus to the sentence and sounds more natural and oral than a relative clause such as "lo que pienso es que" or "lo que creo es que". Therefore, these types of locutions work well as an alternative to equational structures in order to translate cleft-sentences and their prosodic-illocutionary force.

Tone: fall

As we know, the falling tone is, generally, the most frequently used one by English native speakers, its use being very ample. This is the reason why we selected only 24 examples to comment on in our analysis, which include tag questions, exclamations, assertions and questions. Tag questions will have their own subsection later, so we will only comment here on the strategies and success rate for the remaining examples.

We have seen that the strategies to render falling tone in Spanish voice-over are varied, and so is their success; for instance, there are 5 cases of high fall (emphatic meaning) which have been translated without resorting to any emphatic particle, syntactic reorganisation, or emphatic intonation, thus failing to match the pragmatic meaning of the source text, as shown in examples F.T.Ex 5 and F.T.Ex 9. Likewise, we also noticed that the rendition of the original pragmatic meaning can be jeopardised during the delivery of the target text when an unnatural intonation is used, as happens in examples F.T.Ex 4 and F.T.Ex 7 labelled as "dubbitis".

On the other hand, we saw that fall exclamatives in English can be rendered in Spanish using interrogatives (preferably with emphatic intonation) to convey the illocutionary force of the original, as shown in F.T.Ex1 and F.T.Ex2.

Concerning the rendition of falling assertions, their rendition poses no problem, since their use and value are similar in English and Spanish. However, given the popularity and spread use of uptalk[94] among the young English-speaking population - mostly females – (Setter 2019: 80), translators should be aware, when translating texts showing young people speaking, that declarative falls can sometimes be disguised as ring tones.

Tone: rise

There are two translation strategies to render rising intonation in our corpus (leaving tag questions aside) that have failed, in our view, to convey the pragmatic meaning of the original text: translating a rising question as a declarative statement (see R.T.Ex 3), and translating a rising declarative (functioning as a question in the original) as an exclamative (see R.T.Ex 4). On the other hand, we have seen the opposite case, a rising declarative (functioning as a question in the original) rendered as a question in the target text, which successfully renders the prosodic-illocutionary force of the source text. In addition to this, there were instances of rising tones used to signal that more information was yet to come, which were successfully rendered in the Spanish text by means of

---

[94] Uptalk occurs when the intonation of someone's speech goes up at the end of (mainly) statements instead of down (Setter 2019: 77).

falling tones, each clause or IP being divided by a comma or a full stop (we do not know for certain since we had the scripts transcribed and did not work with the original translator's script).

Tone: tag questions

There are several translation strategies for tag questions within our corpus showing mostly a satisfactory result in terms of illocutionary force rendition. Most of the tags have been translated as questions (8 out of 13). From these, both falling and rising tags may be translated as ratified questions (using "¿sí?", "¿no?", "¿verdad?", etc.), which is, in our corpus, the commonest form of translation. Falling tags may be rendered as non-neutral interrogative sentences[95], which in Spanish confirm a hypothesis, and their function is other than requesting information, as in T.Q.Ex 4. Finally, rising tags can be translated as standard questions since they usually expect an answer from the other speaker in the conversation.

Translating tag questions (both rising and falling) as questions coincides with the answers given by professional voice-over translators in our survey, as shown in the figure below:

7. Una frase como por ejemplo: "she's the new teacher, isn't she?" (tag question) la traduzco siempre como una pregunta al español.
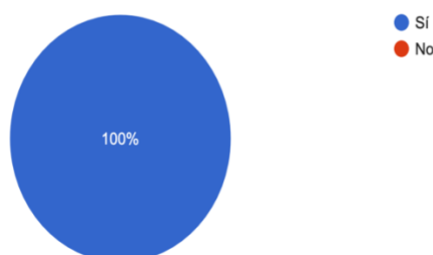4 responses

● Sí
● No

100%

Figure 44. Answer to question 7 in the translators' survey

[95] As stated by the RAE (2011: 469) "la función pragmática de la interrogación no siempre se circunscribe a la demanda de información: se puede usar para dar órdenes, para formular ruegos, para confirmar una hipótesis o para expresar sorpresa."

Nonetheless, English tag questions can also be translated into Spanish by using non-interrogative constructions, still rendering the illocutionary force implicit in the tone direction of the tag. Thus, falling tags, which do not request information, but, in fact, serve as confirmation signals, can be rendered as statements in Spanish, such as T.Q.Ex 3 and T.Q.Ex 4. On the other hand, rising tags which are not rendered as interrogatives fail to render the illocutionary force, since these always have a requesting component and expect a real answer from the receiver, as shown in examples T.Q.Ex 5 and T.Q.Ex 6.

In the end, when dealing with tone direction both in tags and other sentence types, translators should be able to resort to the strategies available in the target language, Spanish in our case, to render the prosodic-illocutionary force of the original, regardless whether these are syntactic, intonational or semantic. The aim is to identify the pragmatic meaning of the tone in the source text and convey it in the target text respecting the oral nature of voice-over discourse and using the most suitable resource in Spanish for the English tone in question.

Tonicity: narrow focus

There are two possibilities when it comes to narrow focus in English: occurring as a consequence of anaphora rule, by which any information that is known or given is deaccented, or as a result of the speaker's contrastive intention, in which case, the nuclear tone is placed on purpose on a specific word. For the first kind, the translator's task is twofold: to render the original deaccented word using the different strategies available (deixis, ellipsis, reaccenting) and to highlight the nucleus by means of an emphatic particle, marked intonation (voice actor), or by means of syntactic/structural reorganisation placing the focused item at the end of the sentence so that it may receive the nucleus, as in a broad focus construction, or resorting to clefting or topicalisation. If narrow focus is the result of contrastive focus, translators just have to render that contrast highlighting the relevant part of the message. As we saw in section 5.2.5.1 (figure 39), positive rendition of contrastive focus only appeared 52% of the time, which, consequently, tells us that the illocutionary force has not been successfully transmitted in almost half of the contrastive focus instances (48% of the time). The rendition strategies are the same as for narrow focus resulting from deaccenting,

although the rendition success differs (69% fully rendered for deaccenting against 52% for contrastive focus). This could be due to translators not being able to identify the prosodic marking of highlighted information in the source text. In addition, most of the cases identified in our corpus show a literal or almost literal translation, a few of them (5 out of 13) being delivered with emphatic intonation, which shows that translators and voice actors, probably, have difficulties conveying the pragmatic meaning of the original text in these cases.

- *Answer to objective b)* [To identify the way discourse markers are considered in voice-over translation.]

In the light of the results of our analysis in section 5.2.7, we can say that in English-Spanish voice-over translation there is no rule of thumb regarding the rendition of discourse markers. We saw that for almost every voice-over genre (documentaries, reality TV, web videos and interviews videos) discourse markers are present in the target text more or less half of the time. As regards idiolect markers, as in the case of chef Gordon Ramsey in the *Kitchen Nightmares* show, these are kept most of the time in the target text, which endows the Spanish version, in our view, with a more authentic and faithful feel. Swearwords, on the other hand, are not generally kept in the target text, only a few times, and usually rendered as "softened" swearwords in the target language.

Considering that discourse markers are common natural elements in oral spontaneous speech and that voice-over is purely oral and mainly used to translate unscripted spontaneous texts, it seems reasonable to keep them in the target text. Thus, in our view, markers, such as fillers, swearwords, idiolect traits, etc., should appear in the voice-over when possible since they provide cohesion to the target text and are a natural way of transmitting orality, unlike the prefabricated one of other audiovisual modes, such as, dubbing; moreover, they create the feeling of authenticity expected by the audience.

- *Answer to objective c)* [To check how synchronic features can affect or be affected as a result of the translation process.]

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 seen in our corpus, they are, actually, frequently missing with no serious consequences in the final effect. For instance, more than half of the cases analysed showing no literal synchrony (53%) fully render the original message, and there were several cases where characters were out of sync in terms of voice quality but they nevertheless succeeded in delivering an equivalent linguistic translation.

As regards isochrony and sound bites, we have seen that there is not a 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. In fact, as the answers to our survey show, voice-over directors do not coincide unanimously when responding to how long sound bites should last, 1-2 and 1-3 seconds being the most repeated answers. Voice talents seem to be aware of this issue, and judging from their answers to the question on sound bites in the survey (see figure 45 below), they make them depending on the chances they have to fit the target text in the time frame set for each text.

Figure 45. Answer to question 11 in the voice talents' survey

Whether there is enough time to include sound bites in the final voice-over version depends on how trimmed the target text can be, compared to the original one. This adjusting or adaptation is one of the fundamental processes in voice-over translation. When we asked voice-over companies about this matter (figure 46), they mostly replied that the adjustment is done by translators themselves and that after the recording process, if something needs to be retouched and adjusted, it can be done in post-production (removing pauses and breaths from the audio file in order to shrink it if the time frame is too tight, or even elongating the video file so as to fit the Spanish recorded text, as happened in the web video *TN January 17* in our corpus). In addition to this, 3 out of 4 of the translators surveyed confirm that taking care of the adaptation is part of their duties as translators, as shown in the graph below:

6. Si traduzco para voice-over (voces superpuestas)
4 responses

● Me ocupo del ajuste para sincronización
● El ajuste lo hace otro profesional en el estudio
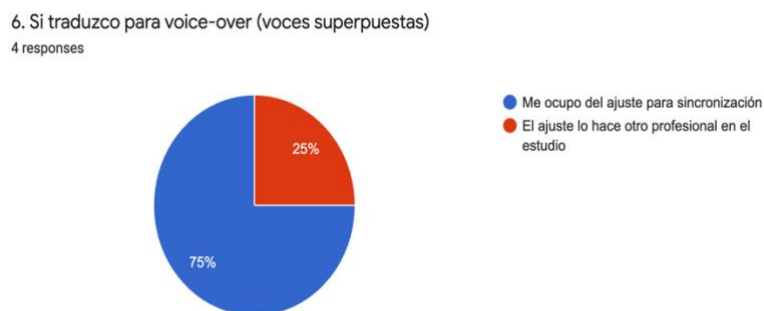
Figure 46. Answer to question 6 in the translators' survey

Nonetheless, 64% of the voice actors questioned state that it is rather common for voice-over scripts to be modified due to synchronisation problems, in the very studio room, during the recording session. In fact, none of the 14 professional voice talents answered this had never happened (first table in figure 47). Apart from that, 35% declare that a linguistic supervisor is never present in the studio during the recording session, 57% say that sometimes there is one, while 7% affirm there is always one (second table in figure 47).

5. ¿Alguna vez se ha cambiado el texto a locutar en la sesión de grabación por problemas de sincronización?
14 responses

- Sí, es frecuente.
- No, nunca me ha pasado.
- Raras veces.

35.7%

64.3%

7. ¿En las grabaciones para voice-over (voces superpuestas) hay un supervisor lingüístico?
14 responses

- Sí, siempre.
- Sí, a veces.
- No. Nunca.
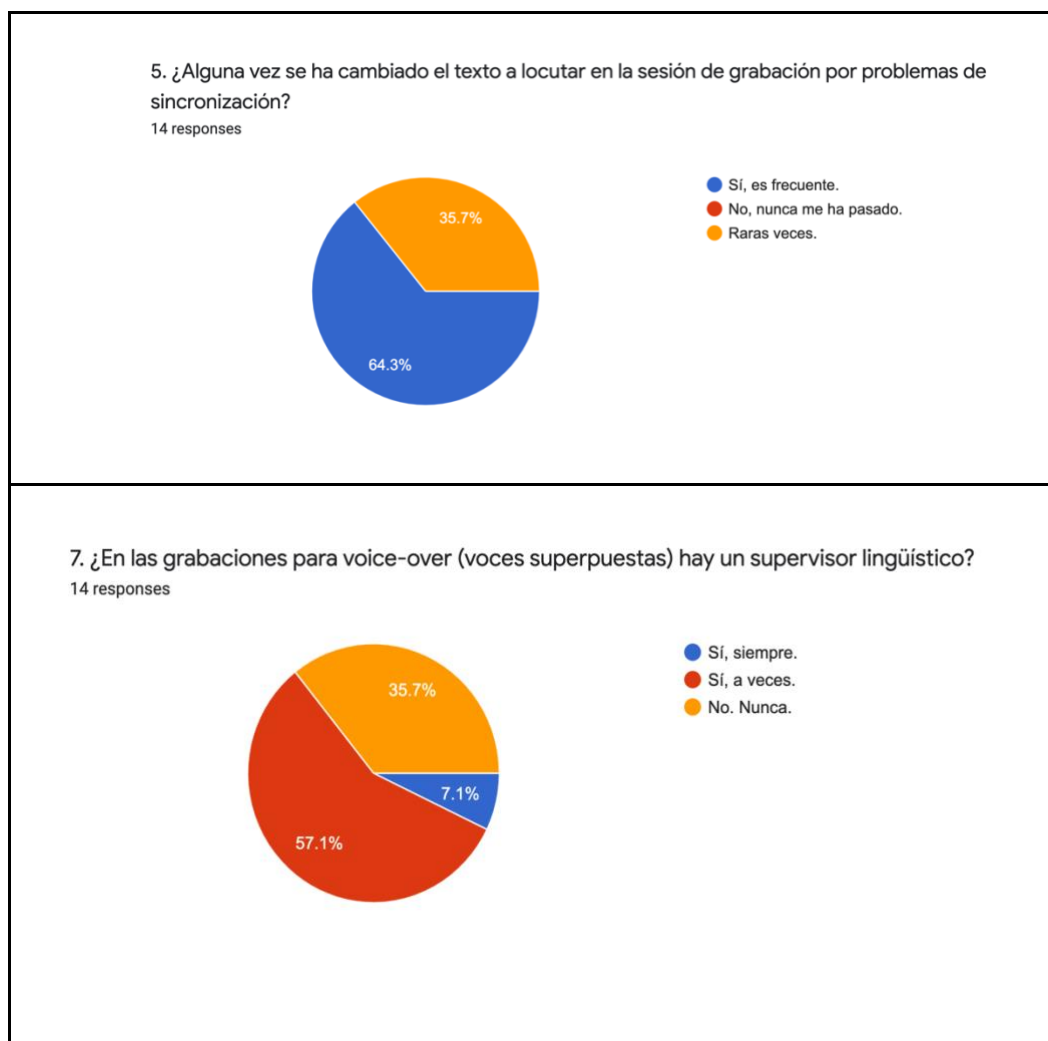
35.7%

7.1%

57.1%

Figure 47 Answers to questions 5 and 7 in the voice talents' survey

This double process of adaptation (that done by the translator and that done in the studio, mostly without linguistic supervisors) means that certain linguistic components which render the original prosodic-illocutionary traits (emphatic particles, linguistic reorganisations, deictic expressions, etc.) may be removed out of ignorance, resulting in a flat target text from a pragmatics perspective.

Content synchrony might be considered to be the least relevant of the synchronic characteristics, since the semantic simultaneity between the source and target text is generally expected in any translation mode. The cases in which there is a mismatch (no content synchrony) in our corpus are mainly due to omissions and modulations, which do not alter the basic semantic meaning. However, we have also

found some instances of mistranslations and false friends which do fail to render the original semantic message. In any case, whether content synchrony is achieved or not is something that only those who compare the two versions of the programme or texts will notice. The target audience will never know if there was a translation error, an omission or a modulation. Only when sound bites are present can the audience compare whether the two texts semantically coincide, but all the content in between these bites is unavailable and, therefore, the audience has to rely on the translator's expertise and reliability.

There is no doubt that the constraints imposed by the visual and temporal components of voice-over texts require specific translation procedures, as in any audiovisual translation mode, but, it is also evident that there is an aesthetic component in some of these (literal and character synchrony), whose purpose is mainly to enhance the audience's audiovisual experience, contributing to the illusion of authenticity and realism that characterises voice-over. Nevertheless, in our results, we can see that voice-over professionals do not seem to feel obliged to comply with these aesthetic synchronic components, which, in our view, implies that voice-over is a malleable audiovisual translation mode. This flexibility, which suggests that translators and production teams can easily adjust the target version to their needs, is probably the reason behind its huge success in recent years.

- *Answer to objective d)* [To check how voice quality is treated in voice-over translation.]

In line with the comments on character synchrony in the previous section, we consider voice quality as a fundamentally aesthetic component of voice-over translation, which can contribute to improving the artistic quality of audiovisual products by making the target text more reliable and authentic. From the data gathered in our corpus, female characters are voiced by female voices, and all male characters but one have been performed by male actors. There are mismatches in terms of pitch range and colour of voice between characters and voices of the same sex, as we have seen in section 5.2.6.5.1, some being more evident than others.

This leads us to the question about how voices are chosen for voice-over projects. The questionnaires supplied to voice actors and casting directors included certain questions on voice selection that are worth discussing at this point. There were three questions on this topic in the questionnaire provided to casting directors, whose answers are shown in the first table of the figure below.
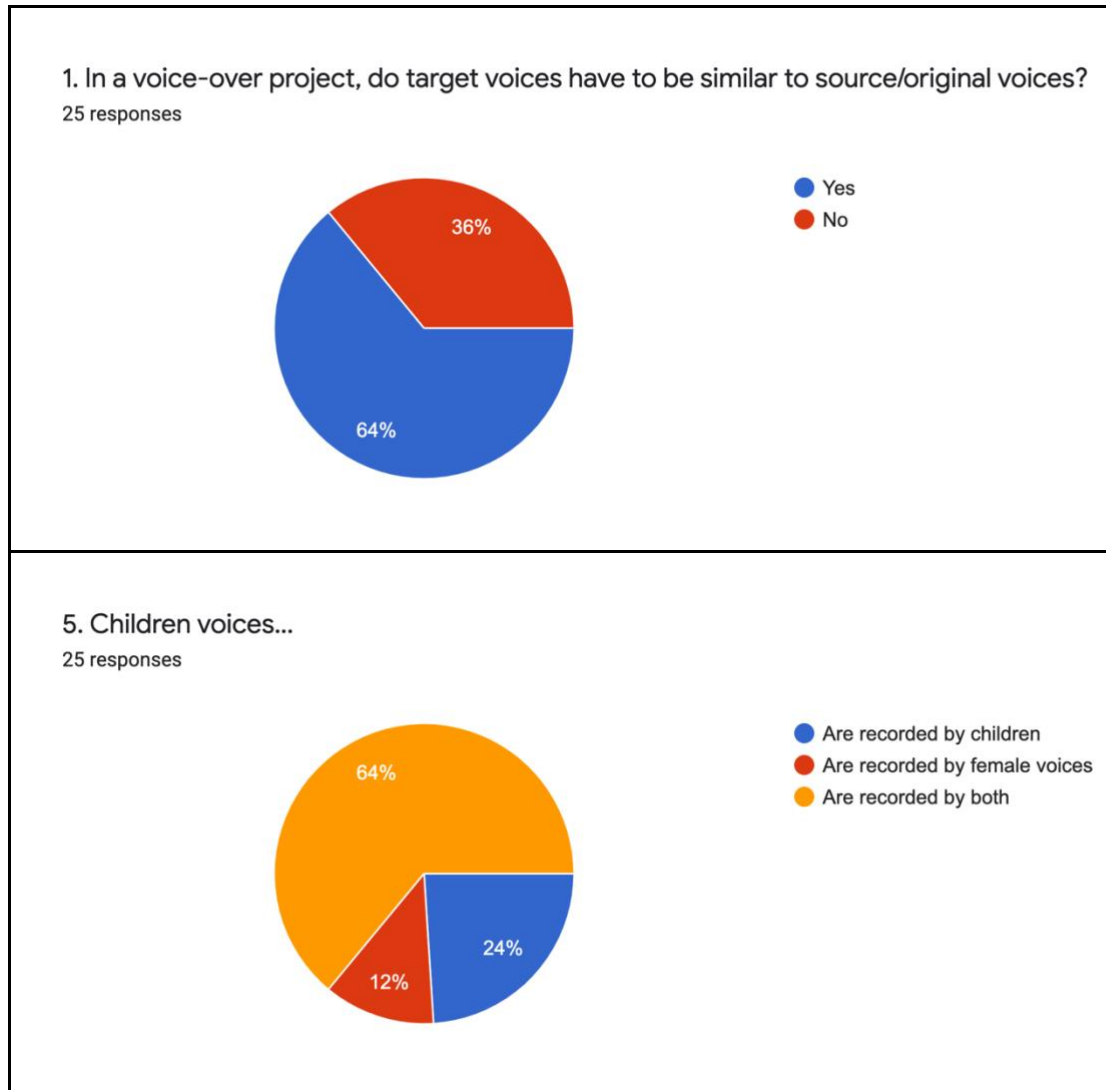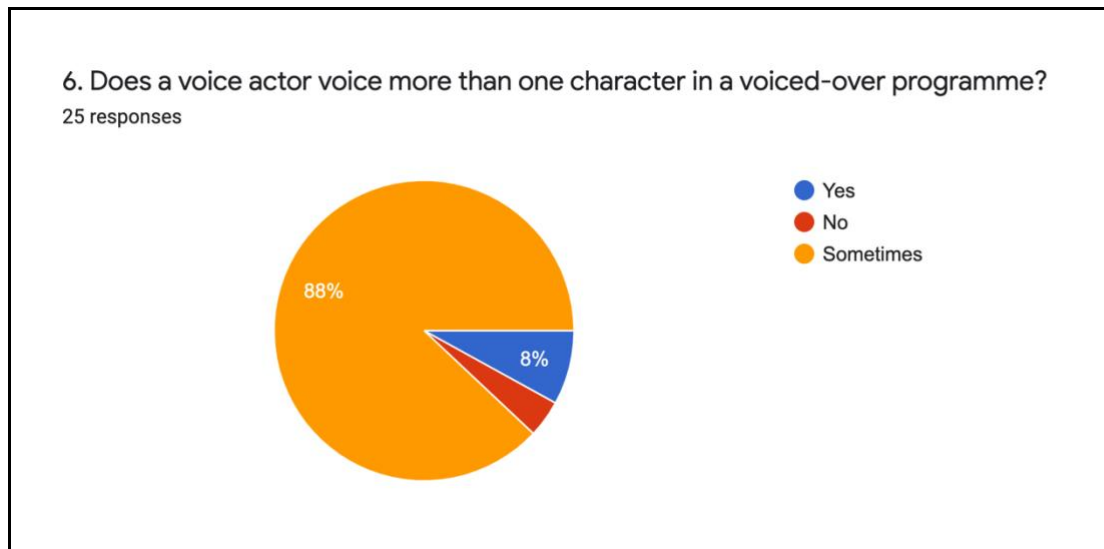
Figure 48. Answers to questions 1, 5 and 6 in the voice talents' survey.

As we can see from the data collected in the survey, most agents consider that voice selection is rather flexible. First, 64% consider that target voices do not have to be similar to the original ones, which could go against the aesthetic principle of character synchrony. Secondly, children's voices can either be voiced by children and female talents, meaning that boys are sometimes voiced by women. This is quite understandable considering that adult females and boys up to their teens have a similar fundamental frequency. Finally, the same voice actor may sometimes perform different characters for the same voiced-over programme, a situation we have identified in our corpus in the web video *TN Philosophy* and the two *Tasteology* documentaries. The rationale behind this is, almost certainly, the reduction of production costs, since fewer voices usually imply less budget spent.

In terms of the principles for choosing a voice over another for a voice-over project, the answers are varied (see question 2 in the survey), ranging from age, pitch, tone of voice, and sex, to versatility - so that the same talent can perform different voices for different characters -, experience, availability, fees, and, if working from a home studio, audio quality.

From a voice actor perspective, voicing characters with different voice qualities seems to be rather common, since 92.9% of the actors surveyed confirm they do it,

although varying their tone of voice (figure 49), thus confirming versatility as a professional requirement within the industry.

3. ¿Ha hecho voces de personajes que no tienen su edad y rango vocal (adolescentes, mayores...)?
14 responses

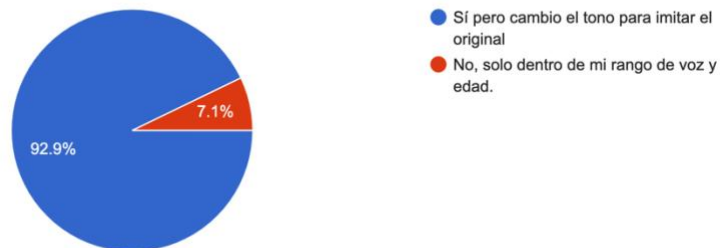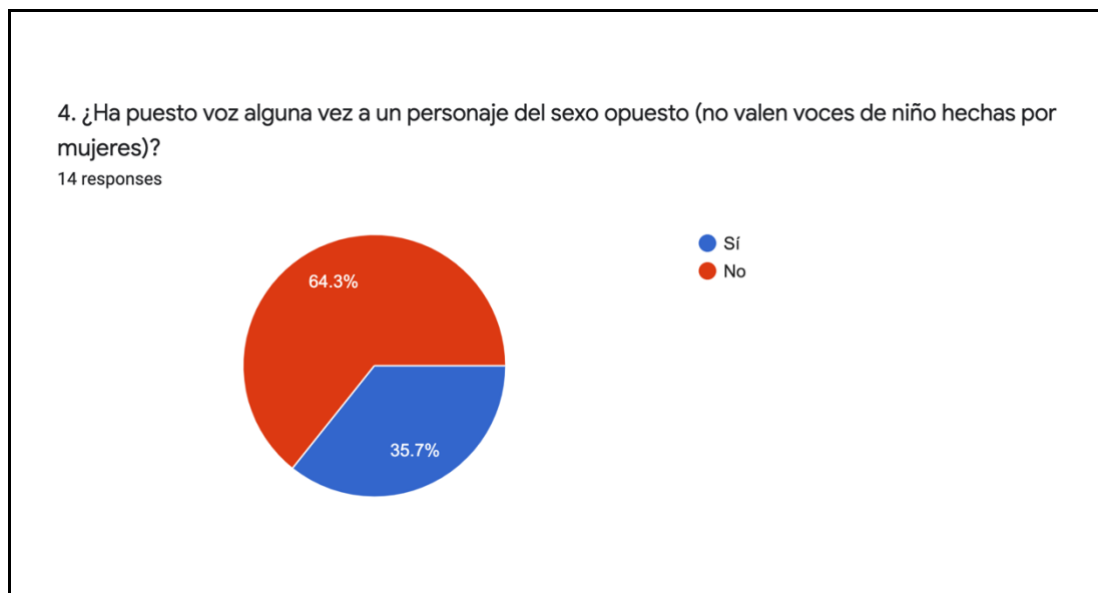Figure 49. Answer to question 3 in the voice talents' survey

In addition to this, 35.7% of the actors (5 out of 14) affirm having voiced characters of a different sex, and 50% consider that their voices are chosen according to the character being performed, as shown in the tables in figure 50 below.

4. ¿Ha puesto voz alguna vez a un personaje del sexo opuesto (no valen voces de niño hechas por mujeres)?
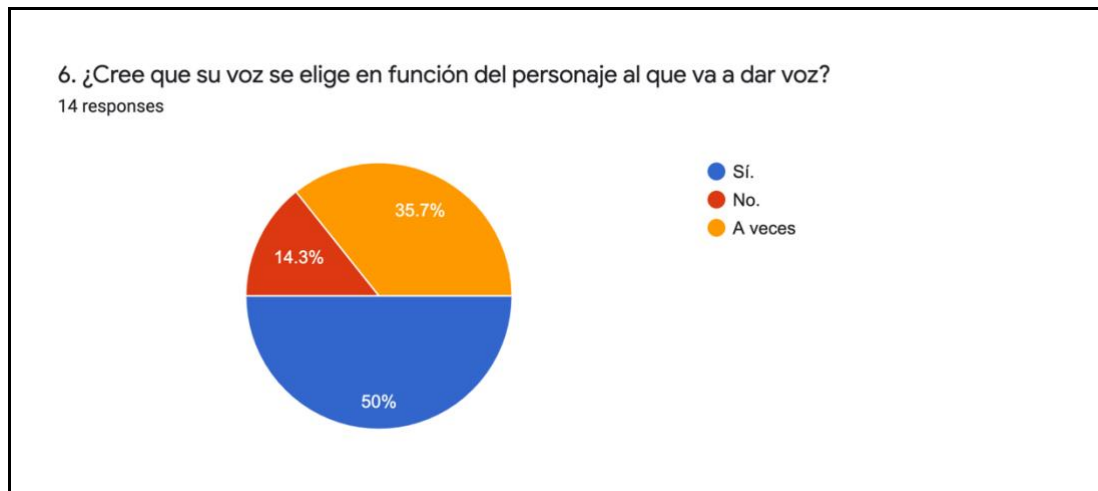14 responses

Figure 50. Answers to questions 5 and 6 in the voice talents' survey

In my experience as a professional voice-over talent and translator for over a decade within the audiovisual industry, and having taking part in the casting process (voice selection) for several voice-over productions, the standard norm is that characters and voices have to match in terms of sex category, with the exception of children, who are voiced by female talents if children actors are unavailable, as mentioned above. It is rather common for male teens to be performed by voice talents in their 20s, 30s and even 40s, if the latter's voice range still allows. The opposite situation, a young voice actor performing an elderly character, is, by contrast, not common and is usually avoided for aesthetic and authentic reasons.

All in all, voice quality seems to be a relevant aspect in the voice-over industry, given the aesthetic and enhancing effect that voices play in any audiovisual product and the fact that, behind each voice-over programme, there is a client who invests some good money in producing it. A wide vocal range is one of the most sought-after qualities that agencies and casting directors seek in voice actors, in this ever evolving and budget-limited industry.

- *Answer to objective e)* [To check how voice-over professionals (translators, actors, adjusters, directors) conceive prosody and prosodic traits in the process of a voice-over production.]

In order to answer this question, we will refer to the questionnaires (see annexe 4) in which numerous questions on prosody were posed to the professionals involved in voice-over translation.

Starting with translators, questions 1 to 5 directly refer to intonation and prosody. Thus, 75% of the translators who answered the survey confirm having studied English phonetics and phonology at some time during their graduate or postgraduate studies and affirm being aware of English intonational elements. When asked how they use these phonetic and phonological components in voice-over translation (see question 2), their answers are more in the line of phonetic synchrony, since one of the translators tries to "look for terminology with a similar phonetic realisation", another aims at "keeping the right pronunciation as in the source language", and another uses it "so that voice actors can pronounce English names". Moving now to how they use prosodic elements in translation (question 4), their strategies range from using prosodic elements which are similar in Spanish, to aiming at a correct Spanish intonation, or looking for tone intensity changes in the pronunciation of the source text. From these answers, we consider that only the last one has to do with English prosody as such, since it refers to tone variation and intensity.

If we now analyse the answers given by the voice actors on the use of intonation (question 1 in figure 51 below), we can see that 42% of the actors follow the original intonation during the delivery of the target text, while 50% adapt it to the client's requirements, which may be varied (sometimes they want a flat intonation, other times a more "acted" performance). What can be drawn from this answer is that 42% of the actors that follow the original intonation are probably more prone to using a dubbitis-like intonation in their performances than those who do not.

1. ¿Qué tipo de entonación usa para grabaciones de voice-over (voces superpuestas)?
14 responses

- Sigo la entonación original
- Uso una entonación neutra
- La que me diga el cliente
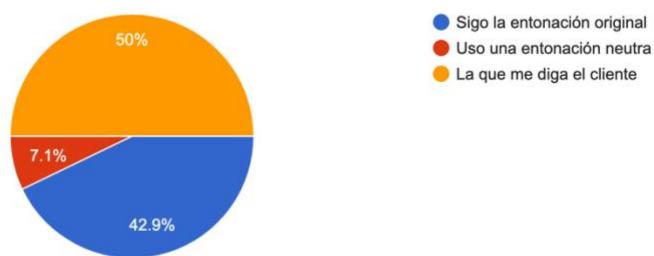
50%

7.1%

42.9%

Figure 51. Answer to question 1 in the voice talent's survey

Voice-over directors provided a different range of answers when asked if Spanish voices should imitate the original intonation (see question 9 in the questionnaire). Most of the directors (11 out of 25) consider that the original intonation should be imitated in the target text if it makes sense and does not sound ridiculous. There are, however, others who think it is not a good idea (5 out of 25) since each language has its melody. Some answers relate the decision to the type of programme and the client's suggestion on whether to follow the original intonation or not. In line with this, questions 7 and 8 refer to the type of delivery for different voice-over categories and whether voice should sound more "acted" depending on the product being recorded. Thus, for documentaries, the most repeated answer is that voices have to sound neutral, followed by "similar to the original ones", and "acted but not much". However, a neutral delivery does not necessarily mean it has to be flat and monotonous, since that could end up ruining the programme. Baños (2019: 278) considers that unlike in voiced-over reality TV programmes, in voiced-over documentaries, emotion is not conveyed, "as such information can be gathered from the original soundtrack and priority is given to what has been said". This can be true for some documentaries, but, in our view, this is not the case of the documentaries we have studied, since *Jamie's comfort food* and *The last days of Anne Boleyn* are quite emotional, being the former delivered almost as if it was a reality TV show.

Finally, regarding foreign dialects and accents, we can see that the commonest form of delivery, according to the professionals in our survey, is to use a neutral Spanish

accent (84%) known as Castilian Spanish[96], while translators declare resorting to both levelling and using Spanish dialects, in equal terms, to render these regional features.
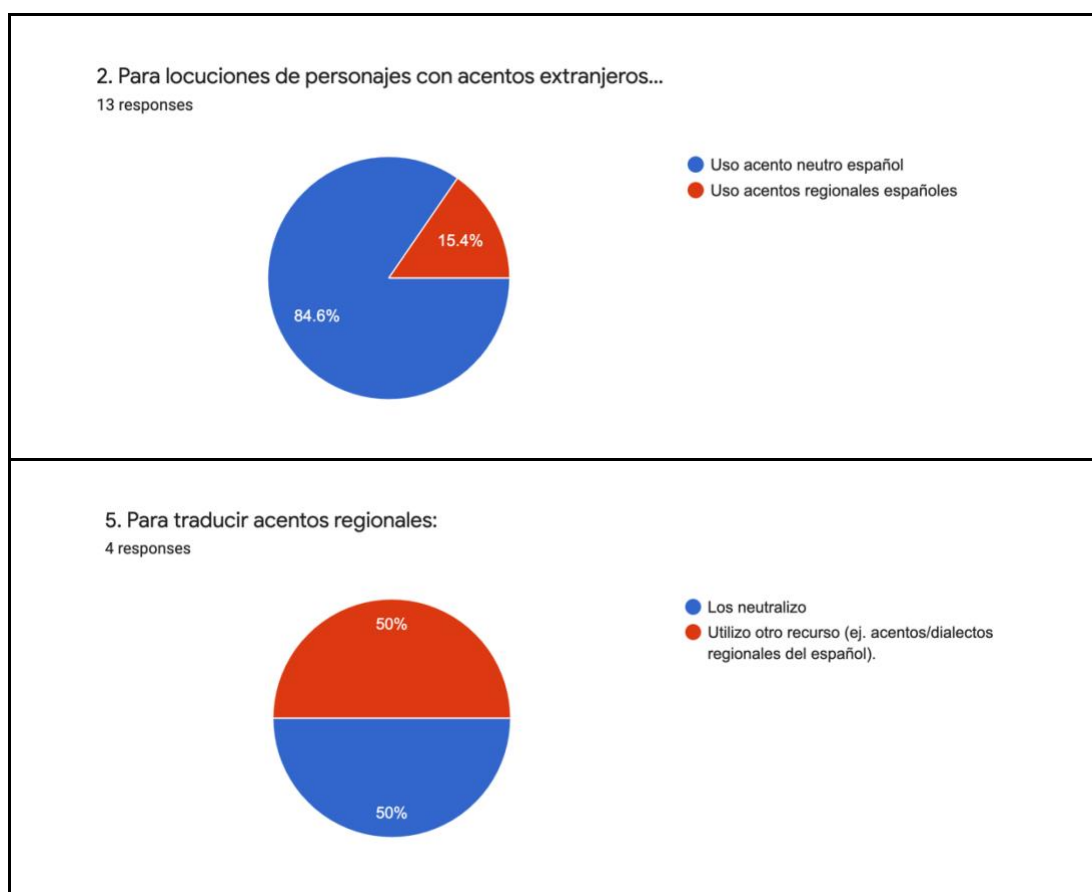
Figure 52. Answers to questions 2 (for the voice talents) and 5 (for translators)

During the analysis of the target texts we did not find any character whose voice accent sounded regional or non-standard, which confirms what has been declared by most of the voice talents. Although half of the translators admit to resorting to regional Spanish dialects in order to render English accents, if the voice actor does not apply an accent when delivering a translated script which contains dialectical expressions, the regional touch will not be perceived by the target audience.

[96] This should not be confused with the variety known as "neutral Spanish" ("español neutro") in the AVT industry, in order to refer to a variety of Latin American Spanish, which is a standard in all Spanish-speaking countries in America (Llorente Pinto 2006).

6.2 Hypothesis fulfilment

In section 4.5 of this dissertation we stated the fundamental question of this work: are translators fully aware of the relevance of English prosodic qualities and of the different resources the Spanish language offers to convey an equivalent illocutionary force in the target text?

After having analysed over 250 examples of utterances with pragmatically relevant prosodic traits, it is safe to affirm that translators are not fully aware of all the intonational and prosodic features that the English language has in order to convey the communicative value of utterances and the possibilities that the Spanish language offers so as to convey a functionally equivalent message. A significant example can be found in 48% of the instances of contrastive focus in English, whose illocutionary force is not rendered in the target text. And yet, intonational contrastive focus is probably one of the most obvious prosodic resources within English intonation, if we compare it with, for instance, deaccenting or intonational subordination.

Thus, the results offered in this dissertation illustrate that, oftentimes, voice-over does not provide a "good" translation, if our understanding of a "good translation" is one that conveys the speaker's intention, as well as the implicatures of the original message, as stated by Mateo Martínez (1996: 83-5). Evidently, this situation does not seem to match the standard and assumed definition of voice-over as a faithful and authentic representation of the original text, since prosody, as has been repeatedly claimed and shown in this dissertation, conveys pragmatic meaning, and this is not always rendered in Spanish voice-over.

In our view, in line with Mateo's reflection on the relevance of intonation teaching in translation studies programmes (2014: 131), raising the awareness of the illocutionary component of prosodic elements in source languages and the challenges they may pose when finding a functionally equivalent construction (syntactic, semantic or intonational) in the target language will certainly improve the quality of translations and, in the case of voice-over, the whole audiovisual experience. We, therefore, suggest that a new synchronic feature could be applied to the translation of voice-over texts:

prosodic-illocutionary synchrony, so that not only the suprasegmental features of language with pragmatic meaning should be considered, but also those prosodic ones pertaining to voice quality. In this way, the aesthetics of this AVT mode will be improved, because as Setter (2019: 85) puts it: "the voice does matter, whether we like it or not".

6.3 Final remarks

In this work we have highlighted those prosodic elements in English that serve a relevant communicative purpose and have analysed and discussed how they can be tackled in Spanish voice-over translation. We did not find examples for *defining/non-defining relative clauses* (Mott 2011: 223; Wells 2006: 202), but the most relevant intonational elements have, in our opinion, been described and analysed (narrow focus, tone direction, reaccenting, and cleft sentences), so we expect that this analysis will provide some food for thought to audiovisual translators, translation researchers and industry professionals in general.

As voice-over translation is a collaborative process in which not only translators but also voice actors, casting agents and directors are involved, it seems pertinent to have the responsibility for achieving a successful voiced-over product shared amongst them. Thus, if scripts are translated literally (almost word for word), it would be a good idea to inform the actors that that has been the case, so that their delivery may match the emphasis heard in the source voice but without transgressing the limits of Spanish intonation, as commonly happens in dubbitis. Also, if scripts need to be modified in the recording studio for isochronic purposes, those in the room who are responsible for such revision should be careful not to delete anything that could alter the rendition of the illocutionary force.

As a final remark, and to offer a brief commentary on the future of the AVT industry after more than a decade in the business, one thing seems clear to me: things will change. This industry is in constant evolution, and as technology develops, industry

professionals have to adapt to the ever-changing situations. Collaborative cloud-based audiovisual solutions are already in motion, and dubbed and voiced-over programmes can be produced (translated, adapted and voiced) from home[97] without depending on traditional recording studios. Moreover, new software [98] powered by AI (artificial intelligence) is already providing lip-sync translation solutions by means of face recognition and lip movement substitution without altering the original video file, showing excellent results[99]. This means that, in a few decades, once such technology becomes available to most audiovisual companies, the way dubbing and voice-over is done today will probably be obsolete, just like what happened with street phone booths after mobile phones stormed the telecommunications market. If this new technology succeeds, the current limits imposed by visual and temporal constraints will be broken and the current technical features that characterise audiovisual texts, such as isochrony and kinetic synchrony, will no longer matter or imply an obstacle in the translation process; the same applies to phonetic synchronisation or lip-sync, since every actor in every translated movie, programme, etc. will articulate in the target language as if they were native speakers. It will be interesting to see how the prosodic features analysed in this dissertation can be rendered using such new technology. But that task will have to wait for now.

---

[97] Zoodigital is a cloud-based AVT solutions company: https://www.zoodigital.com/services/localize/dubbing

[98] AI Translation company website: https://www.synthesia.io/ and an article on its technology: https://www.fxguide.com/fxfeatured/synthesia-lip-sync/

[99] News articles on this new technology technology: https://www.fxguide.com/fxfeatured/synthesia-lip-sync/; and: https://techcrunch.com/2019/04/25/the-startup-behind-that-deep-fake-david-beckham-video-just-raised-3m/

6 Conclusiones

El capítulo final de esta tesis resume las conclusiones obtenidas de los resultados en la sección de análisis y los cuestionarios respondidos por los profesionales involucrados en el proceso de traducción para voces superpuestas (traductores, actores y directores de casting). Como se indica en el capítulo 4.4.2.3, se han proporcionado tres cuestionarios diferentes a traductores, actores y directores de casting (uno por grupo) con el fin de obtener una visión más amplia de lo relevante que es la prosodia en el proceso de traducción profesional para voces superpuestas. Así, hemos obtenido 25 respuestas de directores de casting, 14 de actores de voz y solo 4 de traductores. Hubiéramos preferido tener más respuestas de traductores, pero responder encuestas con fines académicos es una tarea voluntaria, y no se ha podido hacer nada más para incrementar el volumen de participación. Hemos de recordar al lector que, como se indica en la sección 4.4.2.3, la información proporcionada por traductores y actores está en español, mientras que la información proporcionada por los directores y agentes de casting está en inglés, ya que la mayoría de ellos trabajan para empresas de localización internacionales fuera de España.

Comenzaremos ahora a responder a los objetivos establecidos en el capítulo 4; a continuación, comprobaremos si nuestra hipótesis inicial se ha cumplido o no; y, para concluir, se presentarán algunas observaciones finales sobre el futuro de la traducción de voces superpuestas.

6.1 Cumplimiento de objetivos

En las etapas iniciales de este trabajo se fijaron cinco objetivos con el fin de analizar cómo las características entonativas y prosódicas de la lengua inglesa se representan en los programas de voces superpuestas en español. Estos son:

  f) Identificar la forma en que las características prosódicas relacionadas con el estrés y el tono, junto con las 3 Ts (tonalidad, tonicidad y tono) de Wells (2006: 06) se traducen al español en las voces superpuestas, y

si la fuerza prosódico-ilocucionaria original, implícita en el texto origen, se ha representado o perdido en la traducción final.

g) Identificar cómo se consideran los marcadores de discurso en la traducción para voces superpuestas.

h) Comprobar cómo las características relacionadas con la sincronización y sus tipos pueden afectar o verse afectadas como resultado del proceso de traducción.

i) Comprobar cómo se trata la calidad y cualidad de la voz en la traducción para voces superpuestas.

j) Comprobar cómo los profesionales de la traducción para voces superpuestas (traductores, actores, ajustadores, directores) conciben la prosodia y los rasgos prosódicos en el proceso de una producción de voces superpuestas.

- *Respuesta al objetivo a)* [Identificar la forma en que las características prosódicas relacionadas con el estrés y el tono, junto con las 3 Ts (tonalidad, tonicidad y tono) de Wells (2006: 06) se traducen al español en las voces superpuestas, y si la fuerza prosódico-ilocucionaria original, implícita en el texto origen, se ha representado o perdido en la traducción final.]

El capítulo 5 ofreció el análisis cuantitativo y cualitativo de cómo las características entonativas relacionadas con el estrés, el tono y las 3Ts (Wells 2006: 06) se tratan en la traducción para voces superpuestas en español. Ahora resumiremos las estrategias y recursos que los traductores han utilizado para traducir estas características entonativas y su idoneidad para cada caso.

Estrés: deacentuación

En términos de estrés, analizamos dos características típicas del inglés: la deacentuación y la reacentuación. Para la deacentuación, las estrategias de traducción encontradas han sido: deixis anafórica, reacentuación y elipsis. De estos, los más exitosos a la hora de representar tanto el significado semántico como la fuerza prosódico-ilocucionaria han sido la deixis y la elipsis, mientras que la reacentuación ha demostrado ser la menos exitosa. Esto se debe al hecho de que, en inglés, la

deacentuación de un elemento implica un foco estrecho en alguna parte de la oración, y la reacentuación, tan común en español, de las palabras deacentuadas en inglés, pueden atraer el foco si se colocan en posición final dentro de la oración. Sin embargo, si omitimos la palabra deacentuada en inglés por medio de elipsis o usamos un pronombre deíctico en posición temática para referirnos a ella, la transmisión de la fuerza ilocucionaria del mensaje original se puede lograr fácilmente.

Estrés: reacentuación

Los elementos reacentuados de los textos originales en inglés de nuestro corpus se han traducido al español por medio de deixis anafórica, deixis catafórica, elipsis y reacentuación (usar la misma palabra de nuevo). Los cuatro son válidos para representar el significado semántico y la fuerza prosódico-ilocucionaria, ya que, como se ve en la sección 5.2.1.2, los casos de reproducción parcial o nula en el corpus se debieron a una locución inadecuada o a una traducción errónea del texto original en términos de elección de palabras, como en los ejemplos R.Ex.5 y R.Ex.9. De estas cuatro estrategias, la más comúnmente recurrida ha sido la reacentuación, con 41 casos de 63, duplicando la segunda estrategia más utilizada, la elipsis, con 19 casos. Teniendo en cuenta que hay limitaciones temporales en la traducción de voces superpuestas, la elipsis o deixis (anafórica y catafórica) son, en nuestra opinión, más útiles en los casos de reacentuación para así poder ahorrar algunas palabras y sílabas si la isocronía es muy ajustada.

Tono: subordinación entonativa

La subordinación entonativa, como hemos visto, señala porciones de la oración con poca o ninguna información en el contexto de la conversación, y estas unidades son de naturaleza descartable (Mott: 2011: 229), y "están aparte de lo que estamos diciendo" (Wells 2006: 243). Esto implica que estos IPs no son realmente relevantes y podrían omitirse en el texto meta si el traductor es capaz de identificarlos. En nuestro corpus, las dos estrategias utilizadas para reflejar esta característica entonativa del inglés han sido: traducción del IP subordinado (8 casos) y eliminación del IP (3 casos). Ambas estrategias son adecuadas para representar el significado semántico del original, pero dado que estos IP no son sustanciales, en nuestra opinión, la mejor opción es eliminarlos

con fines isocrónicos. Sin embargo, la dificultad parece estar en la identificación de estos IPs cuando hay variación de tono, ya que, de los 11 casos de subordinación, 8 se han incluido en el texto meta, lo que puede implicar que su valor irrelevante probablemente no fue identificado por el traductor.

Tono: citas

Respecto a las citas, observamos que este es un recurso muy habitual en el discurso oral espontáneo en inglés. Las principales estrategias de traducción que se encuentran en nuestro corpus son: mantener la cita en el texto objetivo, y reportarla o parafrasearla. Hay un caso en el que la cita se repitió (reacentuó) en el texto origen, que se representó por medio de deixis en el texto meta, por lo que esta estrategia recae en la categoría de reacentuación en lugar de en las citas, aunque aquí también se ha cuantificado. De los 69 casos de citas de nuestro corpus, 45 se han traducido manteniendo la cita y 23 se han parafraseado o reportado. Parece no haber ninguna regla general establecida para saber cuándo se debe mantener una cita y cuándo no, o cuando los traductores optan por mantener dicha cita en el texto meta o deciden parafrasearla.

Del cuestionario proporcionado a los traductores, la pregunta 8 se dirigió a la traducción de citas, como se muestra en la siguiente figura, que muestra también las respuestas obtenidas:
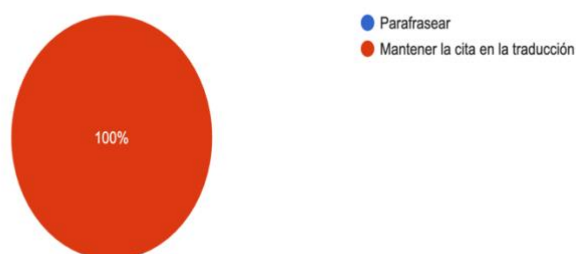

Figura 43. Respuesta a la pregunta 8 de la encuesta para traductores

Como se puede ver en la figura anterior, los cuatro traductores que respondieron a la encuesta, por lo general, mantienen la cita en el texto meta. Este resultado puede servir para explicar por qué la mayoría de las citas se han representado como tales en lugar de reformularse. Además, las citas funcionan como elementos demostradores y representan principalmente actos ilocucionarios (Clark y Gerrig 1990: 779). Esto significa que, si queremos mantener el mismo valor ilocucionario en el texto meta, las citas son la mejor alternativa. Por otro lado, parafrasear también sirve perfectamente para representar el significado semántico del original, pero en términos de fuerza prosódico-ilocucionaria el resultado no es el mismo que el obtenido manteniendo las citas. Tenemos que tener en cuenta la parte visual del texto a la hora de traducir citas, ya que la mayoría de las veces se puede ver a las personas en pantalla gesticulando o imitando a la persona que está siendo citada. Además, si hay *sound bites* que preceden o siguen a la traducción (en forma de paráfrasis), la sensación de autenticidad y realismo se puede perder ya que ambos textos (origen y meta) no coincidirán.

Tonalidad: división sujeto-predicado

A menudo, en inglés, el sujeto de una cláusula se separa del predicado con un solo IP para destacarlo, o porque el sujeto es largo y es preferible tener dos grupos de respiración en lugar de uno. En nuestro corpus hemos encontrado 6 casos de división de sujeto-predicado, que se han traducido de dos maneras en las voces-superpuestas en español: dividir el sujeto y el predicado como en inglés, o unir el sujeto y el predicado en un solo IP. Parece que no hay ningún problema a la hora de representar el significado semántico y la fuerza ilocucionaria del original con estas dos opciones. Sólo hay un ejemplo en el que la fuerza prosódico-ilocucionaria no se ha logrado plenamente ya que no hubo ruptura o insistencia en el sujeto para resaltarlo tanto como en el original.

Hemos notado que la mayoría de los casos de división sujeto-predicado en inglés se han traducido dividiendo el sujeto y el predicado en el texto meta (4 de 6), aunque esta no sea una característica común en español. Esto podría ser el resultado de la interferencia del texto origen durante la grabación de la traducción por parte de los locutores. Teniendo en cuenta que los locutores usan auriculares durante las sesiones de grabación para poder escuchar la versión original y saber cuándo iniciar y terminar sus líneas, esto podría llevarlos a hacer las mismas pausas que las voces inglesas

originales. Como sabemos, no hay necesidad de sincronización labial para voces superpuestas, pero la voz meta ciertamente necesita terminar un poco antes que la original, o, al menos, al mismo tiempo, nunca después, lo que puede ser la razón por la que obtenemos estas pausas en la versión española.

Tonalidad: Topicalización (temas y frases hendidas)

La topicalización es una excelente manera de marcar el énfasis en inglés y una de las estrategias más recurridas para hacerlo son las frases hendidas (tipo y tipo *wh* o *pseudo clefts* en inglés). En nuestro corpus, hemos identificado 11 casos de oraciones hendidas, y uno de pseudo hendida topicalizada (ver Top.Ex 1), para las cuales hay dos estrategias de traducción: el uso de una estructura ecuacional siguiendo el modelo de Braga & Rica Peromingo (2015) para traducir el énfasis estructural inglés en español (véase la página 121), y una traducción literal sin oración de relativo. Hay un caso en nuestro corpus en el que se cortó la línea (este es sólo un caso excepcional en el que se ha editado el vídeo español y se ha eliminado la escena). A partir de las dos estrategias mencionadas anteriormente, la más exitosa ha sido el uso de estructuras ecuacionales, que pueden formarse como construcciones de foco final, como *pseudo-clefts* en inglés, ("What... is...."/Lo que... es..."), o colocando el material resaltado en posición frontal (a la izquierda), seguido por el verbo "ser" y la cláusula relativa con la secuencia de contraste, como en "la magnitud de los cargos es precisamente lo que los hace más convincentes".

La otra estrategia que se ha utilizado para representar frases hendidas en voces superpuestas en español con un resultado menos exitoso ha sido el uso de una traducción literal sin oración de relativo. La cuestión aquí es que sin la cláusula de relativo que caracteriza a las *cleft-sentences* inglesas y a las estructuras ecuacionales en español, el elemento contrastivo puede perderse y, en consecuencia, la fuerza ilocucionaria no puede lograrse plenamente en la traducción. Sin embargo, hemos encontrado dos casos (C.S.EX 3 y C.S.Ex 4) en los que la falta de una cláusula de relativo ha sido sustituida por una locución adverbial ("sin embargo") y una preposicional ("en mi opinión"), que sí introducen el elemento contrastivo y proporcionan foco final. Por lo tanto, "sin embargo", como conector adversativo, une el significado semántico y la información pragmática de la frase, activando una serie

de instrucciones sobre cómo debe interpretarse la frase (Maričić & Đurić 2011: 93). Además, "en mi opinión" proporciona foco final a la frase y suena más natural y oral que una cláusula de relativo como "lo que pienso es que" o "lo que creo es que". Por lo tanto, este tipo de locuciones funcionan bien como una alternativa a las estructuras ecuacionales con el fin de traducir las oraciones hendidas (*cleft-sentences* en inglés) y su fuerza prosódico-ilocucionaria.

Tono: descendente

Como sabemos, el tono descendente es, generalmente, el más utilizado por los hablantes nativos ingleses, siendo su uso muy amplio. Esta es la razón por la que seleccionamos sólo 24 ejemplos para comentar en nuestro análisis, que incluyen preguntas ratificadas (*tag questions*), exclamaciones, aserciones y oraciones interrogativas. Las preguntas ratificadas tendrán su propia subsección más adelante, por lo que aquí solo comentaremos las estrategias y la tasa de éxito para los ejemplos restantes.

Hemos visto que las estrategias para traducir el tono descendente en las voces superpuestas en español son variadas, al igual que su éxito; por ejemplo, hay 5 casos de caída tipo *high fall* (significado enfático) que se han traducido sin recurrir a ninguna partícula enfática, reorganización sintáctica o entonación enfática, por lo que no coinciden con el significado pragmático del texto origen, como se muestra en ejemplos F.T.Ex 5 y F.T.Ex 9. Del mismo modo, también notamos que la interpretación del significado pragmático original puede verse comprometido durante la grabación del texto meta cuando se utiliza una entonación que no es natural en español, como sucede en los ejemplos F.T.Ex 4 y F.T.Ex 7 etiquetados como *doblajitis* o *dubbitis*.

Por otro lado, vimos que los tonos descendentes con fines exclamativos en inglés se pueden representar en español utilizando oraciones interrogativas (preferiblemente con entonación enfática) para transmitir la fuerza ilocucionaria del original, como se muestra en F.T.Ex1 y F.T.Ex2.

En cuanto a la traducción del tono descendente con fines aseverativos, su traducción no plantea ningún problema, ya que su uso y valor son similares en inglés y

español. Sin embargo, dada la popularidad y el uso extendido del *uptalk* entre la población joven de habla inglesa - en su mayoría mujeres - (Setter 2019: 80), los traductores deben ser conscientes, al traducir textos que muestran a jóvenes hablando, que los tonos descendentes con fines declarativos, a veces, pueden camuflarse como tonos ascendentes.

Tono: ascendente

En nuestro corpus hemos encontrado dos estrategias de traducción para reproducir el significado de la entonación ascendente (dejando a un lado las preguntas ratificadas o *tag questions*) que no han podido, a nuestro juicio, transmitir el significado pragmático del texto original: traducir una pregunta ratificada como una oración declarativa (véase R.T.Ex 3), y traducir una oración no interrogativa con tono ascendente (que funciona como una pregunta en el original) como una oración exclamativa (véase R.T.Ex 4). Por otro lado, hemos visto el caso opuesto, una oración declarativa con tono ascendente (que funciona como una pregunta en el original) traducida como una pregunta en el texto de destino, que reproduce con éxito la fuerza prosódico-ilocucionaria del texto origen.

Tono: preguntas ratificadas

Hay varias estrategias de traducción para las preguntas ratificadas dentro de nuestro corpus que muestran principalmente un resultado satisfactorio en términos de reproducción de fuerza ilocucionaria. La mayoría de las *tag questions* en inglés se han traducido como preguntas (8 de 13). A partir de estas, tanto las de tono ascendente como descendente pueden traducirse como preguntas ratificadas (usando "¿sí?", "¿no?", "¿verdad?", etc.), siendo, en nuestro corpus, la forma más común de traducción. Las *tag questions* con tono descendente pueden representarse como frases interrogativas no neutras, que en español confirman una hipótesis, y su función es diferente a la habitual solicitud de información, como en T.Q.Ex 4. Por último, las *tag questions* con tono ascendente se pueden traducir como preguntas estándar, ya que por lo general esperan una respuesta del otro hablante en la conversación.

La traducción de *tag questions* (tanto con tono ascendente como descendente) como preguntas coincide con las respuestas dadas por los traductores profesionales para voces superpuestas en nuestra encuesta, como se muestra en la figura siguiente:

7. Una frase como por ejemplo: "she's the new teacher, isn't she?" (tag question) la traduzco siempre como una pregunta al español.
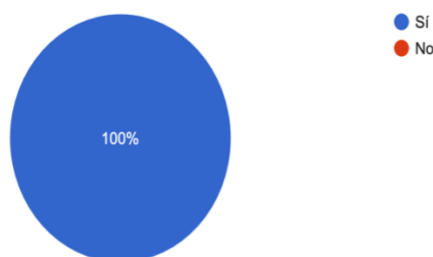
4 responses

- Sí
- No

100%

Figura 44. Respuesta a la pregunta 7 de la encuesta para traductores

Sin embargo, las *tag questions* en inglés también se pueden traducir al español mediante el uso de construcciones no interrogativas, y transmitir la fuerza ilocucionaria implícita en la dirección del tono de la *tag question*. Por lo tanto, las *tag questions* con tono descendente que no solicitan información, si no que, de hecho, sirven como señales de confirmación, pueden representarse como oraciones declarativas en español, como en el caso de los ejemplos T.Q.Ex 3 y T.Q.Ex 4. Por otro lado, las *tag questions* con tono ascendente que no se representan como oraciones interrogativas, no transmiten la fuerza ilocucionaria, ya que éstas siempre tienen un componente interrogativo y esperan una respuesta real del receptor, como se muestra en los ejemplos T.Q.Ex 5 y T.Q.Ex 6.

Cuando se trata de la dirección del tono tanto en *tag questions* como en otro tipo de oraciones, los traductores deben poder recurrir a las estrategias disponibles en el idioma de destino, el español en nuestro caso, para representar la fuerza prosódico-ilocucionaria del original, independientemente de si son recursos sintácticos, entonativos o semánticos. El objetivo es identificar el significado pragmático del tono en el texto origen y transmitirlo en el texto meta respetando la naturaleza oral del discurso de las voces superpuestas y utilizando el recurso más adecuado en español para el tono inglés en cuestión.

Tonicidad: foco estrecho

Hay dos posibilidades cuando se trata de un foco estrecho (*narrow focus*) en inglés: que ocurra como resultado de la *anaphora rule*, por la cual cualquier información que se conoce o se da por sabida se deacentúa, o como resultado de la intención contrastiva del hablante, en cuyo caso, el tono nuclear cae a propósito en una palabra específica. Para el primer tipo, la tarea del traductor es doble: reproducir la palabra deacentuada original utilizando las diferentes estrategias disponibles (deixis, puntos suspensivos, reacentuación) y resaltar el núcleo por medio de una partícula enfática, entonación marcada (actor de voz), o por medio de reorganización sintáctica/estructural colocando el elemento focal al final de la oración para que pueda recibir el núcleo, como en una construcción de foco amplio, o recurriendo a la hendidura. Si el foco estrecho es el resultado del foco contrastivo, los traductores sólo tienen que representar ese contraste resaltando la parte relevante del mensaje. Como vimos en la sección 5.2.5.1 (figura 39), la interpretación positiva del foco contrastivo sólo sucedió el 52% del tiempo, lo que, en consecuencia, nos dice que la fuerza ilocucionaria no se ha transmitido con éxito en casi la mitad de los casos de foco contrastivo (48% de las veces). Los recursos de traducción son los mismos que para el foco estrecho resultante de la deacentuación, aunque su éxito difiere (el 69% transmiten el valor origina en su totalidad para deacentuación, y sólo el 52% lo hace para el foco contrastivo). Esto podría deberse a que los traductores no pueden identificar el realce prosódico de la información resaltada en el texto meta. Además, la mayoría de los casos identificados en nuestro corpus muestran una traducción literal o casi literal, algunos de ellos (5 de 13) que se reproducen con entonación enfática, lo que muestra que los traductores y actores de voz, probablemente, tienen dificultades para transmitir el significado pragmático del texto original en estos casos.

- *Respuesta al objetivo b)* [Identificar cómo se consideran los marcadores de discurso en la traducción para voces superpuestas.]

A la luz de los resultados de nuestro análisis en la sección 5.2.7, podemos decir que en la traducción para voces superpuestas inglés-español no hay ninguna regla general con

respecto a la traducción de los marcadores de discurso. Vimos que para casi todos los subgéneros dentro de las voces superpuestas (documentales, reality TV, vídeos web y vídeos de entrevistas) los marcadores de discurso están presentes en el texto meta más o menos la mitad del tiempo. En cuanto a los marcadores característicos de idiolectos, como en el caso del chef Gordon Ramsey en el programa *Kitchen Nightmares*, estos se mantienen en el texto meta en la mayoría de los casos, lo que dota a la versión española, en nuestra opinión, de una sensación más auténtica y fiel. Las palabrotas, por otro lado, generalmente no se mantienen en el texto meta; sólo unas pocas veces, y, por lo general, se reproducen como tacos inofensivos o no tan groseros, en el idioma meta.

Teniendo en cuenta que los marcadores de discurso son elementos naturales propios del discurso oral espontáneo y que las voces superpuestas son un modo de traducción puramente oral que utiliza principalmente para traducir textos espontáneos sin guiones, parece razonable mantenerlos en el texto meta. Por lo tanto, en nuestra opinión, los marcadores, como *fillers*, palabrotas, rasgos idiolectales, etc., deben aparecer en la voz superpuesta cuando sea posible, ya que proporcionan cohesión al texto meta y son una forma natural de transmitir la oralidad natural, a diferencia de la oralidad prefabricada característica de otros modos audiovisuales, como el doblaje; además, crean la sensación de autenticidad esperada por el público.

- *Respuesta al objetivo c)* [Comprobar cómo las características relacionadas con la sincronización y sus tipos pueden afectar o verse afectadas como resultado del proceso de traducción.]

Las voces superpuestas, como sabemos, son un método de traducción restringido, con restricciones visuales y temporales que afectan al proceso de traducción, lo que significa que el texto traducido tiene que ser sincronizado con los límites impuestos por el texto origen. Estas características técnicas, aunque restrictivas y exigentes, pueden ser explotadas por los traductores para proporcionar una sensación de autenticidad y realismo al texto meta, como es el caso de la sincronización literal, cinética y de personajes. Sin embargo, estos aspectos sincrónicos no necesitan estar presentes o aplicarse todo el tiempo en la traducción para voces superpuestas, ya que, como hemos visto en nuestro corpus, en realidad, con frecuencia no se aplican, lo que no acarrea

consecuencias graves en el efecto final. Por ejemplo, más de la mitad de los casos analizados que no muestran sincronía literal (53%) reproducen completamente el mensaje original, y hubo varios casos en los que los personajes no estaban sincronizados en términos de cualidad de voz, y, sin embargo, lograron reproducir una traducción lingüística equivalente.

En cuanto a la isocronía y *sound bites*, hemos visto que no hay una duración de tiempo establecida para estas, y que sólo están presentes cuando es posible, lo que significa que estos *bites* -que permiten la sincronía literal- se sacrifican cuando hay que satisfacer las necesidades isocrónicas. De hecho, como muestran las respuestas a nuestra encuesta, los directores de voces superpuestas no coinciden unánimemente al responder cuánto tiempo deben durar los *sound bites*, siendo 1-2 y 1-3 segundos las respuestas más repetidas. Los locutores parecen ser conscientes de este problema, y a juzgar por sus respuestas a la pregunta sobre *sound bites* en la encuesta (véase la figura 45 a continuación), lo hacen dependiendo de si el texto meta cabe en el marco temporal establecido para cada texto por el vídeo original.
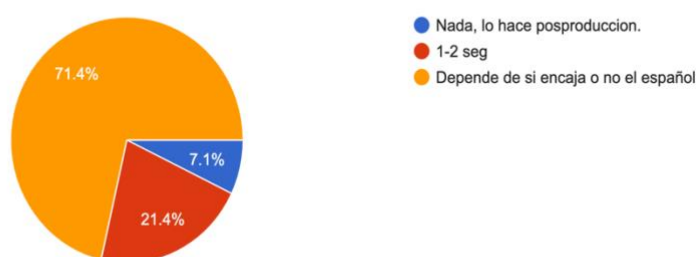

Figura 45. Respuesta a la pregunta 11 en la encuesta de locutores

La posibilidad de incluir *sound bites* en el texto meta depende de lo comprimido que pueda estar el guion en español a locutar, en comparación con el original. Este ajuste o adaptación es uno de los procesos fundamentales en la traducción para voces superpuestas. Cuando preguntamos a las empresas de traducción audiovisual sobre este asunto (figura 46), en su mayoría respondieron que el ajuste es realizado por los propios traductores y que después del proceso de grabación, si algo necesita ser retocado y ajustado, se puede hacer en postproducción (eliminación de pausas y respiraciones del archivo de audio con el fin de reducirlo si el marco de tiempo es demasiado ajustado, o

incluso alargar el archivo de vídeo con el fin de ajustarlo al texto grabado en español, como sucedió en el vídeo web *TN January 17* en nuestro corpus). Además de esto, 3 de los 4 traductores encuestados confirman que la adaptación es parte de sus funciones como traductores, como se muestra en el siguiente gráfico:
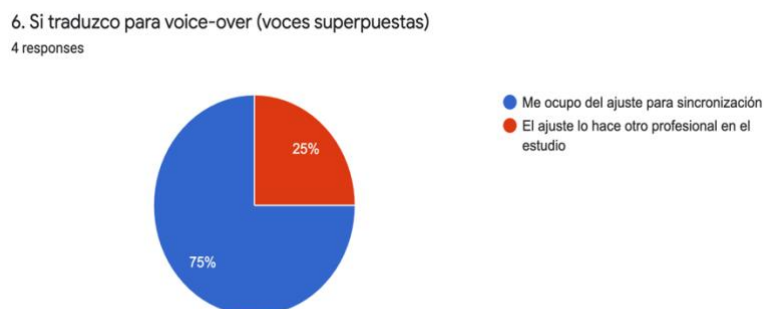

Figura 46. Respuesta a la pregunta 6 de la encuesta para traductores

Sin embargo, el 64% de los locutores encuestados afirman que es bastante común que los guiones a grabar se modifiquen debido a problemas de sincronización, en la misma sala de estudio, durante la sesión de grabación. De hecho, ninguno de los 14 locutores profesionales respondió que esto nunca había sucedido (primera tabla en la figura 47 abajo). Aparte de eso, el 35% declara que un supervisor lingüístico nunca está presente en el estudio durante la sesión de grabación, el 57% dice que a veces hay uno, mientras que el 7% afirma que siempre lo hay (segunda tabla en la figura 47).
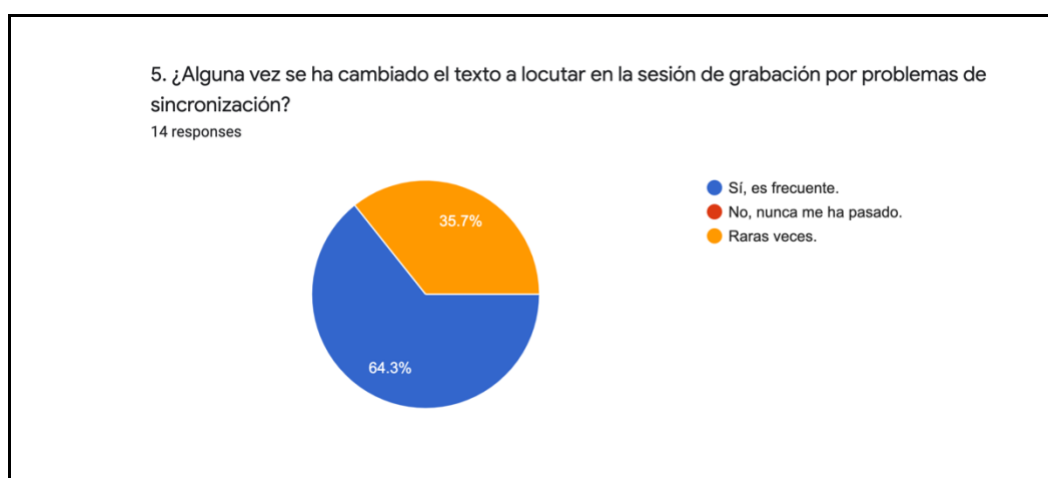
Figura 47. Respuestas a las preguntas 5 y 7 de la encuesta para locutores

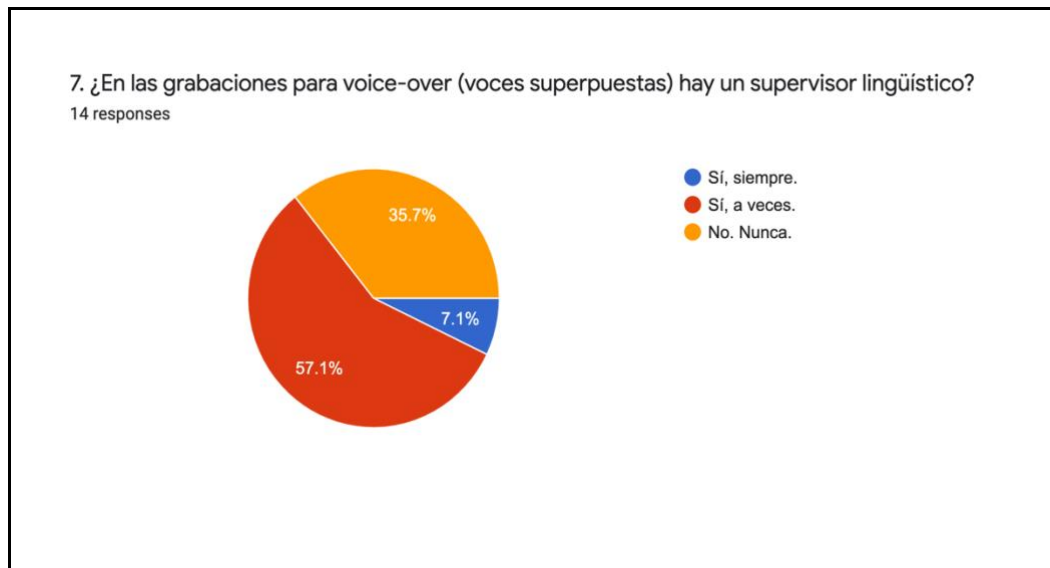Este doble proceso de adaptación (el realizado por el traductor y el realizado en el estudio, en su mayoría sin supervisores lingüísticos) significa que ciertos componentes lingüísticos que reproducen los rasgos prosódico-ilocucionarios originales (partículas enfáticas, reorganizaciones lingüísticas, expresiones deícticas, etc.) pueden ser eliminados por error o ignorancia, lo que resulta en un texto meta plano desde un punto de vista pragmático.

La sincronización de contenido puede considerarse la menos relevante de las características sincrónicas, ya que la simultaneidad semántica entre el texto origen y el meta se asume, generalmente, en cualquier modo de traducción. Los casos en los que hay una discordancia (sin sincronía de contenido) en nuestro corpus se deben principalmente a omisiones y modulaciones, que no alteran el significado semántico básico. Sin embargo, hemos encontrado algunos casos de traducciones erróneas y *false friends* que no representan el mensaje semántico original. En cualquier caso, si se logra o no la sincronización de contenido es algo de lo que sólo se darán cuenta aquellos que comparen las dos versiones del programa o textos. El público objetivo nunca sabrá si hubo un error de traducción, una omisión o una modulación. Sólo cuando hay *sound bites,* la audiencia puede comparar si los dos textos coinciden semánticamente, pero

todo el contenido entre estos *bites* no está disponible y, por lo tanto, el público tiene que confiar en la experiencia y fiabilidad del traductor.

No cabe duda de que las limitaciones impuestas por los componentes visuales y temporales de los textos de voces superpuestas requieren procedimientos de traducción específicos, como en cualquier modo de traducción audiovisual, pero, también, es evidente que hay un componente estético en algunos de ellos (sincronía literal y de personajes), cuyo objetivo principal es potenciar la experiencia audiovisual del público, contribuyendo a la ilusión de autenticidad y realismo que caracteriza este modo de traducción. Sin embargo, en nuestros resultados, podemos ver que los profesionales de las voces superpuestas no parecen sentirse obligados a cumplir con estos componentes sincrónicos estéticos, lo que, en nuestra opinión, implica que las voces superpuestas son un modo de traducción audiovisual moldeable. Esta flexibilidad, que sugiere que los traductores y los equipos de producción pueden ajustar fácilmente la versión meta a sus necesidades, es probablemente la razón de su enorme éxito en los últimos años.

- *Respuesta al objetivo d)* [Comprobar cómo se trata la calidad y cualidad de la voz en la traducción para voces superpuestas.]

En línea con los comentarios sobre la sincronía de personajes en la sección anterior, consideramos la calidad de voz como un componente fundamentalmente estético de la traducción para voces superpuestas, y que puede contribuir a mejorar la calidad artística de los productos audiovisuales al hacer el texto meta más fiable y auténtico. A partir de los datos recogidos en nuestro corpus, los personajes femeninos son locutados o interpretados por voces femeninas, y todos los personajes masculinos excepto uno han sido interpretados por actores masculinos. Hay desajustes en términos de rango de tono y color de voz entre personajes y voces del mismo sexo, como hemos visto en la sección 5.2.6.5.1, siendo algunos más evidentes que otros.

Esto nos lleva a la pregunta de cómo se eligen las voces para los proyectos de voces superpuestas. Los cuestionarios proporcionados a los locutores y directores de casting incluían ciertas preguntas sobre la elección de voces, que vale la pena discutir

llegados a este punto. En el cuestionario proporcionado a los directores de casting se formularon tres preguntas sobre este tema, cuyas respuestas figuran en la primera tabla de la figura que se muestra a continuación.
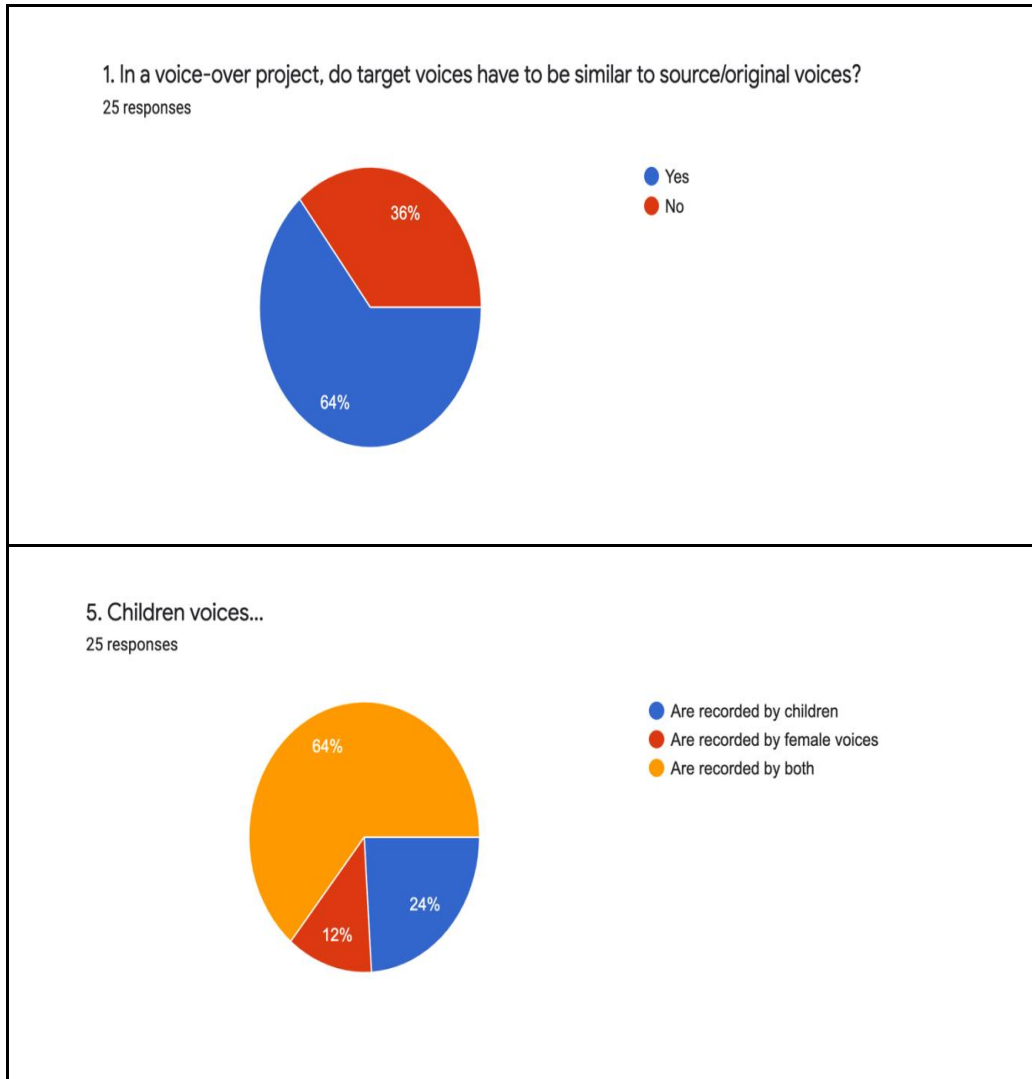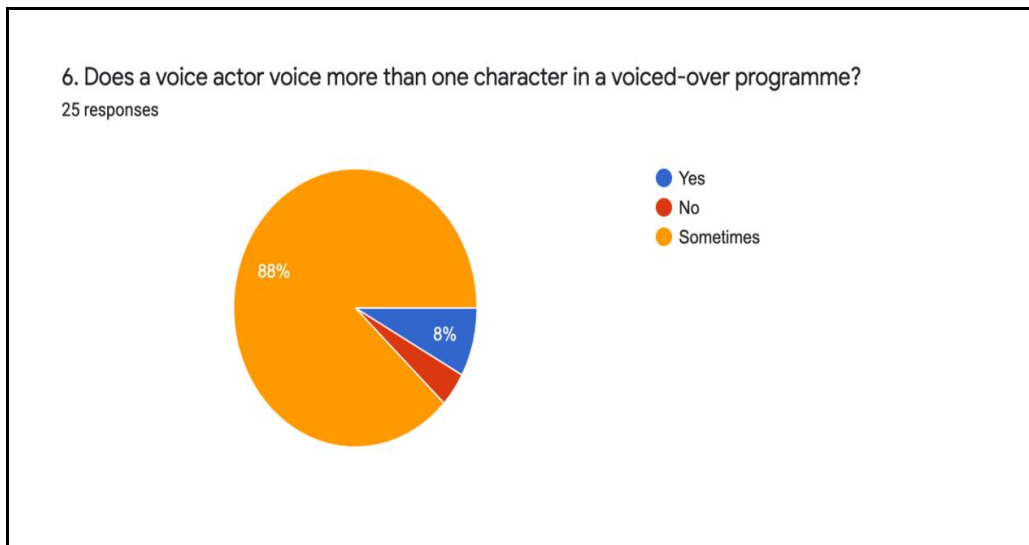
Figura 48. Respuestas a las preguntas 1, 5 y 6 de la encuesta para locutores.

Como muestran los datos recopilados en la encuesta, la mayoría de los agentes consideran que la selección de voces es bastante flexible. En primer lugar, el 64% considera que las voces meta no tienen que ser similares a las originales, lo que podría ir en contra del principio estético de la sincronía de los personajes. En segundo lugar, las voces de personajes infantiles pueden ser grabadas por niños y locutoras femeninas, lo que significa que los niños a veces son representados por voces de mujeres adultas. Esto es bastante comprensible teniendo en cuenta que las mujeres adultas y los niños, en general hasta su adolescencia, tienen una frecuencia fundamental similar. Por último, un mismo locutor puede, a veces, interpretar diferentes personajes para el mismo programa de voces superpuestas, una situación que hemos identificado en nuestro corpus en el vídeo web *TN Philosophy* y los dos documentales de *Tasteology.* La justificación a este hecho se debe, casi con toda seguridad, a la reducción de costes de producción, ya que, a menos voces, menos presupuesto invertido.

En cuanto a los principios para elegir una voz sobre otra para un proyecto de voces superpuestas, las respuestas son variadas (ver en anexo 4 la pregunta 2 en esta encuesta), que van desde la edad, el timbre, el tono de voz y el sexo, hasta la versatilidad - para que el mismo locutor pueda realizar diferentes voces para diferentes personajes -, experiencia, disponibilidad, tarifas y, si se trabaja desde un estudio en casa, calidad de audio.

Desde el punto de vista de los locutores, interpretar personajes con diferentes cualidades de voz parece ser bastante común, ya que el 92,9% de los actores encuestados confirma que lo hacen, aunque variando su tono de voz (figura 49), lo que confirma que la versatilidad es un requisito dentro de la industria.
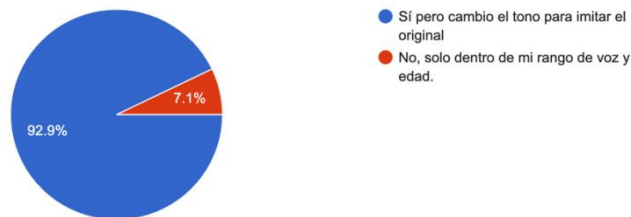

Figura 49. Respuesta a la pregunta 3 en la encuesta para locutores

Además, el 35,7% de los locutores (5 de 14) afirman haber interpretado personajes de diferente sexo, y el 50% considera que sus voces son elegidas de acuerdo con el personaje que se va a interpretar, como se muestra en las tablas de la figura 50 a continuación.

Figura 50. Respuestas a las preguntas 5 y 6 de la encuesta para locutores

En mi experiencia como locutor profesional y traductor durante más de una década dentro de la industria audiovisual, y tras haber participado en el proceso de casting para varias producciones de voces superpuestas, la norma estándar es que los personajes y las voces tienen que coincidir en términos de condición sexual, con la excepción de las voces infantiles, que son locutadas por voces femeninas si los actores infantiles no están disponibles, como se ha mencionado anteriormente. Es bastante común que las voces de adolescentes masculinos sean grabados por locutores adultos de entre 20, 30, e incluso 40 años, si el rango de voz de estos últimos todavía lo permite. La situación opuesta, de un locutor joven que interprete un personaje mayor, como un anciano, no es común y, por lo general, se evita por razones estéticas y de realismo.

Con todo, la cualidad y calidad de la voz parece ser un aspecto relevante en la industria de las voces superpuestas, dado el efecto estético y de buena calidad que las voces aportan a cualquier producto audiovisual y el hecho de que, detrás de cada programa de voces superpuestas, hay un cliente que invierte un dinero produciéndolo. Un rango vocal amplio y versátil es una de las cualidades más solicitadas por las agencias y los directores de casting cuando buscan locutores para proyectos de voces superpuestas. Esta es una industria en constante evolución y con un presupuesto limitado.

- *Respuesta al objetivo e)* [Comprobar cómo los profesionales de la traducción para voces superpuestas (traductores, actores, ajustadores, directores) conciben la prosodia y los rasgos prosódicos en el proceso de una producción de voces superpuestas.]

Para responder a esta pregunta, nos referiremos a los cuestionarios (véase el anexo 4) en los que se plantearon numerosas preguntas sobre prosodia a los profesionales que participan en la traducción para voces superpuestas.

Comenzando con los traductores, las preguntas 1 a 5 se refieren directamente a la entonación y la prosodia. Así, 3 de los 4 traductores que respondieron a la encuesta confirman haber estudiado fonética y fonología inglesa en algún momento durante sus estudios de grado o posgrado y afirman conocer los elementos entonativos del inglés. Cuando se les pregunta cómo utilizan estos componentes fonéticos y fonológicos en la traducción para voces superpuestas (véase la pregunta 2), sus respuestas están más en la línea de la sincronización fonética (o *lip sync*), ya que uno de los traductores intenta "buscar terminología con una realización fonética similar", otro tiene como objetivo "mantener la pronunciación correcta como en el idioma de origen", y otro la utiliza "para que los actores de voz puedan pronunciar nombres en inglés". Pasando ahora a cómo utilizan los elementos prosódicos en la traducción (pregunta 4), sus estrategias van desde el uso de elementos prosódicos que son similares en español, a apuntar a una correcta entonación española, o la búsqueda de cambios de intensidad de tono en la pronunciación del texto origen. A partir de estas respuestas, consideramos que sólo la última tiene que ver con la prosodia inglesa como tal, ya que se refiere a la variación e intensidad del tono.

Si ahora analizamos las respuestas dadas por los locutores sobre el uso de la entonación (pregunta 1 en la figura 51 mostrada a continuación), podemos ver que el 42% de los actores siguen la entonación original durante la locución del guion meta, mientras que el 50% la adapta a las necesidades del cliente, que puede ser variada (a veces quieren una entonación plana, otras veces una actuación más "actuada"). Lo que queda claro a tenor de esta respuesta es que el 42% de los actores que siguen la entonación original son, probablemente, más propensos a usar una entonación tipo *doblajitis* (*dubbitis* en inglés) que aquellos que no lo hacen.

1. ¿Qué tipo de entonación usa para grabaciones de voice-over (voces superpuestas)?
14 responses

- Sigo la entonación original
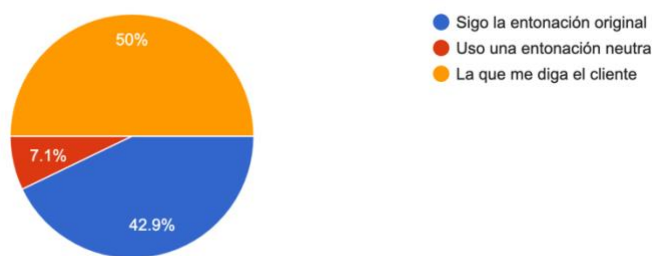- Uso una entonación neutra
- La que me diga el cliente

Figura 51. Respuesta a la pregunta 1 en la encuesta para locutores

Los directores de voces superpuestas proporcionaron una gama de respuestas bastante diversa cuando se les preguntó si las voces españolas debían imitar la entonación original (véase la pregunta 9 en el cuestionario). La mayoría de los directores (11 de 25) consideran que la entonación original debe ser imitada en el texto objetivo si tiene sentido y no suena ridícula. Hay, sin embargo, otros que piensan que no es una buena idea (5 de 25) ya que cada idioma tiene su propia melodía. Algunas respuestas relacionan la decisión con el tipo de programa y la sugerencia del cliente sobre si seguir la entonación original o no. Las preguntas 7 y 8 se refieren al tipo de locución para diferentes categorías de programas de voces superpuestas y si la voz debe sonar más "actuada" o no, dependiendo del producto que se está grabando. Por lo tanto, para documentales, la respuesta más repetida es que las voces tienen que sonar neutras, seguidas de "similares a las originales", y "actuadas, pero no mucho". Sin embargo, una locución neutra no significa necesariamente que tenga que ser plana y monótona, ya que podría terminar arruinando el programa. Baños (2019: 278) considera que, a diferencia de los programas de tele-realidad hechos mayormente con voces superpuestas, en los documentales no se transmite emoción, "ya que dicha información se puede recopilar de la banda sonora original y se da prioridad a lo que se ha dicho". Esto puede ser cierto para algunos documentales, pero, en nuestra opinión, este no es el caso de los documentales que hemos estudiado, ya que *Jamie's comfort food (La comida reconfortante de Jaimie)* y *Los últimos días de Anne Bolena* (*The last days of Anne Boleyn)* son bastante expresivos, siendo el primero grabado casi como si fuera un *reality*.

Por último, en cuanto a los dialectos y acentos extranjeros, podemos ver que la forma más común de traducción, según los profesionales de nuestra encuesta, es utilizar un acento español neutro (84%) conocido como castellano, mientras que los traductores declaran recurrir tanto a la nivelación como al uso de dialectos españoles, en igualdad de condiciones, para representar estas características regionales.
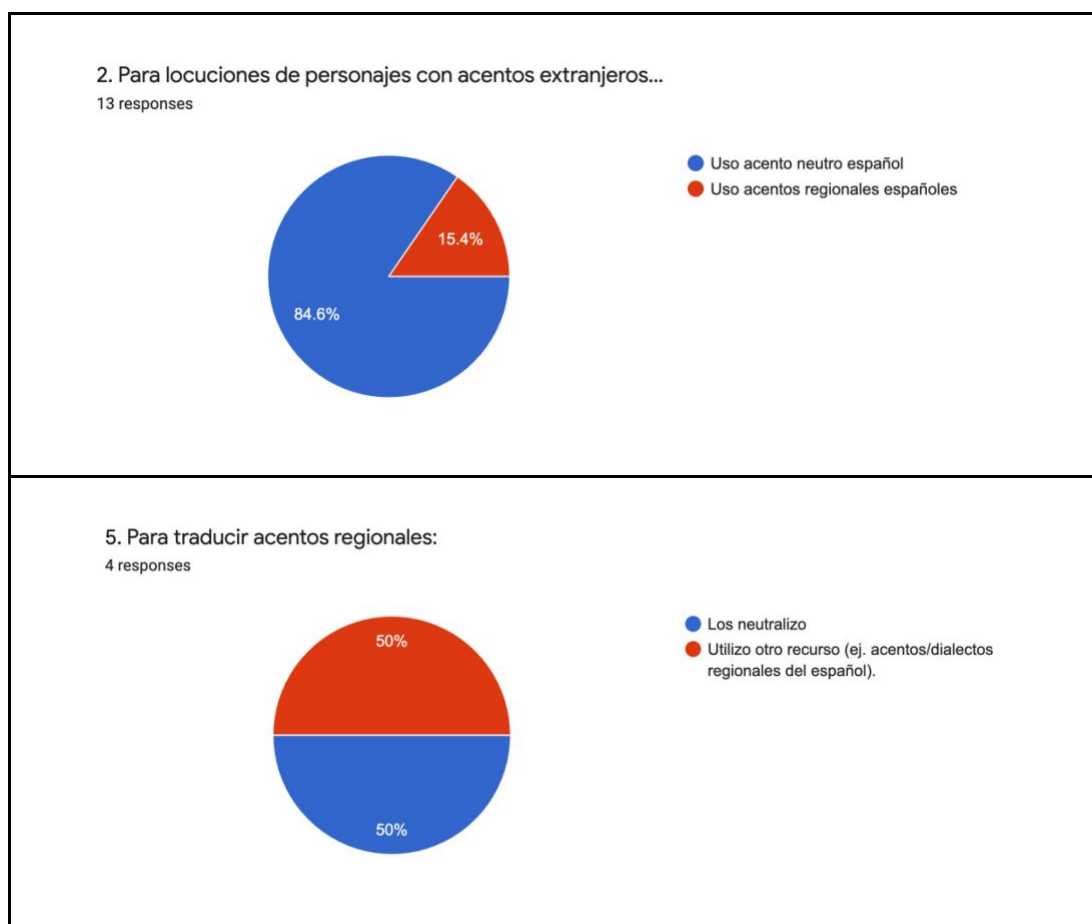
Figura 53. Respuestas a las preguntas 2 (para locutores) y 5 (para traductores)

Durante el análisis de los textos meta no encontramos ningún personaje cuyo acento sonase regional o no estándar, confirmando así lo declarado por la mayoría de los locutores. Pese a que la mitad de los traductores admiten recurrir a dialectos regionales españoles para reproducir acentos en inglés, si el locutor no aplica un acento regional al grabar un guion traducido que contiene expresiones dialectales, el toque regional nunca será percibido por el público meta.

6.2   Cumplimiento de hipótesis

En la sección 4.5 de este trabajo propusimos nuestra hipótesis: ¿son los traductores plenamente conscientes de la relevancia de las cualidades prosódicas del inglés y de los diferentes recursos que ofrece la lengua española para transmitir una fuerza ilocucionaria equivalente en el texto meta?

Tras haber analizado más de 250 ejemplos de expresiones con rasgos prosódicos pragmáticamente relevantes, podemos afirmar que los traductores no son plenamente conscientes de todas las características entonativas y prosódicas que tiene la lengua inglesa para transmitir el valor comunicativo de estas, y las posibilidades que ofrece la lengua española para transmitir un mensaje funcionalmente equivalente. Un ejemplo significativo lo podemos encontrar en el 48% de los casos de foco contrastivo en inglés, cuya fuerza ilocucionaria no se reproduce en el texto meta. Y eso que el foco contrastivo entonativo es, probablemente, uno de los recursos prosódicos más obvios dentro de la entonación inglesa, si lo comparamos con, por ejemplo, la deacentuación o la subordinación entonativa.

Por lo tanto, los resultados ofrecidos en este trabajo ilustran que, a menudo, las voces superpuestas no proporcionan una "buena" traducción, si entendemos como "buena traducción" una que transmita la intención del hablante, y las implicaturas del mensaje original, según lo declarado por Mateo Martínez (1996: 83-5). Evidentemente, esta situación no parece coincidir con la definición estándar y asumida de las voces superpuestas como un modo de representación fiel y auténtico del texto original, ya que la prosodia, como se ha afirmado y demostrado repetidamente en este trabajo, transmite significado pragmático, y este no siempre se representa en el *voice-over* español.

En nuestra opinión, en consonancia con la reflexión de Mateo sobre la relevancia de la enseñanza de la entonación en los programas de estudios de traducción (2014: 131), sensibilizar sobre el componente ilocucionario de los elementos prosódicos en los idiomas de origen y los desafíos que pueden plantear a la hora de encontrar una construcción funcionalmente equivalente (sintáctica, semántica o entonativa) en el idioma meta, mejorará la calidad de las traducciones y, en el caso de

las voces superpuestas, la experiencia audiovisual final. Por lo tanto, en este trabajo, sugerimos que se aplique una nueva característica sincrónica a la traducción de textos para voces superpuestas: *la sincronía prosódico-ilocucionaria*, de modo que no sólo se consideren las características suprasegmentales del lenguaje con significado pragmático, sino también las prosódicas relacionadas con la calidad de la voz. De esta manera, se mejorará la estética de este modo de traducción audiovisual, ya que como afirma Setter (2019: 85): "la voz sí importa, nos guste o no".

6.3   Observaciones finales

En este trabajo hemos destacado aquellos elementos prosódicos de la lengua inglesa que tienen un propósito comunicativo relevante y hemos analizado y comentado cómo se pueden abordar en la traducción para voces superpuestas en español. No encontramos ejemplos para *defining/non-defining relative clauses* (Mott 2011: 223; Wells 2006: 202), pero, en nuestra opinión, se han descrito y analizado los elementos entonativos más relevantes (deacentuación, reacentuación, foco estrecho, dirección de tono, oraciones hendidas, citas, subordinación entonativa, etc.), por lo que esperamos que este análisis proporcione temas de reflexión para traductores audiovisuales, investigadores y profesionales de la industria audiovisual en general.

Como la traducción de textos para voces superpuestas es un proceso colaborativo en el que no sólo participan traductores, sino también locutores, agentes de casting y directores, parece pertinente que la responsabilidad a la hora de lograr un producto final exitoso se comparta entre todos ellos. Por lo tanto, si los guiones se traducen literalmente (casi palabra por palabra), sería una buena idea informar a los actores de que así ha sido, para que su actuación pueda coincidir con el énfasis de la voz de origen pero sin transgredir los límites de la entonación española, como suele ocurrir en la *doblajitis*. Además, si los guiones necesitan ser modificados en el estudio de grabación por motivos isocrónicos, aquellos en la sala responsables de dicha revisión deben tener cuidado de no eliminar nada que pueda alterar la transmisión de la fuerza de ilocucionaria.

Como última observación, y para ofrecer un breve comentario sobre el futuro de la industria audiovisual, tras de más de una década en el sector, tengo una cosa clara: las cosas van a cambiar. Esta industria está en constante evolución, y a medida que la tecnología se desarrolla, los profesionales de la industria tienen que adaptarse a situaciones en constante cambio. Las soluciones audiovisuales colaborativas proporcionadas por la *nube* ya están aquí, y se pueden producir programas doblados y locutados para voces superpuestas (traducción, adaptación, locución y mezcla) desde casa, sin depender de estudios de grabación tradicionales. Además, nuevos programas informáticos impulsados por IA (inteligencia artificial) ya ofrecen soluciones de traducción con sincronización labial por medio de reconocimiento facial y sustitución labial sin alterar en demasía el archivo de vídeo original, con excelentes resultados. Esto significa que, en unas pocas décadas, una vez que dicha tecnología esté disponible para la mayoría de las empresas audiovisuales, la forma en que el doblaje y las voces superpuestas se hacen hoy en día probablemente estará obsoleta, al igual que sucedió con las cabinas telefónicas una vez que los teléfonos móviles irrumpieron en el mercado de las telecomunicaciones. Si esta nueva tecnología tiene éxito, los actuales límites impuestos por las restricciones visuales y temporales se quebrantarán y las características técnicas que caracterizan los textos audiovisuales hoy en día, como la isocronía y la sincronía cinética, dejarán de importar o de ser un obstáculo en el proceso de traducción; lo mismo se aplica a la sincronización fonética o sincronización labial, ya que cada actor de cada película traducida, programa, etc. vocalizará en el idioma meta como si fuese un hablante nativo. Será interesante ver cómo las características prosódicas analizadas en este trabajo se pueden representar en el texto meta utilizando esta nueva tecnología. Pero ese trabajo, por ahora, tendrá que esperar.

7 References

ABERCROMBIE, D. (1967). *Elements of general phonetics.* Edinburgh: Edinburgh University Press.

AGOST, R. (2001). Los géneros de la traducción para el doblaje. In Duro, M. (Ed.) *La Traducción Para El Doblaje Y La Subtitulación,* 229-250.

AIJMER, K. (2002). *English discourse particles: evidence from a corpus.* Amsterdam (NL) – Philadelphia (USA): John Benjamins.

ALARCOS LLORACH, E. (2011 [1994]). *Gramática de la lengua española* (20th edition). Madrid: Colección Nebrija y Bello. Espasa Libros.

ASHBY, P. (2011). *Understanding phonetics*. London: Hodder Education.

ASHBY, P. (2013). *Intonation: Pre-nuclear patterns.* SCEP 2013 Handbook (unpublished manuscript), Summer Course in English Phonetics, UCL, London, 12-13 August 2013.

AUSTIN, J. L. (1962). *How to do things with words*. Oxford: Oxford University Press.

BAGLEY, G. (2001). A mixed bag: Negotiating claims in MTV's the real world. *Journal of Film and Video,* 61-76.

BAKER, M. (2018). *In other words: A coursebook on translation.* Routledge.

BAÑOS, R. (2019). Translating reality TV into spanish: When fast-food TV challenges AVT conventions. *Perspectives, 27*(2), 265-282.

BAÑOS, R., & CHAUME, F. (2009). Prefabricated orality: a challenge in audiovisual translation. *InTRAlinea, Special Issue: The translation of dialects in Multimedia*.

BARZDEVICS, I. A. (2012). Un recorrido por la voz superpuesta. In Martínez Sierra, J. J., (Ed.) *Reflexiones Sobre La Traducción Audiovisual: Tres Espectros, Tres Momentos,* 57-68.

BELL, R. T. (1991). *Translation and translating: Theory and practice.* London: Longman.

BENECKE, B. (2004). Audio-description. *Meta: Journal Des Traducteurs/Meta: Translators' Journal, 49*(1), 78-80.

BERTUCCELLI, M. (1996). *Qué es la pragmática.* Paidós.

BICKERTON, D. (1973). The nature of a creole continuum. *Language, 49*(3), 640-669.

BIRBILI, M. (2000). Translating from one language to another. *Social Research Update, 31*(1), 1-7.

BOSSEAUX, C. (2008). Buffy the vampire slayer: Characterization in the musical episode of the TV series. *The Translator, 14*(2), 343-372.

BRAGA, J. & RICA PEROMINGO, J. P. (2015): *Herramientas y tecnicas para la traduccion ingles-espanol: Los textos literarios*. Escolar y Mayo Editores.

CALDIZ, A. (2012). La prominencia prosódica en español y en inglés: Un analisis contrastivo. Paper presented at the *IV Jornadas Internas De Español Como Lengua Segunda Y Extranjera: Experiencias, Desarrollos, Propuestas.* Universidad Nacional de la Plata. Buenos Aires.

CATTRYSSE, P. (2001). Multimedia and translation: Methodological considerations. In Gambier, Y. (Ed.) *(Multi) media translation: Concepts, practices and research,* 1-12.

CEREZO MERCHÁN, B. (2012). *La didáctica de la traducción audiovisual en españa: Un estudio de caso empírico-descriptivo.* (Unpublished PhD Dissertation), Universitat Jaume I, Castellón de la Plana.

CHAUME, F. (2001). La pretendida oralidad de los textos audiovisuales y sus implicaciones en traducción. In Chaume, F. & Agost, R. (Eds.) *La Traducción En Los Medios Audiovisuales, 7*, 77-87. Universitat Jaume I.

CHAUME, F. (2004a). *Cine y traducción.* Catedra Ediciones.

CHAUME, F. (2004b). Discourse markers in audiovisual translating. *Meta: Journal Des Traducteurs/Meta: Translators' Journal, 49*(4), 843-855.

CHAUME, F. (2005a). Estrategias y técnicas de traducción para el ajuste o adaptación en el doblaje. In Merino Álvarez, R. Santamaría, J. M., Pajares, E. (Eds.) *Trasvases culturales: literatura, cine, traducción* (*4*): 145-153.

CHAUME, F. (2005b). Los estándares de calidad y la recepción de la traducción audiovisual. *Puentes*, *6*, 5-12.

CHAUME, F. (2008). Teaching synchronisation in a dubbing course: some didactic proposals. *The didactics of audiovisual translation*, 129-140. John Benjamins.

CHAUME, F. (2012). *Audiovisual translation: Dubbing.* St. Jerome Pub.

CHIARO, D. (2009). Issues in audiovisual translation. *The routledge companion to translation studies*, 155-179. Routledge.

CID URIBE, M. E. (1996). La entonación de las preguntas ratificadas en español y de sus equivalentes" question tags" en inglés: Una comparación. *Onomázein: Revista*

De Lingüística, Filología Y Traducción De La Pontificia Universidad Católica De Chile, (1), 56-70.

CLARK, H. H. & GERRIG, R. J. (1990). Quotations as demonstrations. *Language,* 764-805. Linguistic Society of America.

CLEVELAND, T. F. (1977). Acoustic properties of voice timbre types and their influence on voice classification. *The Journal of the Acoustical Society of America, 61*(6), 1622-1629.

COLLINS, B., & MEES, I. M. (2013). *Practical phonetics and phonology: A resource book for students.* Routledge.

CÓMITRE NARVÁEZ, I. (1997). Algunas consideraciones sobre la traducción del texto audiovisual. In Santamaría, J. M., Pajares, E.; Olsen, V., Merino, R., Eguíluz, F. (Eds.) *Trasvases culturales: Literatura, cine, traducción (2),* 89-95. Universidad del País Vasco.

COUPER-KUHLEN, E. (1986). *An introduction to English prosody.* Edward Arnold.

CRUTTENDEN, A. (1997). *Intonation.* Cambridge University Press.

CRUTTENDEN, A. (2008). *Gimson's Pronunciation of English* (7th ed.) London: Hodder Education.

CRYSTAL, D. (1975). *The English tone of voice: Essays in intonation, prosody and paralanguage.* Hodder Arnold.

DARWISH, A. & ORERO, P. (2014). Rhetorical dissonance of unsynchronized voices: Issues of voice-over in news broadcasts. *Babel, 60*(2), 129-144.

DE BEAUGRANDE, R., & DRESSLER, W. U. (1981). *Introduction to text linguistics.* Routledge.

DEL RÍO, R. (2001). Problemática del subtitulado y el doblaje en la traducción cinematográfica. In Pascua Febles, I. (Ed.) *La Traducción: Estrategias Profesionales,* 219-228.

DEL VALLE, MARI SOL. (2015a). Anomalías en radio y televisión: ¿A quién le importan? *Cuadernos De Periodistas: Revista De La Asociación De La Prensa De Madrid,* (*31*), 104-119.

DEL VALLE, MARI SOL. (2015b). *La lectura de noticias de los informativos en la radio y la televisión: Análisis de la prosodia y la emisión de la voz.* (Unpublished PhD Dissertation), Universidad Complutense de Madrid. Madrid.

DEPRAETERE, I. & SALKIE, R. (2017). *Semantics and pragmatics: Drawing a line.* Springer.

DEWAELE, J. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development, 25*(2-3), 204-222.

DÍAZ CINTAS, J. & MUÑOZ SÁNCHEZ, P. (2006). Fansubs: Audiovisual translation in an amateur environment. *Jostrans: The Journal of Specialised Translation, 6,* 37-52.

DÍAZ CINTAS, J. & ORERO, P. (2006). Voice-over. *Encyclopedia of language & linguistics*, 477-479.

DÍAZ CINTAS, J. & ORERO, P. (2010). Voiceover and dubbing. In Gambier, Y. & Van Doorslaer L. (Eds.) *Handbook of Translation Studies, vol 1*, 441-445. John Benjamins

DÍAZ CINTAS, J. & REMAEL, A. (2014 [2007]). *Audiovisual translation: Subtitling*. Routledge.

ESPASA, E. (2002). Traducir documentales: Paisajes híbridos. In Sanderson Pastor, John D. (Ed.) *Traductores Para Todo: Actas De Las III Jornadas De Doblaje Y Subtitulación De La Universidad De Alicante,* 29-40.

ESPASA, E. (2004). Myths about documentary translation. In Orero, P. (Ed.) *Topics in audiovisual translation* (Vol. 56). John Benjamins Publishing. 183-198.

ESTEBAS-VILAPLANA, E. (2014a). The evaluation of intonation. In Thompson, G. & Alba-Juez, L. (Eds.) *Evaluation in Context,* John Benjamins 179-194.

ESTEBAS-VILAPLANA, E. (2014b). Phonological models of intonational description of English. In Monroy-Casas, R. and Arboleda-Girao, I. (Eds.) *Readings in English phonetics and phonology*, 231-260. IULMA (Institut Universitari de Llengües Modernes Aplicades) Universitat de Valencia.

ESTEBAS-VILAPLANA, E. (2014c). *Teach yourself English pronunciation: An interactive course for Spanish speakers.* Netbiblo.

FATTAL, I. (2018). Why do cartoon villains speak in foreign accents. *The Atlantic,* available on: https://www.theatlantic.com/education/archive/2018/01/why-do-cartoon-villains-speak-in-foreign-accents/549527/

FAWCETT, P. (1983). Translation modes and constraints. *The Incorporated Linguist, 22*(4), 186-190.

FODOR, I. (1976). *Film dubbing: phonetic, semiotic, esthetic and psychological aspects.* 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, 13*(2), 289-304.

FRANCO, E., MATAMALA, A., & ORERO, P. (2010). *Voice-over translation: An overview.* Peter Lang.

FUCHS, R. (2014). You got the beat: Rhythm and timing. In Monroy-Casas, R. & Arboleda-Girao, I. (Eds.) *Readings in English phonetics and phonology,* 165-188. IULMA (Institut Universitari de Llengües Modernes Aplicades) Universitat de Valencia.

GAMBIER, Y. (1996). *Les transferts linguistiques dans les médias audiovisuels* Presses Univ. Septentrion.

GAMBIER, Y. (2006). Subtitling. *Encyclopedia of Language and Linguistics,* 258-263.

GAMBIER, Y. & SUOMELA-SALMI, E. (1994). Subtitling: A type of transfer. In Eguiluz, Merino, Olsen, and Pajares (Eds.) *Trasvases culturales: literatura, cine, traducción*, (1), 243-252. Universidad del País Vasco.

GAMBIER, Y., & GOTTLIEB, H. (2001). *(Multi) media translation: Concepts, practices, and research.* John Benjamins Publishing.

GARCÍA-LECUMBERRI, M. L. (1995). *Intonational Signalling of Information Structure in English and Spanish: A Comparative Study.* (Unpublished PhD Dissertation), University of London.

GARRIDO, C. (2013). La traducción, hermana pobre de los documentales de la naturaleza en españa. registro y análisis de las inadecuaciones traductivas presentes en la versión castellana de life, de la BBC. *Quaderns: Revista De Traducció, 20,* 211-233. Departament de Traducció i d'Interpretació de la Universitat Autònoma de Barcelona

GILLIES, A. (2017). *Note-taking for consecutive interpreting: A short course.* Routledge.

GIVENS, D. B. (2016). Reading palm-up signs: Neurosemiotic overview of a common hand gesture. *Semiotica, (210),* 235-250.

GÓMEZ GONZÁLEZ, M., & SÁNCHEZ T. (2016). *English pronunciation for speakers of Spanish: From theory to practice* .Walter de Gruyter GmbH & Co KG.

GORBMAN, C. (2014). The master's voice. *Film Quarterly, 68* (2), 8-21.

GÓRSKA, K. (2015). Ni doblaje ni subtitulación. el gran éxito del voice-over en la televisión polaca. In *Quaderns: Revista De Traducció*, *10*, 63-72.

GRANATO, L. (2014). Studies on the intonation of English: A critical review. *Readings in English Phonetics and Phonology, 1,* 189-230.

GRICE, H. P. (1975). Logic and conversation. In Cole, P. & Morgan J. L. (Eds.) *Syntax and Semantics: Speech acts,* vol 3, 41-58. Elsevier.

GUTIÉRREZ ORDÓÑEZ, S. (1997). *Temas, remas, focos, tópicos y comentarios.* Arco Libros.

GUTT, E. (1990). A theoretical account of translation-without a translation theory. *Target: International Journal of Translation Studies, 2*(2), 135-164. John Benjamins.

HALLIDAY, M. & HASAN, R. (1985): *Language, context, and text: Aspects of language in a social-semiotic perspective.* Waurn Ponds, Vic: Deakin University: distributed by Deakin University Press.

HATIM, B. & MASON, I. (1990). *Discourse and the translator.* Longman.

HATIM, B. A. & MUNDAY, J. (2004). *Translation: An advanced resource book.* Routledge.

HERVEY, S. G. (1998). Speech acts and illocutionary function in translation methodology. In Hickey, L. (Ed.) *The Pragmatics of Translation, Topics in translation 12,* 10-24. Multilingual Matters Ltd.

HEWINGS, M. (2007). *English pronunciation in use. Advanced book with answers, 5 audio CDs and CD-ROM.* Cambridge University Press.

HICKEY, L. (1993). Stylistics, pragmatics and pragmastylistics. *Revue Belge de Philologie et d'Histoire, 71*(3), 573-586.

HICKEY, L. (1998). Perlocutionary equivalence: Marking, exegesis and recontextualisation. In Hickey, L. *The Pragmatics of Translation, Topics in Translation 12*, 217. Multilingual Matters LTD.

HIDALGO NAVARRO, A. (2001). Modalidad oracional y entonación. notas sobre el funcionamiento pragmático de los rasgos suprasegmentales en la conversación. In *Moenia: Revista lucense de lingüística y literatura*, 7, 271-292.

HORAK, J. (2006). Wildlife documentaries: From classical forms to reality TV. *Film History: An International Journal, 18*(4), 459-475.

HOUSE, J. (2001). Pragmatics and translation. In Fernández Nistal, P. & Bravo Gozalo, J. M. (Eds.) 2001 *Pathways of Translation Studies*. Valladolid: Universidad de Valladolid. 55-77.

HOUSE, J. & BLUM-KULKA, S. (1986). *Interlingual and intercultural communication: Discourse and cognition in translation and second language acquisition studies.* Gunter Narr Verlag Tübingen.

HURTADO ALBIR, A. (2001). *Traducción Y Traductología. Introducción a La Traductología.* Cátedra.

JIMÉNEZ, M. A. (1999). *La traducción a la vista. un análisis descriptivo.* (Unpublished PhD Dissertation). Universidad Jaume I. Castellón de la Plana.

KAUFMANN, F. (2004). Un exemple d'effet pervers de l'uniformisation linguistique dans la traduction d'un documentaire: De l'hébreu des immigrants de «Saint-jean» au français normatif d'ARTE. *Meta: Journal Des Traducteurs/Meta: Translators' Journal, 49*(1), 148-160.

KELLY, G. (2000). *How to Teach pronunciation.* Ed. Jeremy Harmer. Pearson Education Limited.

KOLLER, W. (1989). Equivalence in translation theory. *Readings in Translation Theory,* 99-104.

KOZLOFF, S. (1988). *Invisible storytellers: Voice-over narration. American fiction film*. University of California Press.

LADD, D. R. (2015). The American four-level analysis of intonation contours: Historical postscript. *Historiographia Linguistica, 42*(1), 119-137. John Benjamins.

LANDERS, C. E. (2001). *Literary translation: A practical guide.* Multilingual Matters.

LARSEN, I. (2001). The fourth criterion of a good translation. *Notes on Translation-Summer Institute of Linguistics, 15*(1), 40-53.

LATINUS, M. & BELIN, P. (2011). Human voice perception. *Current Biology, 21*(4), R143-R145.

LEECH, G., & SVARTVIK, J. (1994). *A communicative grammar of English* (2nd ed.). London: Longman Group Limited.

LEVY, J. (1967). Translation as a decision process. In J. Levy (Ed.) *To honor Roman Jakobson: Essays on the occasion of his seventieth birthday, 1171-1182.* Mouton.

LINDSEY, G. (1981). Intonation and pragmatics. *Journal of the International Phonetic Association, 11*(1), 2-21. Cambridge University Press.

LLORENTE PINTO, M. D. R. (2006). ¿Qué es el español neutro?. In *DLE. Artículos del Departamento de Lengua Española, Cuadernos del Lazarillo, 31*. GREDOS (Gestión del Repositorio Documental de la Univerisdad de Salamanca).

LÓPEZ GUIX, J. G. & WILKINSON, J. M. (2001). *Manual De Traducción: Inglés-Castellano: Teoría Y Práctica.* Barcelona. Gedisa Editorial.

LUNDY, L. K., RUTH, A. M. & PARK, T. D. (2008). Simply irresistible: Reality TV consuåmption patterns. *Communication Quarterly, 56*(2), 208-225.

LUQUE, A. F. (2017). Aspectos profesionales y técnicos de la traducción audiovisual, con especial referencia al caso de españa. *TRANS. Revista De Traductología, (5),* 143-152.

LUYKEN, G. & HERBST, T. (1991). *Overcoming language barriers in television: Dubbing and subtitling for the european audience.* Manchester. European Institute for the Media.

MARICIC, S. & DURIC, M. (2011). "Pero", "sin embargo" y "no obstante" como conectores discursivos. *Facta Universitatis Series: Linguistics and Literature, 9*(2), 89-98.

MARTÍ FERRIOL, J. L. (2006). *Estudio empírico y descriptivo del método de traducción para doblaje y subtitulación.* (Unpublished PhD Dissertation). Universitat Jaume I. Castellón de la Plana.

MARTÍ FERRIOL, J. L. (2013). *El método de traducción. doblaje y subtitulación frente a frente.* Publicacions de la Universitat Jaume I.

MARTÍNEZ CARO, E. (2002). *Funciones pragmáticas, orden de constituyentes y acentuación en inglés y en español: Estudio del corpus.* Universidad Complutense de Madrid, Servicio de Publicaciones.

MARTÍNEZ-CELDRÁN, E., FERNÁNDEZ-PLANAS, A. M. & CARRERA-SABATÉ, J. (2003). Castilian Spanish. *Journal of the International Phonetic Association, 33*(2), 255-259. Oxford University Press.

MARTOS, A. (2016). Lo que siempre deberías hacer cuando hables por teléfono. *20 Minutos.* Retrieved from https://blogs.20minutos.es/comunicacion-no-verbal-lo-que-no-nos-cuentan/tag/sonrisa-telefonica/

MATAMALA, A. (2008). Teaching voice-over translation: A practical approach. In Díaz Cintas, J. (Ed.) *The Didactics of Audiovisual Translation,* 115-127. John Benjamins.

MATAMALA, A. (2009a). Main challenges in the translation of documentaries. In Díaz Cintas, J. (Ed.) *New Trends in Audiovisual Translation,* 109-120. Multilingual Matters.

MATAMALA, A. (2009b). Translating documentaries: From neanderthals to the supernanny. *Perspectives: Studies in Translatology, 17*(2), 93-107.

MATAMALA, A. (2019). Voice-over: Practice, research and future prospects. In Pérez-González, Luis (Ed.) *The Routledge Handbook of Audiovisual Translation,* 64-81. Routledge.

MATÉ, M. G. (2017). El llamado "español latino" de los doblajes cinematográficos en la encrucijada entre el español mexicano, el "español general" y el "español neutro". In Silke, J. & Gesine, M. La *Traducción Desde, En y Hacia Latinoamérica: Perspectivas Literarias y Lingüísticas*. Iberoamericana/Vervuert, 247–274.

MATEI, M. (2010). Discourse markers as functional elements. *Bulletin of the Transilvania University of Braşov, 3*, 52.

MATEO MARTÍNEZ, J. (1996). La fuerza ilocucionaria y su relevancia en la traducción del inglés al español. *Pragmalingüística* (3-4), 77-88.

MATEO, M. (2002). Los sobretítulos de ópera: Dimensión técnica, textual, social e ideológica. Paper presented at the *Traductores Para Todo: Actas De Las III Jornadas De Doblaje Y Subtitulación De La Universidad De Alicante,* 51-74.

MATEO, M. (2007). Surtitling today: New uses, attitudes and developments. *Linguistica Antverpiensia, New Series–Themes in Translation Studies,* (6)

MATEO, M. (2014). Exploring pragmatics and phonetics for successful translation. (*VIAL) Vigo International Journal of Applied Linguistics, (11)*, 111-135.

MAYORAL, R. (2001). El espectador y la traducción audiovisual. In Chaume, F. and Agost, R. (Eds.) *La traducción en los medios audiovisuales. Castellón: Universitat Jaume I, Servei de Publicacions:* 33-48

MAYORAL, R. (2012). El estudio de la traducción audiovisual. In Martínez Sierra, J. J., (Ed.) *Reflexiones Sobre La Traducción Audiovisual: Tres Espectros, Tres Momentos,* 179-185.

MAYORAL, R., KELLY, D. & GALLARDO, N. (1988). Concept of constrained translation. non-linguistic perspectives of translation. *Meta: Journal Des Traducteurs/Meta: Translators' Journal, 33*(3), 356-367.

MEAD, J. A. (2005). *Fascination of reality television with the college student audience: The uses and gratifications perspective on the program genre.* Paper presented at the Association for Education in Journalism and Mass Communication Conference, San Antonio.

MOTT, B. (1991). *A course in phonetics and phonology for Spanish learners of English.* Barcelona: PPU.

MOTT, B. (2009). Fronted nuclei: Narrow focus in spanish and english. *Bucharest Working Papers in Linguistics, (2),* 73-81.

MOTT, B. (2011 [2000]). *English phonetics and phonology for Spanish speakers* (2nd ed.). Barcelona: Edicions Universitat Barcelona.

MULLER, S. (2005). *Discourse markers in native and non-native English discourse.* Amsterdam (NL) – Philadelphia (USA): John Benjamins.

MÜLLER, S. (2005). *Discourse markers in native and non-native English discourse.* John Benjamins Publishing.

NABI, R. L., BIELY, E. N., MORGAN, S. J. & STITT, C. R. (2003). Reality-based television programming and the psychology of its appeal. *Media Psychology, 5*(4), 303-330.

NICHOLS, B. (2017). *Introduction to documentary.* Indiana University Press.

NIDA, E. A. & TABER, C. R. (1969). *The theory and practice of translation.* Brill.

NISTAL, P. F. & GOZALO, J. M. B. (2001). *Pathways of translation studies.* Universidad de Valladolid, Centro Buendía de Cultura, Formación Continua y Extensión Universitaria.

NOLAN, F. (1991). Forensic phonetics. *Journal of Linguistics, 27*(2), 483-493.

O'CONNOR, J. D. & ARNOLD, G. F. (1973). *Intonation of colloquial English: A practical handbook.* Longman, London.

O'DONNELL, R. M. (1990). Research needs and gaps: Hispanic Americans. In Haper, M. S. (Ed.), *Minority aging: Essential curricula content for selected health and*

allied health professionals. 425-432. Washington, DC: U.S. Government Printing Office.

ORERO, P. (2004). The pretended easiness of voice-over translation of TV interviews. *JoSTrans, the Journal of Specialised Translation, 2*(8), 76-96.

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, 213-222.* Granada: Comares.

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, 255-264.

ORERO, P. (2006b). Voice-over: A case of hyper-reality. 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. 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. *Audiovisual translation*, 130-139. Springer.

PAGEON, D. (2007). *The world of the voice-over: Writing, adapting and translating scripts; training the voice; building a studio.* Actors World Production.

PAQUIN, R. (1998). Translator, adapter, screenwriter. Translating for the audiovisual. *Translation Journal, 2*(3), 1-5.

PARROTT, M. (2010 [2000]). *Grammar for English language teachers* (2nd ed.) Cambridge University Press.

PÉREZ-GONZÁLEZ, L. (2014). *Audiovisual translation: theories, methods and issues.* Routledge.

PERTEGHELLA, M. (2002). Language and politics on stage: Strategies for translating dialect and slang with references to shaw's pygmalion and bond's saved. *Translation Review, 64*(1), 45-53.

PETTIT, Z. (2005). Translating register, style and tone in dubbing and subtitling. *JoSTrans Journal of Specialised Translation, 4*(4), 49-65.

PIASTRA, L. (1989). La traducción cinematográfica. Paper presented at the *Fidus Interpres: Actas De Las Primeras Jornadas Nacionales De Historia De La Traducción,* 344-352.

PIKE, K. L. (1945). *The Intonation of American English.* Ann Arbor: University of Michigan Press.

PINTO, M. (2001). Quality factors in documentary translation. *Meta: Journal Des Traducteurs/Meta: Translators' Journal, 46*(2), 288-300.

PLANTINGA, C. (2005). What a documentary is, after all. *The Journal of Aesthetics and Art Criticism, 63*(2), 105-117.

PYM, A. (1997). Koller's äquivalenz revisited. *The Translator, 3*(1), 71-79.

QUILIS, A. (1997). *Principios de fonología y fonética españolas.* Arco libros.

RABADÁN, R. (1991). *Equivalencia y traducción: Problemática de la equivalencia translémica inglés-español.* León: Universidad, 1991.

RABADÁN, R., & FERNÁNDEZ NISTAL, P. (2002). *La traducción inglés-español: Fundamentos, herramientas, aplicaciones.* Leon: Universidad de León.

RAE (2011) *Nueva gramática de la lengua española. fonética y fonología.* In Blecua, J. M. (Ed.) Real Academia de la Lengua Española. Madrid: Colección Nebrija y Bello. Espasa Libros.

RAE (Real Academia Espanola). (2005). Diccionario panhispánico de dudas - uso de la coma. Retrieved from: http://buscon.rae.es/dpd/srv/search?id=SwqUf6YOtD6TfcfDSC%7CV1EqcYbX4D61AWBBrd [last accessed on 20/02/2020]

REESE, B. & ASHER, N. (2007). Prosody and the interpretation of tag questions. Paper presented at the *Proceedings of Sinn Und Bedeutung, 11,* 448-462.

REMAEL, A. (2001). Some thoughts on the study of multimodal and multimedia translation. In Gambier, Y. and Gottlieb, H. (Eds.) *(Multi) media translation: Concepts, practices, and research.* Amsterdam, 13-22.

REYES, G. (1995). *El abecé de la pragmática.* Arco libros.

REYNOLDS, M. & SIZEMORE, M. (2000). Spanish voice-over: An effective option for producing educational videotapes for the hispanic elderly. *Educational Gerontology, 26*(5), 467-474.

RICA PEROMINGO, J. P. (2014). La traducción de marcadores discursivos (DM) inglés-español en los subtítulos de películas: Un estudio de corpus. *JoSTrans, the Journal of Specialised Translation, 21*, 177-199.

RICA PEROMINGO, J. P. (2016). *Aspectos lingüísticos y técnicos de la traducción audiovisual (TAV),* Bern, Switzerland: Peter Lang CH.

ROACH, P. (2009 [1983]). *English phonetics and phonology: A practical course.* (4th edition). Cambridge University Press.

ROMERO RAMOS, L. (2005). La traducción de dialectos geográficos y sociales en la subtitulación: Mecanismos de compensación y tendencia a la estandarización. In Merino Álvarez, R., Santamaría, J. M., and Pajares, E. (Eds.) *Trasvases culturales: literatura, cine, traducción* (4), 243-259. Universidad del País Vasco.

ROTURIER, J. (2015). *Localizing apps: A practical guide for translators and translation students.* Routledge.

SÁNCHEZ MOMPEÁN, S. (2012). Dubbing Animation into Spanish: Behind the Voices of Animated Characters, *JoSTrans, the Journal of Specialised Translation, 23*, 270-291.

SÁNCHEZ MOMPEÁN, S. (2017). *The rendition of English intonation in Spanish dubbing.* (Unpublished PhD Dissertation), Universidad de Murcia, Murcia.

SÁNCHEZ MOMPEÁN, S. (2020). *The Prosody of Dubbed Speech. Beyond the Character's words.* Palgrave Studies in Translating and Interpreting. Palgrave Macmillan.

SCEP Handbook (2013). (Unpublished manuscript), Summer Course in English Phonetics, UCL, London, 12-13 August 2013.

SCHWARZ, B. (2011). Translation for dubbing and voice-over. In Malmjaer, K. and Windle, K. (Eds.) *The Oxford handbook of translation studies*. 394-409. Oxford University Press.

SCOTT, K. (2013). *Functions of intonation.* SCEP 2013 Handbook (unpublished manuscript), Summer Course in English Phonetics, UCL, London, 12-13 August 2013.

SCRIVENER, J. (2011 [1994]). *Learning teaching: The essential guide to English language teaching.* (3rd edition) Macmillan Education.

SEARLE, J. R. (1969). *Speech acts: An essay in the philosophy of language.* Cambridge University Press.

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.

SEPIELAK, K. (2016). Synchronization techniques in multilingual fiction: Voiced-over films in poland. In Gross, L. (Ed.) *International Journal of Communication,* 10, 1054-1073.

SETTER, J. (2019). *Your voice speaks volumes: It's not what you say, but how you say it.* Oxford. Oxford University Press.

SHLESINGER, M. (2010). Relay interpreting. In Gambier, Y. and Van Doorslaer, L. (Eds.) *Handbook of Translation Studies,* 1, 276-278.

SILEO, A. (2018). Il simil sync o semi-sinc nel panorama italiano: lo slittamento e offuscamento dei confini previsti dal CCNL. *Altre Modernità*, 258-267.

SINCLAIR, J. (1996). Preliminary recommendations on corpus typology. *EAGLES Document TCWG-CTYP/P* (Available from Http://Www.Ilc.Pi.Cnr.it/EAGLES/Corpustyp/Corpustyp.Html) [last accessed on 20/02/2020]

SOH TATCHA, C. (2009). Doublage cinématographique et audiovisuel: Équivalence de son, équivalence de sens. *Meta: Journal Des Traducteurs/Meta: Translators' Journal, 54*(3), 503-519.

SOKOLI, S. (2005). Temas de investigación en traducción audiovisual: La definición del texto audiovisual. In Zabalbeascos, P., Santamaría, L. and Chaume, F. (Eds.) *La Traducción Audiovisual: Investigación, Enseñanza Y Profesión.* Granada: Comares, 177-185.

SOLÉ, M. J. (1989). Translating intonation. Paper presented at the *Actas XI Congreso AEDEAN: Translation Across Cultures: La Traducción Entre El Mundo Hispánico Y Anglosajón, Relaciones Lingüísticas, Culturales Y Literarias,* 181-194.

SOLÉ, M. J. (1991). Stress and rhythm in English. *Revista Alicantina De Estudios Ingleses,* no.04, 145-162,

SPERBER, D. & WILSON, D. (1986). *Relevance: Communication and Cognition.* Blackwell.

STEWART, W. A. (1964). Urban negro speech: Sociolinguistic factors affecting English teaching. In R. Shuy, (Ed.) *Social Dialects and Language Learning: Proceedings of the Bloomington, Indiana, Conference.* National Council of Teachers of English.

SWAN, M. (2005). *Practical English usage* Oxford University Press.

TELLO, I. (2011). *La traducción del dialecto: Análisis descriptivo del dialecto geográfico y social en un corpus de novelas en lengua inglesa y su traducción al*

español. (Unpublished PhD Dissertation). Universitat Jaume I. Castellón de la Plana.

TENCH, P. (2015). The intonation systems of English. Bloomsbury Publishing.

TENCH, P. *The pronunciation of grammar.* Retrieved from: http://digitalmarketingcardiff.com/unclepaul/wp-content/uploads/2013/08/pronunciation_of_grammar.pdf [last accessed on 20/02/2020]

TOURY, G. (1980). *In search of a theory of translation.* Porter Institute for Poetics and Semiotics, Tel Aviv University.

TOVAR, A. (2014). *Estrategias de focalización de la gramática del español.* (Unpublished BA Dissertation). Universidad Autónoma de Barcelona.

TRAGER, G. L. & SMITH, H. L. (1951). *An outline of English structure.* Norman: Battenburg Press.

UR, P. (1996). *A course in language teaching trainer's handbook: Practice of theory* Ernst Klett Sprachen.

VARSHNEY, U., SNOW, A., MCGIVERN, M. & HOWARD, C. (2002). Voice over IP. *Communications of the ACM, 45*(1), 89-96.

VERHOEVEN, J. (2002). The communicative setting and markers in speech. *Phonetic work in progress*, 177, 190. Antwerp: Antwerp University.

VIDAL, M. V. E. & VICTORIA, M. (1999). Los enunciados interrogativos: Aspectos semánticos y pragmáticos. In Bosque, I., & Demonte, V. (Eds.) *Gramática*

Descriptiva De La Lengua Española: Entre La Oración Y El Discurso. Morfología, 3, 3929-3992.

VINCE, M. & EMMERSON, P. (2009 [1993]). *First certificate language practice* (4th edition). Oxford: Macmillan Education.

WANG, L. (2007). Theme and rheme in the thematic organization of text: Implications for teaching academic writing. *Asian EFL Journal, 9*(1), 164-176.

WELLS, J. C. (1982). *Accents of English: An Introduction.* Cambridge University Press.

WELLS, J. C. (2006). *English intonation: An introduction.* Cambridge University Press.

WELLS, R. S. (1945) The pitch phonemes of English. *Language, 21,* 27-39.

WHARTON, T. (2012) Prosody and meaning: theory and practice. In: Romero-Trillo, J., (Ed.) *Pragmatics, Prosody and English Language Teaching.* Dordrecht, Netherlands : Springer. pp. 97-116.

WIGMORE, J. H. (1926). A new mode of identifying criminals. *Journal of Criminal Law and Criminology, 17 (2),* 165-166.

WRIGHT, J. A. & LALLO, M. J. (2013 [2009]). *Voice-over for animation.* Focal Press.

YANG, L., LIU, J., YANG, X. & HUA, X. (2007). Multi-modality web video categorization. Paper presented at the *Proceedings of the International Workshop on Workshop on Multimedia Information Retrieval,* 265-274.

YULE, G. (1996). *Pragmatics.* Oxford University Press.

ZABALBEASCOA, P. (2001). La traducción de textos audiovisuales y la investigación traductológica. In Chaume, F. and Agost, R. (Eds.) *La Traducción En Los Medios Audio-Visuales,* 49-56. Publicacions de la Universitat Jaume I.

ZABALBEASCOA, P. (2008). The nature of the audiovisual text and its parameters. In Díaz Cintas, J. (Ed.) *The Didactics of Audiovisual Translation, 7*, 21-37. John Benjamins.

ZABALBEASCOA, P., IZARD, N. & SANTAMARIA, L. (2001). Disentangling audiovisual translation into Catalan from the Spanish media mesh. In Zabalbeascoa, P., Izard, N. & Santamaría, L. (*Multi) Media Translation: Concepts, practices and research, 34*, 101-112. John Benjamins.

8 Annexes

The folder with the seven annexes for this dissertation can be downloaded from the following web link: https://we.tl/t-Bg3fzOcThP