Exploring Applied Strategies
for English-Language Dubbing

 Giselle Spiteri Miggiani⊠
University of Malta

_____

Abstract

The sudden boom of non-English language content on popular streaming platforms has obliged media localisation companies to react and adapt quickly to provide the market with English dubbing lip-synch services. This comes with several challenges due to the absence of a long-established English dubbing tradition and professional practice, as well as the lack of consolidated norms and conventions, or a textual repertoire to act as a point of reference. This paper builds on a prior theoretical study (Spiteri Miggiani, 2021) that seeks to identify potential norms as well as challenges in English-language dubbing. It goes a step further by exploring strategies and techniques that could address the identified challenges while aiming to satisfy the generally accepted quality standards that govern dubbing globally. The proposed strategies apply to the dubbing workflow as a whole with a special focus on the dubbing text adaptation process. The study aims to set the grounds for further research required to corroborate these strategies through applied studies in an academic setting, or through collaboration with localisation companies where they can be tested and observed in professional practice.

Key words: English-language dubbing, dubbing, adaptation, training, applied strategies, quality standards, lip-synch, rhythm, intonation, sound.

⊠ mspit12@um.edu.mt, https://orcid.org/0000-0002-6382-9938

1. Introduction

New ways of consuming audiovisual content have given rise to new trends, audiences, and consequently new localisation demands (Jenner, 2019). The rapidly growing increase in availability of non-English language productions on over-the-top (OTT) platforms, such as Netflix, Amazon Prime Video and Hulu, has given birth to a new media localisation sub-industry, that of English-language dubbing.

One of the reasons for the increase in non-English language content is that a number of streaming platforms need to comply with an EU requirement whereby 30 percent of distributed content in Europe has to be European, as stated in **Article 13** of the *Audiovisual Media Services Directives* issued in November 2018 in the *Official Journal of the European Union*. This directive implies producing, acquiring or funding European content.

Original productions in Spanish, Italian, Swedish, Danish, German, French and so on, come with English language subtitled and dubbed streams in English-speaking countries or other territories where English is largely understood and target language versions may not be provided. The need to provide a dubbed stream may be due to the fact that not all Anglophone viewers are necessarily accustomed to subtitles, considering the substantial amount of English language original content that has been available so far. A consumer research study carried out by Netflix seems to demonstrate that US viewers who watch the English dubbed streams of foreign language series are more likely to continue watching till the end. Netflix also claims that deliberately defaulting to the English dubbed dialogue has actually led to increased viewership in the US (Roettgers, 2018). Despite this, viewer response doesn't always seem so favourable; English dubbing has been described by consumers as "dubby" and "awkward," to the point that Netflix has been trying to address the issue by seeking strategies that may enhance the quality, and consequently, the credibility of its English-dubbed streams (Goldsmith, 2019).

The sudden boom of foreign language content has obliged localisation companies to quickly adapt to provide the market with English dubbing lip-synch services. This comes with several challenges due to the absence of a long-established English dubbing tradition and modus operandi, as well as the lack of consolidated norms and conventions, or a textual repertoire to act as a point of reference.

A prior theoretical study intended at identifying potential norms as well as challenges in English dubbing explores whether English dubbed patterns are modelled on the standards that govern other dubbing languages and cultures (Spiteri Miggiani, 2021). The study provides an empirical qualitative analysis of a selection of English-dubbed fiction series of various source languages streamed on Netflix. The sample data encompasses the pilot episodes of: *The Hook Up Plan* (French, *Plan Coeur*) (Couvreur et al., 2018–present); *Money Heist* (Spanish, *La Casa de Papel*) (Pina at al., 2017–present); *Better Than Us* (Russian, *Luchshe, chem lyudi*) (Dzhunkovskiy et al., 2018–2019); *How to Sell Drugs Online (Fast)* (German) (Kässbohrer et al., 2019–present); *Fauda* (Hebrew/Arabic) (Benasuly Amit et al., 2015–present). These are examined *in primis* as target culture audiovisual texts in their own right

and only later in comparison with their original counterparts. The identified patterns are analysed vis-à-vis a revisited classification of quality parameters (outlined later on), adapted further from the commonly accepted quality standards proposed and discussed by Chaume (2007), among other scholars. This study reveals that due attention is given to dialogue synchronisation in terms of matching duration, pauses and lip movements, even though the last-mentioned also presents inconsistencies. Naturalness in the target language and speech tempo are prioritised and pursued, though sometimes affecting other quality parameters, such as rhythmic synchrony that emerges as a major challenge throughout. Other recurrent patterns across the data sample analysed include unnatural intonation (mostly at a speech melody level), and instances of hollow sound quality, particularly in outdoor scenes.

The present study attempts to tackle each of the challenges that have emerged, and that shall be outlined below, by proposing applied strategies that aim to achieve the generally accepted dubbing quality standards. It focuses on interlingual lip-synch dubbing, as opposed to phrase-synch dubbing or intralingual lip-synch dubbing (the latter being used mostly in the case of American English products dubbed into British English, or vice versa, especially for children's programmes). While most of the strategies proposed below are applicable to other dubbing languages as well, the primary focus is on English dubbing, being one of the "new entries" in the localisation industry.

This "newness", for both viewers and professionals, and the lack of consolidated norms, conventions and textual stock equivalents can be used to the advantage of the Anglophone dubbing industry. English dubbing is still in time to shape its own "dubbing personality" because it is not yet contaminated by an engrained tradition, approach, or previous influences and strategies. This tabula rasa scenario provides the opportunity to determine which quality standards to prioritise and how to achieve them.

2. Challenges Faced by the Industry in English-Language Dubbing

Lack of habituation needs to be taken into account. Anglophone audiences are not accustomed to the consumption of dubbed products, and this certainly has an influence on viewer response. In other words, a positive or negative English dubbing experience does not depend solely on the quality of the dubbed product (Spiteri Miggiani, 2021). This premise is supported by the *dubbing effect* theory posited in Di Giovanni and Romero-Fresco (2019), and Romero-Fresco (2020), whereby viewers who are not accustomed to dubbing do not seem to possess that "self-defence" mechanism that unconsciously draws their attention away from the mouth and lip movements on screen, allowing them to enjoy the product. It is likely that Anglophone viewers do not possess such viewing mechanisms since these need to be developed over time and through continuous dubbing consumption (Zabalbeascoa, 1993, p. 248).

It might take a while before the audience is able to accept the *cultural discord* (Spiteri Miggiani, 2021)—that visual and aural mismatch or contrast between the body language on screen and the

verbal language heard (mainly on a suprasegmental level). The gestures, facial expressions and body language in general may not be easily associated with English-speaking cultures, and more importantly, with an English accent and pronunciation (for instance, a British accent heard over an Italian or Spanish actor that gesticulates). That said, habituation may help ease this contrast, as well. It all depends on how much the Anglophone audience is willing to close an eye when determining its own tolerance threshold and embracing a dubbing suspension of disbelief (Romero-Fresco, 2006).

In any case, the response towards dubbing does not depend solely on its audience. The responsibility also falls on the professional practice itself and the quality levels attained. The difficulties brought about by lack of familiarisation with dubbing apply also to the professionals involved. These were suddenly faced with the need to quickly adapt to new localisation demands with perhaps little or no time for extensive training. Many dubbing actors and translators/adaptors may very well be at their first dubbing attempts, and the lack of a well-consolidated point of reference raises the level of difficulty.

Defining quality in dubbing, without falling into the trap of prescriptivism or a narrow view based on a limited sample of dubbing cultures and languages, is no easy task. Along the years, researchers have put together a number of basic quality parameters that seem to make dubbing work in the eyes of viewers and practitioners, among these Whitman-Linsen (1992), Ávila (1997), Chaume (2007, 2012, 2020), Spiteri Miggiani (2019). Chaume (2007, 2020) discusses the importance of acceptable lip-synch, credible and natural-sounding dialogue, fidelity to the original product, semiotic cohesion between words and images, clear sound and volume, and adequate role interpretation.

Achieving such dubbing quality standards is the end goal of the strategies proposed in this paper. The identification and analysis of English dubbing challenges and patterns carried out in the previous study (Spiteri Miggiani, 2021) is based on a revisited list of quality parameters adapted further from Chaume (2007), while also proposing a subdivision into two categories: textual and non-textual.

Textual parameters:

- Adequate lip synchronisation.
- Natural-sounding dialogue.
- Cohesion between dubbed dialogue and visuals.
- Fidelity to source text.
- Agreeable phonaesthetics.

Non-textual parameters:

- Suitable voice selection.
- Convincing voice performance.
- Natural-sounding intonation.
- Appropriate sound quality.

Both categories are in actual fact interrelated, and every quality parameter has an impact on the overall result. Therefore, this is not meant to be a clear-cut division, but is simply intended to highlight and distinguish the main agents responsible for each category. The strategies proposed apply to the dubbing workflow as a whole, though they focus particularly on the dubbing text adaptation process. Generally speaking, it can be said that the textual quality parameters fall under the responsibility of dubbing translators and adaptors, while the other professionals are directly responsible for the non-textual parameters. This does not exclude intervention in the text (for enhanced synchrony, among other reasons) on behalf of dubbing directors, actors, supervisors and other dubbing professionals that belong mainly to the non-textual category. The overall synchronisation of a dubbed product does not depend only on text adaptation, but also on the dubbing actors, assistants and sound engineers. The text, of course, has an impact on both the actors' performance and on their use of tone, though the typical prosodic delivery that usually characterises dubbed speech – referred to as *dubbitis* by Sánchez-Mompeán (2020, p. 148) – seems to be mainly a result of the voice-acting process (Ávila, 1997; Chaves, 2000; Sánchez-Mompeán, 2020), as well as the recording modus operandi itself (Spiteri Miggiani, 2019), that is, the actual workflow and practical techniques, as will be discussed further on. For this reason, naturalness in the use of language and naturalness in the actors' use of intonation are considered distinct quality standards since they depend on separate factors and professionals.

The following sections draw on the classification of quality parameters outlined above in order to individually examine a number of identified challenges, together with possible strategies that can be applied by the professionals involved.

3. Non-Textual Parameters: Respective Challenges and Strategies

3.1. Sound Mixing and Editing

In dubbing, the achievement of optimal sound implies clarity, lack of noise or interference, and adequate volume levels. The newly recorded voices need to blend into the original music and effects track; this encompasses background noise and any dialogue retained in the original. When dialogue alternates between dubbed scenes and scenes that retain the original language, the shift in audio needs to be imperceptible, otherwise it will result in an annoying distraction, possibly breaking the dubbing suspension of disbelief.

It is important to ensure sufficient background noise (also referred to as general walla or buzz or ambience) because this adds depth and realism to the sound quality, therefore contributing to the credibility of the end product (Spiteri Miggiani, 2019, p. 143–148). Sound technicians need to make sure that volume levels are sufficient, and at times, background noise needs to be enhanced in the

dubbed track, especially when part of the original background noise track is lost. This may imply having to record and include additional—perhaps indistinct—dialogue lines, a strategy that also enables the dubbed version to feature general background murmur in the target language. Likewise, due attention needs to be given to television sets, computers, radio sets, mobiles or other devices playing audiovisual content. The possibility and feasibility of replacing these voice tracks must be taken into consideration.

A recurrent pattern identified in the aforementioned study (Spiteri Miggiani, 2021) is an evident hollowness in sound, and dryness in the dubbed voices, as though these were detached from their surroundings. This *soundproof studio effect* seems to emerge mostly in outdoor scenes. Enhancing reverberation can help merge voices with room tones or environment acoustics.

Another key element in dubbed tracks is consistency in volume levels when two characters are at the same distance from the camera. A discrepancy in the volume of speech between one character and another distracts the viewers, reminding them that the conversation has been fabricated in a studio. This may simply be due to the various actors not keeping the same distance from the studio microphone.

Likewise, long shots and distance voice effects need to be catered for accordingly, and this could mean instructing the dubbing actors to move further away from the microphone while recording, especially when the desired effect cannot be created digitally. Utterances that switch between long shots and close up shots would require alternating volume levels, while effects would need to be added when speech is uttered over the phone or behind a glass, or any other sound filter.

3.2. Voice Selection, Performance, Delivery

Choosing voice qualities in accordance with physique, character role, gender and age of on-screen actors ensures credibility (Bosseaux, 2015, p. 59). Voices need to meet viewer expectations as to what a character should sound like (Whitman-Linsen, 1992). Dubbing children seems to pose a challenge, at times resulting in voices that may sound slightly artificial (or almost synthetic), most likely due to the pitch settings required to make the actors in question sound younger. This is because children are often dubbed by female adults, or children who are slightly older than the characters played. Other than that, the previous study (Spiteri Miggiani, 2021) reveals compliance with global dubbing standards when it comes to voice selection and voice variety. On the whole, the characters are allocated to different voice actors (with perhaps some exceptions where the same voice actor dubs more than one character in the same series episode), while due attention is given to the selection of suitable voice qualities according to physique, character role, gender and age.

In order to enhance body-voice adherence, a strategy which is occasionally used in English dubbing is that of having the original actors dub themselves in English, where there is adequate language proficiency. Apart from body-voice adherence, an advantage provided by this method is that it can help minimise the *cultural discord effect* mentioned earlier. This method may result in a combination

of English language flavours in the dubbed versions—the foreign flavour belonging to the original actors who dub themselves, and the English accent belonging to the American or British dubbing actors who dub the rest of the characters. Hence, due caution needs to be applied, since variety in accent is also used to mark cultural or linguistic "otherness." If only part of the original acting cast dubs itself, the dubbing director would need to ensure that both native English accents and foreign English accents blend in unnoticeably, without "marking" specific characters when this difference is not meant to emerge in the plot.

A key feature that is closely linked to role performance and voice delivery, and that requires due attention, is *intonation*, in particular speech melody and emphasis. As corroborated in the data and results of the previous study (Spiteri Miggiani, 2021), a general flat *speech melody* tends to emerge as a pattern. This happens even when the original actors dub themselves. In general, a wider range of rising and falling tones would contribute towards a more natural delivery. On the contrary, an effort to seek naturalness in *emphasis* is evident thanks to the use of elongation and stressed syllables, especially in the case of monosyllabic words. This, however, results in synchronisation issues in those instances when no elongation is present in the original delivery. In other dubbing languages, this dragging effect is used as a strategy to achieve synchronisation (Baños Piñero, 2009; Sánchez-Mompeán, 2020), while in English it appears to be counterproductive. That said, the actors' delivery depends highly on the text adaptation, as will be discussed in section 4.1.1.

Another factor that may impinge on the actors' delivery is the recording modus operandi itself. Listening to the original audio via headphones as they record the target text can influence the actors' delivery who may unconsciously mirror the original language intonation (Spiteri Miggiani, 2019, pp. 74–75). The use of the rythmo band and its karaoke-like features, enables dubbing actors to do without the original audio and ensures detachment from the original speech melody. This has therefore proven to be a useful device to help avoid imitation of the source language intonation that might sound less natural in the target language. Very often, dubbing actors can choose whether or not to listen to the original audio during the recording takes. Not listening to the original voices would therefore help avoid source delivery interference. That said, in those studios where rythmo band tools are not adopted, it is necessary for the actors to listen to the original audio while voicing the target dialogue in order to capture the rhythm. In this case, awareness and relentless focus is the only way to enhance naturalness in intonation.

Last but not least, a matching delivery rate is usually required to meet lip synchronisation demands, thus implying that dubbing actors cannot use a slower pace than that of the original language in an attempt to attain a more natural sounding English. Again, delivery rate falls naturally into place once this has been catered for in the text adaptation phase.

4. Textual Parameters: Respective Challenges and Strategies

The textual quality standards are inextricably intertwined since they all influence each other. Certain choices in favour of lip synchronisation may have an impact on the extent of naturalness in the target dialogue, and vice-versa: prioritising natural-sounding language may entail a sacrifice in the degree of corresponding lip movements. All this can be the result of a *conscious weighing process* (Krings, 1986, p. 268; Chesterman, 2000, p. 82), on behalf of dialogue writers, or of consequences that possibly arise due to unconscious cognitive processes triggered by the adaptation workflow itself (Spiteri Miggiani, 2019, pp. 88–91). Sometimes, fidelity to the original text is necessarily sacrificed to satisfy some other quality parameter, most likely cohesion between words and visuals. It is by no means an easy task to strike a balance between quality standards; they resemble elastic strings pulling from all directions, each at a different level of tension.

Translators who may need to learn dubbing adaptation skills to meet client demands may ease the process by breaking down the text adaptation workflow into single tasks, and move through them in a specific order. With time and experience, adaptation becomes an automated process and in actual fact most tasks occur simultaneously on a cognitive level. The following list provides a possible text adaptation workflow model, following an initial raw translation of the original dialogue.

Text adaptation workflow:

- Language-focused preadaptation.
- Identification of pauses and rhythm in a given dialogue line.
- Insertion of dubbing notations and tempo markers (if applicable).
- Dialogue line rewriting to match duration (by resizing the length).
- Dialogue line rewriting to ensure that body language and words correspond.
- Dialogue line rewriting to match lip movements.
- Postadaptation language fine-tuning.

The above model highlights the numerous rewriting attempts required to attain an optimal version. English-language dialogue writers would very likely need to adapt a raw translation carried out by another professional, and may not be familiar with the original language at hand. One of the explanations to this type of workflow that subdivides the text preparation into two phases (translation and adaptation) is the current localisation talent shortage. It is not easy to find trained translators or dialogue writers who are able to translate as well as adapt audiovisual content from multiple languages into English for dubbing purposes. Separating the two roles makes it possible to entrust the raw translation from French, Turkish, Spanish, Swedish and so on, to various translators, while the adaptation of the various language productions can be carried out by same dialogue writer.

This role separation brings about its challenges. To this end, a language-focused revision of the target text (without the video) prior to initiating the technical adaptation phase, is strongly recommended in order to ensure that the text flows naturally in English and can act as a reliable "base" for

adaptation. Language revision and fine tuning would need to be repeated at the very end of the text adaptation process, as well.

The *dubbing notations* (Spiteri Miggiani, 2019, p. 129–143) refer to technical notes inserted in the dubbing script provided that these are requested by the client. Their purpose is to support the actors during the recording process. These highlight the visibility, or otherwise, of mouth articulatory movements on screen (e.g., OFF or ON, to indicate whether off-screen or on-screen utterances), pauses, or paralinguistic features (non-verbal mouth uttered sounds such as *sighs*, *eats*, *breathes*) (Chaume, 2012). Based on in-house observation, it appears, so far, that Anglophone localisation companies rarely adopt such notations and, at most, simply carry over the indications present in the original script. The increase in the use of software tools for dubbing is slowly eliminating the need to include any notations in the adapted text. Dubbing software generally adopts the rythmo band tool which already offers several visual aids to the actors. Pauses, for example, are simply indicated by gaps in between words. The use of the rythmo band is widespread among US and UK dubbing studios and appears to be the norm.

4.1. Lip Synchronisation Strategies

Adequate lip synchronisation or lip-synch is no doubt one of the quality standards that can have a huge impact on viewer reception. In professional practice the term is used with a broad meaning and encompasses:

- **Timing:** matching the duration, that is, the start and end of dialogue lines (referred to as *isochrony* in audiovisual translation (AVT) studies).
- **Tempo:** matching the internal rhythm and mouth flap frequency (can be referred to as *rhythmic synchrony*) (Spiteri Miggiani, 2019, 2021).
- **Lip movements:** matching the mouth articulatory movements, in particular bilabial consonants, labiodental consonants and lip-rounded vowels (referred to as *phonetic synch* in AVT studies).

Two adaptation approaches that possibly account for challenges in English dubbing lip sychronisation are:

1) priority given to natural-sounding language that could derive from the absence of a *dubbese* tradition in English;

2) priority given to natural-sounding speech tempo, given that English speech delivery is slower than that of a number of source languages (Pellegrino et al., 2011).

As for the first approach, prioritising natural sounding language seems to come at the cost of lip-synch. The existence of a standardised English dubbing language, that has been accepted over time by viewers and also tries to seek an acceptable balance between all quality parameters (lip-synch

included), could actually be beneficial since it would guide the dubbing adaptors' choices. With regard to the second approach, matching the speed and tempo of most source languages would result in an unnatural pace in English for both dialogue writers and actors, as well as for Anglophone viewers. However, adopting a more target-oriented natural speed will have an impact on the lip-synch outcome if this implies adopting a slower speech tempo. This constitutes a significant challenge in English text adaptation that raises the degree of difficulty for dialogue writers.

Identifying the reasons behind lip-synch challenges in English can help dialogue writers channel their attention, choices, and awareness. The following sections tackle possible strategies based on the explanations proposed above.

4.1.1. Timing and Tempo

Rhythmic synchrony (and consequent matching mouth flap frequency) is crucial and depends highly on the speech tempo adopted by the actors. The importance of tempo is discussed in detail by Fodor (1976, p. 31), Spiteri Miggiani (2019, pp. 76–79) and Sánchez-Mompeán (2020, p. 30). The text adaptation strategy proposed is to reproduce the same rhythmic pattern, reciting the target line at the same speed and volume while mirroring the same recitation style and respecting the original pauses, including any short breath intake before pronouncing a phrase. All these have an impact on speech tempo. Whispering, for instance, tends to increase the speed of dialogue, while a loud angry tone tends to slow down pronunciation and articulation. Also, focusing on clear articulation can slow down the speech rate of dubbing actors and this needs to be taken into account by dialogue writers. Dialogue writers need to develop this type of *rhythm sensitivity* in order to provide the voice actors with a script that can achieve optimal synchronisation when combined with the dramatisation process.

Rather than aiming at a similar mathematical count in terms of syllables or phonemes, dialogue writers are enticed to capture the original speech tempo and apply it to the target lines (Spiteri Miggiani, 2019, p. 76; Sánchez-Mompeán, 2020, pp. 28–32). Emulating the actors, therefore, rehearsing the rhythmic pattern by repeating and imitating the original dialogue line, aloud, can be helpful. Likewise, replacing the words with mere gibberish or a repeated sound can help as well, since the aim is to capture the tempo and flow in order to overwrite it with a new set of words, in English.

Capturing the speech tempo is more likely to ensure that the target language reproduces the same number of mouth flaps. It is important to note that a target text line could match the timing but not the tempo: it could match the total duration and length of a source text utterance, but not necessarily the rhythm and duration of the individual sounds within the dialogue line. In English-language dubbing, lack of rhythmic synchrony seems to be the main issue as far as lip synchronisation is concerned. This generally implies noticeable empty mouth flaps half-way through dialogue lines.

As mentioned earlier, a possible explanation to this is the attempt to reproduce natural sounding *speed* (and *dialogue*) in English. English tends to have a slower pace when compared to some other

European languages such as Spanish or French (Pellegrino at al., 2011). The speech category that can perhaps be referred to when analysing non-fictional productions is conversational or spontaneous speech (as opposed to news, radio, speeches, interviews etc.) since this is the type of speech or prefabricated orality (Baños-Piñero & Chaume, 2009) that screenwriters generally try to mimic. Spontaneous speech rate varies according to language; a cross-linguistic comparative study that analyses information density in relation to syllabic rate and information rate in a selected number of languages reveals Japanese as having the fastest speech rate with an average of 7.84 syllables per second (sps) followed by Spanish (7.82 sps), French (7.18 sps), Italian (6.99 sps) and English (6.19 sps). German (5.97 sps) and Mandarin (5.18 sps) have a slower speech rate than English (Pellegrino et al., 2011).

German and Mandarin aside, the adoption of a natural English language delivery speed when dubbing from Spanish, Italian or French will most likely result in empty mouth flaps and a noticeable slower auditory articulatory pace when compared to the faster and visible articulatory movements on screen. This means that if lip synchronisation is chosen as a priority in English language dubbing, English-dubbing viewers would need to get accustomed to listening to a perhaps-less-natural faster speech rate. In time, this would need to fall within their dubbing tolerance threshold.

When it comes to rhythm, not much flexibility can be applied because it is a key feature that actors, too, rely on when recording their lines. This helps to guarantee the synchronisation of the dubbing scripts to the visuals. It is the only thread that connects dialogue writers and actors—a common framework or stencil that acts as a point of reference for both. Synchronisation is achieved when actors and dialogue writers apply the same rhythm and tempo while reciting the dubbing script, therefore when the two are in symbiosis. It could be argued that matching speed is less important during off-screen dialogue lines. Off-screen shots certainly allow more freedom, but dialogue writers should ideally try to respect the same duration and tempo in any case, since this enables actors to preserve that symbiosis and rhythmic feel of the text. This is especially important in certain types of shots, such as a continuous shot-reverse-shot sequence alternating between one character and another in a conversation whereby the actor who is speaking is not always visible on screen.

Needless to say, the use of software that integrates the rythmo band method offers the above-mentioned rhythmic framework that acts as a common tool for both dialogue writers and dubbing actors, thus enhancing symbiosis and accurate synchronisation when it comes to speech tempo. The graphical representation of words, stretched and compressed according to the tempo, as they travel beneath the images, are an extremely useful visual aid.

The frequent use of monosyllabic words in English certainly challenges timing and tempo synchronisation in dubbing. These will naturally create a mismatch when placed in coincidence with disyllabic or multisyllabic words in the source language and their respective visible mouth flaps on screen. Besides, the elongation that usually accompanies monosyllabic words creates a slightly delayed effect as against the visible mouth movements. This characteristic, coupled with the conciseness of the English language in general, often calls for amplification, that is, the expansion of

text, to match timing and tempo. In other words, the translated and adapted lines in English dubbing tend to be shorter than their original counterparts.

4.1.2. Language Adaptation Devices

Amplification and reduction are among the main translation strategies mentioned by Chaume (2012, pp. 72–73, 2020, p. 112) in his non-language specific discussion on isochrony. Amplification, therefore, is more likely to be frequent in English-language dubbing. Expansion may prove to be more challenging than reduction since this must occur in a natural-sounding way, therefore also adhering to another quality standard.

Drawing on classical rhetoric and poetics, the four principles of modification or *quadripartita ratio* proposed by Quintilian (2001 [95 CE]) provide a model that can be applied to dubbing adaptation: *adiectio* (addition), *detractio* (omission or reduction or condensation), *immutatio* (permutation or change in form) and *transmutatio* (change in order) (Spiteri Miggiani, 2019). *Adiectio* or addition can clearly be used as an expansion device, and the most common form of addition is *geminatio*, that is, repetition. The two types of repetition are *anaphora* and *epiphora*: repeating the same word or sequence of words at the beginning or at the end of a neighbouring clause, respectively. Repetition is possibly one of the easier strategies to resort to when trying to reach the desired length and duration, but it is important to bear in mind that classical rhetoricians propose repetition with the intention of emphasis as a desired effect. Therefore, intention and viewer perception need to be taken into account when resorting to repetition.

Other expansion techniques include explicitation (adding further detail to a given phrase or line); periphrasis (the use of a longer expression with the same meaning, e.g., "I am going" instead of "I will," "I was wondering if" instead of "could I…?"); shifting from active voice to passive voice; the use of longer synonyms; fillers and discourse markers; phatic expressions; interjections; question tags and vocatives.

Fillers such as "um," "er" and other filled pauses or hesitation markers, and discourse markers such as "you know," "so," "well," "like," "I mean" can have an interactional function (e.g., politeness) or a cognitive function (e.g., a character processing thoughts). Such tools provide handy text expansion escamotages, however their overabundant use may also be counterproductive and can result in an unnatural and cacophonic dialogue.

The use of interjections (e.g., oh, wow, oops, ugh, ouch etc.) could also have a counterproductive effect if added in the target language when these are absent in the original text because they interrupt the *rhythmic flow*. Ideally, they should match the timing and positioning of interjections in the original text especially at the beginning of dialogue lines. This tends to ensure more accuracy in rhythmic synchrony which is then carried over the rest of the dialogue line. In general, it is important to replace original interjections with their equivalent in the target language, be it American English

or British English, ideally without *over-domesticating* since this would contribute to enhancing the *cultural discord* mentioned earlier.

Phatic expressions are an extremely useful device, not only for amplification purposes, but also to enhance rhythmic synchrony, and therefore match mouth flap movements, especially when trying to compensate for monosyllabic words in English. These include expressions used as greetings or leave-taking (e.g., good morning, how are you?, hi, I have to go now, goodbye); statements that aim to start conversations or that are used for small talk or politeness (e.g., excuse me, can I borrow you for a minute?, I was wondering if, by the way, let me see, I agree with you, I see your point etc.); and phrases that are used for apologising, giving compliments, thanking or criticising (e.g., thanks, I appreciate it, I'm sorry, there's another way to look at this, congratulations, you did it well etc.). All these can make versatile additions to the text without communicating extra semantic content that is not present in the original. Some of the above techniques have also been tackled by Chaume (2008, pp. 135) though not applied to English.

4.1.3. Lip Movements

The strategies to match mouth articulatory movements do not differ from those adopted in other target languages. Lip-synching entails matching mainly labial consonants (p, b, m, v, f) as well as labialised consonants and vowels (o, u, w) especially in close-up shots and at the beginning and end of dialogue lines (Fodor, 1976; Chaume, 2012, 2020; Spiteri Miggiani, 2019; Sanchéz-Mompeán, 2020). This is done by choosing words in the target text that contain the desired consonants or vowels. Bilabial consonants (p, b, m) and labiodental consonants (f, v) are almost interchangeable — a word containing any of these consonants may be used to match a word containing any bilabial or labiodental consonant in the original text. It could be useful to bear in mind that when translating from German, the /w/ is articulated and voiced as a labiodental consonant. When translating from Spanish, the /v/ is articulated as a voiced bilabial, while the /w/ could either be voiced as a bilabial or articulated as a lip rounded vowel, depending on the word. Thus, these need to be treated as such.

Table 1 illustrates an ideal phonetic synch scenario. Needless to say, it is impossible to match all mouth articulatory movements, and doing so would then impinge on the language in terms of correctness and naturalness. A balance between parameters needs to be sought, even if this may imply prioritising and sacrificing one over the other depending on the specific scene, camera shots and semantic content.

Table 1.

Phonetic synch interchangeability chart

Lip Articulatory Movement	Source Text	Dubbed Text
Bilabial consonants	/m/ /p/ /b/	m/ /p/ /b/ /f/ /v/
Labiodental consonants	/f/ /v/	/m/ /p/ /b/ /f/ /v/
Labial consonants	nil	nil
Rounded vowels & semivowels	/o/ /u/ /w/	/o/ /u/ /w/

Note. The consonants and vowels in the dubbed text column can replace the ones in the source text column. When there are no labial consonants in the source text (nil), the ideal scenario is to avoid having unnecessary labial consonants in the dubbed text.

Source: Author

While many studies focus on labial consonants, mouth opening movements that accompany vowels are not to be underestimated. A final word ending in /o/ at the end of a line, perhaps in a close-up shot, would ideally require a word that features the same or a similar vowel (/o/, /u/ rather than /i/, /e/). A mismatch can be aesthetically distracting.

The previous study (Spiteri Miggiani, 2021) reveals distraction resulting from the presence of labial consonants in the English dubbed version when these are absent in the original text. In some instances, several rewinds are necessary in order to understand the words articulated and, interestingly enough, these instances often coincide with mismatched labial consonants. When applied to dubbing, some cognitive studies seem to point at miscomprehension when a *labial visual mode of articulation* coincides with a *non-labial auditory stimulus*, and vice versa, when a *non-labial visual mode of articulation* coincides with a *labial-auditory stimulus*. Surprisingly enough, the latter seems to have the stronger miscomprehension effect (Möttönen & Sams, 2008).

In order to achieve effective lip-synch, optimal positioning of the said consonants and vowels is imperative; they need to be as close as possible in pronunciation-timing to the original words in question, otherwise a correct match of these labials and vowels would be futile. Misplaced labial consonants do not create a visual coincidence with the original. Also, the effort to place labial consonants in a dubbed text, when the shots do not require phonetic synch, can be avoided, (for instance, long shots or off-screen dialogue where the mouth is not visible). It is imperative to identify when lip-synch is to be given priority. The reproduction of the same rhythmic synchrony or speech tempo as the original is the only way to ensure lip movement coincidence. This needs to be applied by both dialogue writers and dubbing actors.

Most often, creative solutions are necessary in order to respect phonetic synch. The ending of a dialogue line in Spanish may read "diez y media" (ten thirty), while the English version may need to read "ten thirty a.m." to match the /m/ in "media" (especially if this can be moved up slightly in the

sentence structure). The word "también" in Spanish may very well be translated as "also" or "as well," though variations on meaning may be necessary. "Me too," "same here," "besides," "furthermore," "moreover" reworked into the target text might need to replace the exact equivalent. Indeed, one of the pitfalls of dubbing text adaptation is the risk of getting carried away by phonetic synch and deviating significantly from the original meaning, or also losing touch with the target language and the way it sounds. Getting carried away by phonetic synch can also be a consequence of the various cognitive processes triggered by the adaptation workflow itself (Spiteri Miggiani, 2019, p. 88–91). This is why going through the final adapted text draft without using the video or audio is a recommended self-revision and fine-tuning phase in the dialogue writing process (see section 4). Such *postadaptation language fine-tuning* can also help to identify any translational routines and consequently choose to retain them, or otherwise, with detached awareness.

4.2.   Natural Dialogue and Text-Induced Intonation

Ample attention seems to be given to natural sounding dialogue in English dubs. In some instances, a small degree of *over-domestication* emerges as a result of the use of slang expressions and target-oriented colloquialisms (or Americanisms) that do not usually belong to standardised dubbed language in other dubbing cultures (Spiteri Miggiani, 2021).

In other dubbing languages, natural sounding language as a quality standard and norm does not imply reproducing spontaneous target-culture oral discourse. This is not to say that English dubbing necessarily needs to play by the same set of norms. Long-established dubbing cultures can be taken as a model on the grounds of positive viewer response in those territories, however English dubbing can also choose to distance itself by seeking customised strategies, while being aware of the norms that usually govern dubbing elsewhere. Anglophone viewer response will in any case vary because viewers are not accustomed to dubbing in general, irrespective of the approach adopted. Having said that, establishing a norm – be it in line with other dubbing languages, or not – is fundamental to start building audience habituation. To do this, a certain amount of consistency across English dubbed products is necessary. A possible way to establish norms and consistency is to create a standardised model to be followed.

Considering the "newness" of English dubbing, it may not be that easy to identify English-language *dubbese* patterns, especially if there is an effort to reproduce naturalness in dialogue to the extent of resorting to over-domestication. Despite this, clearly identifiable source calques seem to emerge, mainly on a syntactic and structural level, as well as instances of literal translation (Spiteri Miggiani, 2021). In this, English dubbing seems to mirror other dubbing languages thus pointing towards the fact that the adaptation workflow itself – adopted across cultures – is largely responsible for the target text outcome.

That being said, it is important to consider that the idea of "naturalness" (Romero-Fresco, 2006) in dubbing dialogue does not refer to spontaneous oral discourse. The so-called *dubbese* (Pavesi, 1996)

is prefabricated orality, just like its original counterpart, in that it tries to imitate spontaneous speech but the end result is necessarily a language that is somewhere along the written and spoken continuum (Baños-Piñero & Chaume, 2009). Ironically, the absence of an English-language *dubbese* that the audience is familiar with could be one of the factors leading to challenges related to the other parameters. Once the English dubbing audience becomes accustomed to a dub-sounding language and accepts it as "natural," it will be easier to adapt the text while achieving the other quality parameters, in particular lip synchronisation. It is not realistic to aim at the same degree of achievement of both spontaneous-sounding language and lip synchronisation because dubbing adaptation implies trying to fit one language system into another, while hoping that the cut-and-replace process will go unnoticed by the viewers. In other words, a small extent of *dubbese* is probably inevitable if all the other quality standards are to be met. Conscious awareness is key to maintaining an acceptable balance.

Indeed, dubbed language in other cultures is generally quite standardised (Baños Piñero, 2006). Interestingly, Pavesi (2016) states that some of the typical dubbese features—mainly a certain amount of *formulaicity*—are actually what enhance viewers' processing of the dialogue and "feelings of shared identity and belonging to the same lingua-cultural community" (Pavesi, 2016, p. 101). Formulaicity is generally carried through in dubbed texts via translational routines, that is, translation options that through repetition become stock solutions in the target language. These generally also include semantic, structural and pragmatic calques as well as loan creations deriving from the original text (Pavesi, 2016, p. 102).

In this context, it is worth highlighting, once again, the challenge that lies in trying to reproduce a natural prosodic dialogue delivery as opposed to the typical delivery (dubbitis). The target text adaptation can undoubtedly help the actors' performance and use of intonation; the way dialogue lines are structured and organised can lead to a specific emphasis or rising or falling tones. An optimal strategy is to avoid obscure and long-winded target lines that may be difficult to read or understand, hence sharpening the language in a way that enables the actors to identify the intention and consequently the reading intonation at a first glance, possibly without the intervention of the dubbing directors.

Of course, other strategies can be adopted to enhance the functionality of a dubbing script and contribute to the actors' delivery, namely, a dubbing-customised use of punctuation, over and above the typical tempo markers (/, //, ..), when these are applicable. This could imply adopting punctuation (mainly commas) in such a way as to help actors recite the text with the desired rhythm, irrespective of grammar rules. Commas, full stops and semicolons can be seen and used as functional tools to guide the actors' *rhythmic flow* rather than devices used solely to subdivide sentences on a semantic level. After all, punctuation will not be seen by viewers but is only meant as a working tool for actors. It is also one of the few devices that dialogue writers can use to communicate with the actors. Adding commas where they usually would not be required in a written text for reading may actually highlight emphasis, a desired intonation or a slight pause. Likewise, adding accents may indicate correct emphasis in the case of words that have homonyms.

The use of the rythmo band, which is widespread in English dubbing studios, sometimes entails an extra challenge for dubbing actors because the words travel beneath the images at the same pace as the speech utterances. Dubbing actors may not always be able to view the whole dialogue segment recorded in a given loop. It is helpful when the software enables the actors to also view the whole text segment, so that they can identify the intention as well as the punctuation at the end of each line, for example, in the case of interrogative statements. This might help reduce line rehearsing before finalising a recorded version. When the software application does not enable actors to view the whole dialogue segments over and above the dialogue bites visible at the bottom, it might be advisable to enhance punctuation accordingly. This may imply, for instance, repeating question marks or exclamation marks at the beginning of dialogue lines (similarly to the use of inverted punctuation in Spanish (¡, ¿) in order to set the intention and intonation prior to the first visible words in the rythmo band.

4.3.   Phonaesthetics

Taking phonaesthetics into account results in less cacophonic sounds (e.g., consonant clusters or too many s's in a dialogue line thus creating a hissing sound). Avoiding such clusters also sustains the actors in the articulation and pronunciation of dialogue lines. Consequently, less time is wasted during recording. This also contributes to the creation of a pleasant-sounding target text especially in monologues or conversations where the dramatic and emotional impact also depend on language style or poetic features. Avoiding repetition of the same words, redundancy, as well as unnecessary rhyme, all enhance the phonaesthetic quality of a dubbing script.

5. Conclusions

Exploring English-language dubbing has highlighted two main challenging factors that can make or break the required balance between quality parameters, and consequently influence the viewer experience. These are: rhythm and sound. The rhythm falls mainly under the responsibility of dialogue writers, and later of voice actors under the guidance of dubbing directors (and sometimes dubbing assistants), while sound is taken care of by audio technicians and dubbing directors. Once the dialogue writer has assigned the appropriate rhythm to the adapted text, in the initial phase of the workflow, then the rest will follow and fall into place – including the actors' performance and synchronisation to the original. On the other hand, sound mixing and editing take place in the final phase of the workflow and can ultimately enhance or undo the work carried out in the previous phases.

When moving through the dubbing workflow, professionals need to be aware of how, where, and why the workflow itself and the chosen modus operandi may have an impact on the quality parameters and final outcome. This awareness can redirect attention and priorities, and enhance supervision or quality control where needed.

English-language dubbing has yet to establish and consolidate its norms across the various roles and tasks in the entire process. It is therefore still in time to shape its own practice as well as viewer response. This can be done by referring to the norms and quality standards that govern other successful dubbing cultures, while at the same time identifying and acknowledging English-specific demands and customised strategies. Indeed, this study proposes a number of possible strategies that hopefully can be corroborated, or otherwise, through further research. This can be done through applied studies in an academic setting or through collaboration with localisation companies, where such strategies can be tested and observed directly in professional practice. This kind of knowledge transfer between academia and industry would be of great value—it implies direct application of scholarly research in the field, thus contributing towards the shaping of an emerging localisation industry.

References

Audiovisual Media Services Directives (2018, November 28). *Official Journal of the European Union*, volume 61, article 13. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2018.303.01.0069.01.ENG&toc=OJ%3AL%3A2018%3A303%3ATOC

Ávila, A. (1997). *El Doblaje* [*Dubbing*]. Cátedra.

Baños Piñero, R. (2006). *Estudio descriptivo-contrastivo del español oral en una serie de TV de producción propia y en una serie de TV de producción ajena. El caso de Siete Vidas y Friends* [*A descriptive-contrastive study of oral Spanish in a local TV serial production and a foreign TV serial production. The case of Siete Vidas and Friends*] [Unpublished Master thesis, Universidad de Granada].

Baños Piñero, R. (2009). *La oralidad prefabricada en la traducción para el doblaje. Estudio descriptivo-contrastivo del español de dos comedias de situación: Siete Vidas y Friends* [*Prefabricated orality in translation for dubbing. A descriptive-contrastive study on sitcoms: Siete Vidas and Friends*]. [Doctoral dissertation, Universidad de Granada]. https://hera.ugr.es/tesisugr/18319312.pdf

Baños-Piñero, R., & Chaume, F. (2009). Prefabricated orality: A challenge in audiovisual translation. In M. Giorgio Marrano, G. Nadiani & C. Rundle (Eds.), *The translation of dialects in multimedia* [Special Issue]. *inTralinea.* http://www.intralinea.org/specials/article/1714

Benasuly Amit, L.B., Yekutieli, S., Issacharoff, A., & Raz, L. (Executive Producers) (2015–present). Fauda [Netflix Series]. Tender Products.

Bosseaux, C. (2015). *Dubbing, film and performance: Uncanny encounters*. Peter Lang.

Chaume, F. (2007). Quality standards in dubbing: A proposal. *TradTerm 13,* 71–89. http://www.revistas.usp.br/tradterm/article/view/47466/51194

Chaume, F. (2008). Teaching synchronisation in a dubbing course. In Díaz-Cintas, J. (Ed.), *The didactics of audiovisual translation* (pp. 129–140). John Benjamins.

Chaume, F. (2012). *Audiovisual translation: Dubbing.* St. Jerome Publishing.

Chaume, F. (2020). Dubbing. In L. Bogucki & M. Deckert (Eds.), *The Palgrave handbook of audiovisual translation and media accessibility* (pp. 103–132). Palgrave Macmillan.

Chaves, M. J. (2000). *La Traducción cinematográfica: El doblaje [Cinematic translation: Dubbing]*. Universidad de Huelva Publicaciones.

Chesterman, A. (2000). A causal model for translation studies. In M. Olohan (Ed.), *Intercultural faultlines* (pp. 15–27). St. Jerome Publishing.

Couvreur, D., Florentiny, F., Lang, C., Thomopoulos, A. & Weinger, S. (Executive Producers) (2018–present). *The Hook Up Plan* [Netflix Series]. Unrated Studios.

Di Giovanni, E., & Romero-Fresco, P. (2019). Are we all together across languages? An eye tracking study of original and dubbed films. In I. Ranzato & S. Zanotti (Eds.), *Reassessing dubbing: Approaches and current trends* (pp. 125–145). John Benjamins.

Dzhunkovskiy, A., Iloyan, E., Kessel, A., Shlyappo, V., Sorokin, R., Tkachenko, M., Trotsyuk; A. & Zhalinskiy, D. (Executive Producers) (2018–2019). *Better Than Us* [Netflix Series]. Sputnik Vostok Production; Yellow, Black and White.Fodor, I. (1976). *Film dubbing. Phonetic, semiotic, esthetic and psychological aspects.* Buske Verlag.

Goldsmith, J. (2019, July 19). Netflix wants to make its dubbed foreign shows less dubby*. The New York Times*. https://www.nytimes.com/2019/07/19/arts/television/netflix-money-heist.html

Jenner, M. (2019). *Netflix and the re-invention of television.* Palgrave Macmillan.

Kässbohrer, P., Murmann, M. & Montag, L. (Executive Producers). (2019–present). *How to Sell Drugs Online (Fast)* [Netflix Series]. Bildundtonfabrik; Netflix.

Krings, H. P. (1986). Translation problems and translation strategies of advanced German learners of French. In J. House and S. Blum-Kulka (Eds.), *Interlingual and intercultural communication* (pp. 263–75). Gunter Narr.

Möttönen, R., & Sams, M. (2008). Audiovisual interaction. In D. Havelock, S. Kuwano, & M. Vorländer (Eds.), *Handbook of signal processing in acoustics* (pp 731–745). Springer Verlag.

Pavesi, M. (1996). L'allocuzione nel doppiaggio dall'inglese all'italiano [*Allocution in dubbing from English to Italian*]. In R. Bollettieri Bosinelli & C. Heiss (Eds.), *Traduzione multimediale per il cinema, la televisione e la scena* [*Multimedia translation for cinema, television, and theatre*] (pp. 117–130). Clueb.

Pavesi, M. (2016). Formulaicity in and across film dialogue: Clefts as translational routines. *Across Languages and Cultures, 17*, 99–121.

Pellegrino, F., Coupé, C., & Marsico, E. (2011). A cross-language perspective on speech information rate. *Language, 87*(3), 539–558.

Pina, Á., Colmenar, J., Manubens, N., Martínez Lobato, E. & Martínez Sonia (Executive Producers). (2017–present). *Money Heist* [Netflix series]. Vancouver Media.

Quintilian. (2001 [95 CE]). In D. Russell (Ed.), *The orator's education,* books 9–10. Harvard University Press.

Roettgers, J. (2018, March 8). Netflix's secrets to success: six cell towers, dubbing and more. *Variety*. https://variety.com/2018/digital/news/netflix-success-secrets-1202721847/

Romero-Fresco, P. (2006). The Spanish dubbese: A case of (un)idiomatic friends. *The Journal of Specialised Translation*, *6*, 134–151.

Romero-Fresco, P. (2020). The dubbing effect: An eye-tracking study on how viewers make dubbing work. *The Journal of Specialised Translation, 33*, 17–40.

Sánchez-Mompeán, S. (2020). *The prosody of dubbed speech. Beyond the character's words.* Palgrave Macmillan.

Spiteri Miggiani, G. (2019). *Dialogue writing for dubbing. An insider's perspective.* Palgrave Macmillan.

Spiteri Miggiani, G. (2021). English-language dubbing: Challenges and quality standards of an emerging localisation trend. *The Journal of Specialised Translation*, *36*, 2-25

Whitman-Linsen, C. (1992). *Through the dubbing glass.* Peter Lang.

Zabalbeascoa, P. (1993). *Developing translation studies to better account for audiovisual texts and other new forms of text production* [Doctoral dissertation, University of Lleida]. http://hdl.handle.net/10803/8185