Measuring quality in translation for dubbing: a quality assessment model proposal for trainers and stakeholders

Giselle Spiteri Miggiani

DOI: 10.18355/XL.2022.15.02.07

Abstract
Quality assessment in the field of Audiovisual Translation (AVT) has been addressed by several scholars, particularly in relation to interlingual subtitling (Pedersen, 2017; Robert & Remael, 2016), intralingual live subtitling (Romero-Fresco & Martínez Pérez, 2015) and interlingual live subtitling (Robert & Remael, 2017; Romero-Fresco & Pöchhacker, 2017), but to-date no model in relation to dubbing has been proposed. As with other AVT modes, the need for a quality assessment method in dubbing arises in academic and in-house training contexts. Moreover, localization companies often resort to 'entry tests' before engaging translators. Self-assessment also proves to be one of the main challenges for trainees in a dubbing training context, and any quality assessment tools can possibly be of help. This paper proposes a tentative quality assessment model that attempts to pin down the 'errors' in a dubbing dialogue script while measuring the quality via a percentage score system. The model focuses on the translation and adaptation phase in the dubbing workflow and is therefore based on a set of textual quality parameters. These are drawn on a revisited taxonomy of dubbing quality standards (Spiteri Miggiani, 2021a, 2021b), further adapted from Chaume (2007), which takes into account the dubbed end product as a whole. The model combines an end product-oriented approach with workflow-oriented standards and expectation norms, therefore taking the industry perspective into account. This implies considering the functionality of a dubbing script as a macro quality parameter in its own right. The application of this tentative model has so far been limited to the author's academic and in-house training settings. This paper, therefore, is simply intended as a point of departure to pave the way towards applied and collaborative research that could test, validate, and further develop the proposed model.
Key words: Dubbing, translation, adaptation, quality standards, quality assessment

1. Introduction

As the language service provision industry becomes increasingly standardized, the issue of quality, from an-end product and process perspective, is being given more importance (Künzli, 2021). Trying to pin down quality and establishing common ground across countries, languages, cultures, and productions can benefit both the professional practice field and training contexts. Among the researchers that have dedicated attention to quality in translation, and more specifically quality assessment, revision, and translator training, some of the most noteworthy, in addition to several others, are Mossop (2001, 2007, 2011, 2014, 2016), Kelly (2005), Tennent (2005), Kunzli (2006), Hurtado-Albir (2007, 2015, 2017), Huertas Barros et al. (2019), and Huerta Barros and Vine (2019). The field of Audiovisual Translation studies (AVT) has also produced its fair share of didactic-oriented research (Bartrina & Espasa, 2003, 2005; Chaume, 2003, 2008; Zabalbeascoa et al., 2005; Díaz Cintas, 2008; Cerezo Merchán, 2012, 2018; Cerezo Merchán & Higes Andino, 2018; Martinez Sierra, 2008b; Bolaños García-Escribano, 2020, among others) with the intent of helping trainees identify quality-related issues and ultimately enhance the target language outcome.

Existing AVT research and publications that propose applicable quality assessment models focus mainly on subtitling (pre-recorded and live), as shall be outlined further

on. To the author's knowledge, no measurable model for dubbing has been put forward so far. The need to somehow measure quality levels in translation for dubbing arises in academic and in-house training contexts where the translated work of students and trainees requires evaluation. Likewise, localization companies that demand an 'entry test', before recruiting novice dubbing translators and adaptors require some form of measure to determine whether the newly rewritten target language text would 'work' in the dubbing studios. Sometimes, companies also demand trainers to provide them with trainee ranking following in-house training. The terms 'translation' and 'translators', in this context, are used in their broader meaning and also refer to the adaptation phase, that is, the preparation of a dialogue list for lip-synch dubbing purposes in a new target language.

The ideal scenario is one where available dubbing actors (together with a dubbing director or language department representative) can try out the translated test samples in the studios, perhaps during or in between recording sessions. This, however, is not always feasible, especially when several candidates need to be evaluated. Moreover, screening test video samples in the studios may imply interrupting the business workflow.

In any case, companies rarely entrust novice dubbing translators with a new commission to be broadcast or made available on any given platform before their skills have been tried and tested. Companies are also after time and cost-effective dubbing scripts. Time spent having to adjust the target text prior to or during the recording process often implies extra costs, whether this means paying actors for the extra time required to complete the recording of the scheduled takes; summoning them back for another session to re-record dialogue lines, or having in-house language editors or dubbing directors review and rewrite the target texts. The translated scripts are primarily functional tools integrated into the dubbing workflow and used by several professional agents (Spiteri Miggiani, 2019: 14). From an industry viewpoint, an additional quality factor is, therefore, the ability of the target language script to ensure the smooth running of the end-to-end dubbing workflow with minimal disruptions, as these would slow down the entire process and incur extra costs. Avoiding this is sometimes prioritized over the quality of the translation itself. Therefore, script time and cost-effectiveness can be considered a quality standard.

2. Quantifying quality

This paper attempts to provide a measurable model for the quality of a dubbing script taken both as a working tool used by practitioners as well as an end product that ultimately needs to be enjoyed by viewers. Therefore, it will be based on two levels of quality: the practical level from a workflow/client perspective and the reception level from an end viewer perspective. It will be referred to as the Textual Parameters (TP) model and is intended for two purposes: (1) trainee evaluation and ranking; (2) feedback and skills development. Variants of the same model are provided to suit both purposes. Whether quality can be quantified in the first place can be debatable, of course, since it necessarily implies a certain degree of subjectivity.

There have been similar initiatives to address the issue of quality in interlingual subtitling (Nikolić, 2021, Pedersen, 2017; Robert & Remael, 2016), intralingual live subtitling (Romero-Fresco & Martínez Pérez, 2015) and interlingual live subtitling (Robert & Remael, 2017; Romero-Fresco & Pöchhacker, 2017). Among the existing models, the FAR model (Pedersen, 2017) proposes an error-based assessment method based on: Functional equivalence (semantics, style), Acceptability (grammar, spelling, idiomaticity), and Readability (segmentation and spotting, reading speed, line length punctuation, use of italics). It is, therefore, a product-oriented model that focuses on the viewer's perspective and assigns a penalty point to each error to achieve a final score. Pedersen (2017: 213-214) also highlights the importance of in-house guidelines as evaluation tools, even though these may vary and are not easy to retrieve. Errors in

the FAR model are rated as minor, standard, or serious. This was inspired by the NER model (Romero-Fresco & Martínez Pérez, 2015) applied to intralingual live subtitling. The NER model is based on error-detection, too, specifically: Number of words, Editing errors, and Recognition errors, while the NTR model (Romero-Fresco & Pöchhacker, 2017) applied to the interlingual live subtitling is based on the Number of words, Translation errors, and Recognition errors. Recognition errors do not depend on the subtitlers but on the speech recognition software used. The NTR model, too, relates errors to content or style and classifies them as minor, major or critical. A formula is then applied to draw a subtitle accuracy rate, and 98% is taken as a benchmark for 'acceptable' subtitling (Romero-Fresco & Pöchhacker, 2017).

Künzli ropes professional subtitlers to provide their input on quality parameters and quality assurance measures (Kunzli, 2017). This has led to the CIA model of interlingual subtitling quality based on Correspondence (between source product and TL subtitles), Intelligibility and Authenticity, each one having its own subset of parameters (Künzli, 2021). The assumption is that the achievement of these three quality dimensions leads to a flowing subtitling experience. For its didactic application, Künzli suggests assigning a different number of maximum points to each of the three quality dimensions, subtracting a penalty for each error, and then converting the final score according to the grading system in the specific country. Künzli stresses the fact that, unlike most models, his approach focuses on the point of view of professional subtitlers rather than that of researchers (Künzli, 2021).

The dubbing model proposed in this paper tries to merge a practitioner and scholarly perspective in an attempt to provide a quality assessment method that can be applied and have a practical use both in an academic and professional context.

Despite there being no available and divulgated measurable model for dubbing, error-based taxonomies exist in specific academic institutions where students are trained in translation for dubbing. These can offer useful insights. A case in point is Frederic Chaume's error taxonomy (Baremo de corrección para la asignatura de Doblaje, UJI intranet) designed for undergraduate students at Universitat Jaume I, as well as Juan José Martínez Sierra's assessment rubric applied at the Universidad de Valencia. Chaume's taxonomy focuses on language errors (grammar, punctuation), typos, and translation errors (mistranslations, missing content, literal translation, awkward translations) and adopts a major/minor binary error degree, while Martinez Sierra's taxonomy focuses on technical aspects such as synchronization-related errors, character allocation, format, time codes, dubbing symbols and take segmentation. In the latter, a points system is used whereby a range of 0.25 to 1.5 points are deducted, with a maximum error cap applied to repeated typos and script formatting errors. The two taxonomies combined to cover the translation and adaptation processes in dubbing.

3. Dubbing Quality Standards: a revisited taxonomy and related errors in quality assessment

Chaume (2007) identifies six main quality standards in a dubbed product: 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. Other authors have discussed similar dubbing quality standards in general, among these Ávila (1997), Whitman-Linsen (1992), Chaves (2000), Chaume (2012, 2020), Spiteri Miggiani (2019, 2021a, 2021b), while some others have focused on specific parameters, such as voice suitability or character synchrony (Bosseaux, 2015; Martínez Sierra, 2008a), on the prosodic features of dubbed speech (Sánchez-Mompeán, 2020) or natural-sounding dialogue (Pavesi 1996, 2016; Romero Fresco 2006; Baños Piñero, 2009; Baños-Piñero & Chaume, 2009).

The TP model proposed in this paper is based on a revisited taxonomy of dubbing quality standards further adapted from Chaume (2007). This revisited taxonomy — discussed in detail by Spiteri Miggiani (2021a, 2021b) — divides the quality parameters into two categories: textual and non-textual.

Textual quality parameters:
(1) Adequate lip synchronization;
(2) Natural-sounding dialogue;
(3) Cohesion between dubbed dialogue and visuals;
(4) Fidelity to source text;
(5) Agreeable phonaesthetics;

Non-textual quality parameters:
(1) Suitable voice selection;
(2) Convincing voice performance;
(3) Natural-sounding intonation;
(4) Appropriate sound quality.

The TP model focuses on the translation phase in the dubbing workflow and is therefore based on the textual quality parameters only (See Table 4). These are mostly the concern of dubbing translators, while the parameters in the second category are mainly the responsibility of other professionals. This subdivision simply serves a practical purpose. The two categories are interdependent, and in actual fact, no one parameter can stand on its own. That said, ironically, one parameter could possibly be enough to produce an undesirable effect and consequently break the suspension of disbelief (Romero-Fresco, 2006) aimed for in a dubbed product. For this reason, the TP model does not recommend an acceptable or ideal percentage score that should supposedly be attained by trainees.

Acceptable lip-synch is taken in its broader professional practice, meaning, that is, corresponding timing, tempo, and lip movements. Timing refers to the matching duration of speech utterances and pauses, otherwise referred to in academia as isochrony (Chaume, 2012). Tempo refers to the speech delivery rate which has an impact on the mouth flap frequency, otherwise referred to as rhythmic synchrony (Spiteri Miggiani, 2019, 2021a). It is important to note that matching timing does not necessarily imply matching tempo. A dialogue line can be of the same length and duration as the original line without necessarily featuring all the necessary internal 'beats' and mouth flap movements. Lip movements refer to the mouth and lip articulatory movements, otherwise referred to as phonetic synch (Chaume, 2012). The movements that usually require visual correspondence in the dubbed version are generally bilabial consonants (/p/,/b/,/m/), labiodental consonants/fricatives (/f/,/v/) and lip rounded vowels or semi-vowels (/o/,/u/,/w/) (Chaume, 2012; Spiteri Miggiani, 2019; 2021b).

Lack of lip-synch, or dischrony (Fodor, 1976), where this is considered necessary (mainly depending on the visuals) will be taken as an error in the TP model. The quality assessor needs to identify the instances where this is imperative. Matching the mouth and lip movements is particularly important in close-up shots, while long shots or off-screen utterances allow more freedom. That said, respecting the same duration and rhythm can still be of utmost importance in off-screen speech to help the actor remain in rhythmic symbiosis (Spiteri Miggiani, 2019) with the text. Besides, when the rhythm band software is used, this does not allow any deviation from the set timing and tempo. Often, lack of lip-synch is what breaks the silent pact with the viewers and consequently the suspension of disbelief, making the dubbed product less credible in the eyes of the audience. Moreover, it can lead to miscomprehension

because of the modified auditory perception caused by the visual conflict (Möttönen and Sams, 2008).

Natural-sounding dialogue also referred to as a credible or realistic oral register (Chaume, 2007, 2012; Martínez Sierra, 2008a), is another parameter that can be associated with an error category: a language-related one as opposed to the previous one of a more technical nature. The degree of 'naturalness' (Romero Fresco, 2006) that is attainable is limited, and a certain extent of prefabricated orality (Baños-Piñero & Chaume, 2009) is expected. This is unconsciously accepted by viewers as long as a certain tolerance threshold is respected. When translators prioritize lip synchronization, this often comes at the cost of naturalness. It could also possibly account for the so-called dubbese register (Pavesi, 1996) in the dubbed target language output. Typical dialogue writing workflows and their impact on the translation cognitive processes may also account for unnatural dialogue and an overabundant presence of dubbese (Spiteri Miggiani, 2019: 85-94).

A certain extent of dubbese goes unnoticed thanks to the tolerance threshold mentioned earlier. Being a register in its own right (Marzà & Chaume, 2009: 36), dubbese does not necessarily have to be classified as an error as such, though it can be taken into account when evaluating the degree of naturalness in the dialogue. The ideal scenario is one where a balance between viewer habituation and their acceptance threshold is achieved, as well as a balance between consolidated formulaic expressions and overly spontaneous expressions. An effort to avoid dubbese at all costs may sometimes lead to over domestication (Spiteri Miggiani, 2021a).

Natural-sounding dialogue as a parameter takes suitable language style and registers variety as a given. The latter, once again, is subject to dubbing norms that usually imply a certain degree of standardization (Baños Piñero, 2009; Brincat, 2015; Martínez Sierra, 2008a). For the sake of the assessment model proposed —which also caters for feedback apart from the evaluation of an end product— correct language and grammar (taken as a given in any end product) will also be incorporated into the natural dialogue quality parameter.

Correspondence between the uttered target language speech and visuals ensuring semiotic cohesion is also an important parameter. Translators cannot manipulate the images and are therefore obliged to mold the text into the visual 'stencil'. This also implies having the newly recorded words correspond temporarily with specific body language (including facial gestures) to ensure semiotic cohesion. The latter is also referred to as kinesic synchrony (Chaume, 2012). Despite being classified as a type of synchrony, a distinction is being made between the synchronies that focus on verbal-to-visual technical cohesion (timing, tempo, lip movements) and verbal-to-visual non-technical cohesion (body language), where there is a semantic dependence of the translation on the meaning conventionally implied by specific gestures or facial expressions. When the new verbal codes and signs do not correspond to the visual, non-verbal codes and signs (Delabastita, 1989; Chaume, 2004) on a semantic and semiotic level, this can be considered an error in quality assessment.

Fidelity to the source text is understood as faithfulness to the intended effect, form, content, or function of the audiovisual text as a whole, thus prioritizing a homologous approach (Nord, 2005: 81; Chaume, 2016). Quality assessment errors can therefore comprise any form of significant loss or text manipulation that has an impact on viewer perception in general. Errors can also be more straightforward, such as missing content that is highly relevant to the plot or mistranslations that convey another meaning. Having said that, translators are often subject to target culture restrictions (or even censorship) deriving from local territory policies, norms, or ideologies (Spiteri Miggiani, 2019: 176-180). These can vary according to the medium, target audience and broadcasting time slots. They may therefore affect the fidelity parameter, though in such cases, lack of fidelity cannot be considered a quality

assessment error. On the contrary, the error would lie in not following the client's guidelines and target culture policies despite perhaps leading translators to stray away from the source text.

Phonaesthetics refer to the avoidance of cacophony in the dialogue, as well as tortuous or long winding dialogue lines that are not necessarily 'unnatural' but can distract viewers and decrease their level of engagement. This may be particularly relevant in emotionally-engaging monologues or dialogue exchanges where the way the dialogue 'sounds' may at times be more important than achieving exact equivalence. Detectable errors can include the overabundant use of /s/ in the same dialogue line since it creates a possibly annoying hissing sound; consonant clusters (that may also be difficult to articulate for actors); awkward sentence structures, and unintentional rhyme or repetitions. These features would not necessarily be grammatically incorrect: they simply may not be pleasant-sounding. Phonaesthetics can also have an impact on script functionality when the target language includes utterances that are difficult to pronounce or articulate for the actors.

The main difference between this revisited taxonomy of quality parameters and the previous most commonly applied one (Chaume, 2007) lies in the subdivision into two categories to distinguish textual from non-textual parameters. As mentioned earlier, this division simply serves a practical purpose, especially in research. In this context, isolating the textual parameters enables a quality assessment that focuses on the dubbing script and the translation process while still associating types of errors to a set of quality standards. Moreover, phonaesthetics is a new entry in the textual category, and the lip-synchronization parameter is intended as comprising additional synchrony: tempo or rhythmic synchrony. The non-textual category, on the other hand, features two additional parameters: natural-sounding intonation (Sánchez Mompeán, 2020) and voice selection relating to voice variety and voice qualities (Fodor, 1976; Whitman Linsen, 1992; Martinez Sierra, 2008a; Bosseaux, 2015). The sound quality parameter has been expanded to include further specifications. Instead of relating only to clear sound and volume, avoidance of undesired noises and voice replacement errors, it also lays emphasis on adequate post-synchronization sound mix and editing, as well as due attention is given to background noise. The latter provides depth and realism to the soundtrack and can therefore enhance the credibility of the dubbed product. The sound mix and editing run across the entire product and actually glue together all the dubbing 'ingredients'. They have the power to break the suspension of disbelief while undoing all the other achieved standards.

The aims across all parameters focus mainly on the achievement of credibility and suspension of disbelief, therefore ultimately approaching the end product from a viewer's perspective in terms of reception and perception. Ensuring comprehension is also an important aim that seems to be a common thread across most parameters. Indeed, errors associated with most of the quality standards may have an impact on comprehension levels. For instance, both poor articulation (voice performance) and lip-synchronization can hinder comprehension.

3.1 A workflow-oriented macro parameter

Since the parameters outlined so far and adopted as a basis for the proposed model are end product-oriented, it is necessary to include another parameter for quality assessment purposes. As mentioned earlier, from a client perspective, a script's time and cost-effectiveness are fundamental. Localization companies are as concerned about the smooth running of the dubbing process and workflow as much as they are about the quality of the end product. In an in-house training context, companies would most likely highlight errors that may not be considered as such by translators or researchers in the field, who are more focused on the quality of the end product.

This workflow-oriented quality approach accounts for the inclusion of a sixth quality standard: a macro parameter that encompasses a variety of functionality-related aspects. Within the company walls, a dubbing script is primarily a working tool that needs to be functional. Related errors may not be classified as such in a didactic context, but they tend to slow down the dubbing workflow, sometimes incurring extra costs. For example, incorrect character allocation (writing the wrong character name next to a given dialogue line) may imply having to summon back an actor and schedule an extra recording slot.

The script functionality parameter can include the following:

(1) Compliance with company guidelines and specifications in relation to local, territorial policies, including blacklisted expressions e.g., vulgar language, expressions related to food or mental health disorders (Spiteri Miggiani, 2019); inclusive and sensitive language; commercial and legal issues; censorship; foreign language; accents; songs; on-screen graphics, and so on. Non-compliance or lack of awareness generally implies revision carried out by in-house dubbing agents.

(2) Compliance with standard conventions or in-house specifications in relation to dubbing notations (Spiteri Miggiani, 2019), otherwise known as dubbing symbols (Chaume, 2012), tempo markers or pauses (Spiteri Miggiani, 2019); character allocation; loop segmentation (when applicable); dialogue segmentation, background dialogue; time codes; typos and spelling; script formatting and layout; dialogue segmentation; punctuation, and consistency within the same script or across scripts belonging to the same serial production, therefore non-compliance with glossary sheet requirements, inconsistencies in names or nicknames, forms of address and terminology (See Table 6). These are considered functional aspects because they mostly have an impact on the actors' role and progress, or on the dubbing workflow in general, and not on the recorded output that viewers 'receive'. Therefore, they are more workflow-oriented than end-product-oriented. Interestingly, these are more pertinent to traditional dubbing workflow contexts as opposed to cloud dubbing or app-based workflows. In the latter case, many of these functional errors are avoided by default, thanks to the integrated assistive tools (Spiteri Miggiani, forthcoming).

In general, aiming to meet all quality standards across the entire dubbing script may not be a realistic goal. When dealing with the translation on a microlevel, compromises are sought constantly. One dialogue line may prioritize one parameter, e.g., lip synchronization, while the next dialogue line may prioritize fidelity to the text or naturalness at the cost of better phonetic synch. What matters ultimately is a harmonious balance and the general feel and outcome of the product when everything is put together. The model proposes a measurable quality assessment model for practical purposes and may also provide a standard benchmark based on chosen criteria when several translators are being tested in the same context.

4. The Textual Parameters Quality Assessment model (TP model)

The TP model is based on the classification of dubbing quality parameters outlined in Section 2 and focuses on the Textual Parameters. In addition to these, script functionality will also be accounted for, in other words, those aspects that may enhance or disrupt the workflow on a practical level and that are often adjusted or solved by the other agents involved. These would not necessarily have an impact on the end product but are considered a quality factor by localization companies. This tentative model, therefore, combines product-oriented standards with workflow-oriented standards and expectations.

The aim is to provide a score for a translated sample that could possibly also be used for translator ranking when this is necessary. This proposal is not intended as an exhaustive quality assessment method and cannot factor in all possible variables and details. It is simply an attempt to provide a tool that may be helpful to trainers.
The TP model is based on an error-detection system, the so-called 'errors' implying non-compliance with the textual parameters taken as an optimal scenario. A formula is applied to calculate a final percentage score to grade the translation.
Equal weight is assigned to all parameters, the main reason being that assigning a value to each one would be subjective. Not only is it subjective to the evaluator in question, but in general, academics, professional translators, and clients may not value parameters in the same way. The importance given to parameters can also vary according to the target culture (Künzli, 2021). For example, phonaesthetics, in Italy, are given much attention and are sometimes prioritized over other parameters. Moreover, the value of each parameter or associated error also depends very heavily on the production's priorities and specific context. Equal weighting also helps to simplify the model, thus enabling the possibility to include other variables instead (such as varying source texts and degree of difficulty).
The equal weighting system can, however, be considered a limitation. Therefore, despite giving an equal mathematical value to each parameter, the TP model can easily incorporate an error grading system. The errors encountered can be classified as minor or major (across quality parameters) by the evaluators themselves, and the score is calculated accordingly, as will be illustrated further on. Therefore, rather than assigning an absolute value to the parameters or associated errors, these are contextualized and graded on a case-by-case basis. This flexibility is required since several factors can influence the weighting assigned to errors, for instance, the variability of productions, clients, cultures, or specific scenes. Indeed, assigning a decontextualized absolute value to each quality standard may also be viewed as a limitation.
Another limitation of the model is that it is error-based and does not reward outstanding or 'better' solutions when the work of translators is compared. As explained further on, this approach is in line with the industry perspective that focuses on having scripts that 'work'. Companies will most likely complain about errors but will less likely reward translators for brilliant solutions, except perhaps by ensuring further translation commissions.
The TP model considers the possibility of a variety of text samples being assessed in the same context. The reason is that different source languages, and consequently different audiovisual samples may be assigned and distributed among the same group of translators/trainees who work with the same target language. A common criterion or quantifiable denominator across the samples is necessary to elicit a quality assessment score. The total number of words in each text sample is therefore taken as a base measurement. The higher the total number of words, the higher the probability of errors. Taking sample variety into account implies factoring in different degrees of difficulty of the source texts. The proposed formula, therefore, incorporates this as a possible variable.

4.1 The error-based formula

The quality assessment is based on an error detection method that relies on non-compliance with the following textual parameters.
(1) Adequate lip synchronisation;
(2) Natural-sounding dialogue;
(3) Cohesion between dubbed dialogue and visuals;
(4) Fidelity to source text;

(5) Agreeable phonaesthetics;
(6) Script functionality.

Since the parameters are considered to have equal mathematical weight, a first and simple basic option is to insert a generic error code [E] throughout the sample target text every time a so-called error is detected. At the end of the text sample review, the total number of errors is counted and used in the formula provided hereunder to calculate a final percentage score. In the case of a university context, the percentage scores obtained will need to be further adapted to the marking system adopted. A sample is provided in Table 1.

Basic formula:
S% = 100 – (E/W)*100

- S is the total percentage score indicating quality levels;
- E is the total number of errors;
- W is the total number of words in the source text sample.

Adapted dialogue	Percentage score
FARSHID (OFF-ON) Here it is. / The wonder of wonders. // I remember the first time I heard you play. [E] (gasps) [E] I'd never heard anything like it. Really, it was [E]… When I think about it, I want to cry. [E]/ I… [E] Do you see this tear running? [E] / Just there. Do you see?	Total no. of words: 57 (W) Total number of errors: 6 (E) S% = 100 – (E/W)*100 **89.47%** = 100 – (6/57)*100

Table 1: Text sample applying the basic formula

If the evaluators in question want to add a major/minor grading to the errors, they would need to determine the number of major errors and the number of minor errors and give weight to each type of error. A simple option to create a mathematical distinction between major and minor errors could be to count each major error as 3 errors and each minor error as 1. This can be customized according to preferences. Minor errors can encompass minor to average severity, and any errors that are considered insignificant can be flagged for feedback purposes but ignored when quantifying the total number of errors.
The formula would therefore be adapted as follows:

S% = 100 - [(Emaj*3+Emin)/W]*100

- S is the total percentage score indicating quality levels;
- Emaj is the number of major errors;
- Emin is the number of minor errors;
- 3 is the weight given to major errors (customizable);
- W is the total number of words in the source text sample.

The higher the number of major errors in a text sample, the higher the number of errors quantified in the formula, thus lowering the final score. Table 2 hypothesizes 4 major errors and 2 minor errors in the text sample.

Adapted dialogue	Percentage score
FARSHID (OFF-ON) Here it is. / The wonder of wonders. // I remember the first time I heard you play. [E] (gasps) [E] I'd never heard anything like it. Really, it was [E]… When I think about it, I want to cry. [E]/ I… [E] Do you see this tear running? [E] / Just there. Do you see?	Total no. of words: 57 (W) Total number of errors: 4 major + 2 minor S% = 100 – [(Emaj*3+Emin)/W]*100 **75.4%** = 100 **-** [(4x3 + 2)/57]*100

Table 2 Text sample applying basic formula with major/minor error grading

The formula can be further adapted to include a third variable: the degree of difficulty of the source text in the case of a variety of texts and/or source languages used in the same evaluation/ranking event. The formula would therefore be adapted as follows:

S% = 100 – [(Emaj*3 +Emin)*O/W]*100

- S is the total percentage score indicating quality levels;
- Emaj is the number of major errors;
- Emin is the number of minor errors;
- W is the total number of words;
- O is the error 'offset', a parameter which varies according to the degree of difficulty of the text; O will be taken as a number between 0.5 and 1, based on 3 degrees of source text difficulty: Low: O = 1, Medium: O = 0.75, High: O = 0.5.

The error offset is included to account for a situation where various text samples are being used in the same ranking event, and therefore some texts may be more difficult than others. The texts could be classified according to a low, medium, and a high degree of difficulty and the offset is meant to balance the score accordingly. When the same sample text (or the same degree of difficulty) is applied for all trainees, the basic formula can be resorted to.
If evaluators prefer to assign equal weight to all errors and opt out of the minor/major grading of errors, while incorporating the degree of difficulty variable, the formula can be simplified as follows:

S% = 100 – [(E*O)/W] * 100

A sample is provided in Table 3. This sample is hypothetically rated as having a high degree of difficulty while incorporating the minor/major grading of errors. The sample hypothesizes 4 major errors and 2 minor errors as in Table 2.

Adapted dialogue	Percentage score
FARSHID (OFF-ON) Here it is. / The wonder of wonders. // I remember the first time I heard you play. [E] (gasps) [E] I'd never heard anything like it. Really, it was [E]… When I think about it, I want to cry. [E]/ I… [E] Do you see this tear running? [E] / Just there. Do you see?	Total no. of words: 57 (W) Total number of errors: 4 major + 2 minor (Emaj + Emin) Level of difficulty: high 0.5 (O) S%=100 – [(Emaj*3+Emin)*O/W]* 100 **87.7% =** 100 - [(4*3+2) x0.5/57]*100

Table 3 Text sample applying both variables: high degree of difficulty and minor/major error grading to the formula

4.2 Determining the degree of difficulty of a text sample

Determining the degree of difficulty of a text sample may very well be subjective. Analyzing and identifying the challenges posed by a given source of audiovisual text (both visuals and dialogue) can help rate the text. A number of translation challenges can be taken into account, ranging from synchronization constraints — determined by specific types of shots or delivery speech rate — to linguistic transfer challenges determined by culture-bound elements, intertextuality, wordplay, specialized jargon, and several other issues. A broad subdivision of the encountered challenges into technical and non-technical constraints (Spiteri Miggiani, 2019: 30-32) — to determine an approximate number of issues — can help compare text samples to then categorize them as low, medium, or high degree of difficulty. Alternatively, the source text analysis model proposed by Chaume (2012: 170-177), in turn, based on Baker (2011), Hatim and Mason (1990, 1997) and Newmark (1998), can be a useful point of reference.

Another possible approach to the error detection system is to analyze the source audiovisual text in advance on a micro level to determine the specific challenges, add markers in the original written text for the evaluator's reference, and then compare the source text with the written target outcome. That said, this approach may possibly lead to a focus on the written text while it is imperative to analyze the target text as an audiovisual whole together with the visuals. The visual representation of phonemes or visemes (Fisher, 1968) may not necessarily correspond to what we expect when analyzing the written text alone. In other words, a certain extent of detachment from the written source text may be helpful. The TP model, therefore, focuses on a target text-oriented approach that aims to evaluate the target text in its own right (Toury, 2012), and, more importantly, in relation to the video material and not to the original written source text. This, in any case, does not exclude a subsequent comparison with the source dialogue to further pin down the issues and understand where they may derive from.

Another question may therefore arise. How can trainers identify errors, especially if the focus lies mostly on the target output? There are three feasible options:

(1) The trainer/reviewer could test the target language samples individually by reciting them against the video, just like any translator would do during the translation process;

(2) In a training context, a very productive way is to mute the audio in the room and have the trainees (with the use of earphones for the original audio) try out their texts, therefore simulating a dubbing recording session, while the rest of the trainees observe;

(3) For evaluation and ranking purposes, having the translators voice and record their own target texts amateurly is an easy and effective assessment tool and also contributes to the didactic experience because it helps to develop self-assessment skills. Trainees are compelled to look at their own recordings, self-evaluate the results, and adjust and re-record accordingly (Chaume, 2008; Spiteri Miggiani, 2019). Emulating the dubbing actors draws attention to the functional aspects of a dialogue script. Self-assessment is often one of the challenges encountered by translators, therefore, the parameters and related errors upon which this model is built can possibly be useful self-evaluation tools.

4.3 Error rubric and legenda of codes for a detailed feedback approach

As mentioned in Section 1, the TP model can either be used solely for evaluation and ranking purposes or it can be combined with a feedback and skills development objective. In the former case, since the quality parameters are assigned equal mathematical weight, a simple [E] will suffice to mark any detected error to then provide a final score. If, on the other hand, the model is used to provide feedback to trainees or companies, specific issues can be identified, addressed, and progress can be registered both on an individual level as well as on a broader level (if the main training objective is to enhance quality in a specific dubbing context).

To this end, this paper proposes an error rubric based on the set of textual quality standards outlined earlier. Two variants are provided depending on the degree of detail required throughout the evaluation and feedback process.

(1) TP assessment rubric, variant A: an error category and code are assigned to each quality parameter.

(2) TP assessment rubric, variant B: codes are also assigned to specific errors within each error category.

Using a variety of codes pertaining to the same quality parameter or error category does not cause any imbalance in the model's quantification method because each parameter is given equal numerical weight. Whether lip synchronization is flagged with a generic [E] or with an error-specific code, it makes no difference mathematically because all types of errors are calculated and used equally in the formula. TP assessment rubric variant A is provided in Table 4; a sample is provided in Table 5.

Quality parameter (Textual parameters)	Generic error code	Error category	Description
Adequate lip synchronisation	[S]	Synchronisation	Technical errors related to timing, tempo, lip articulatory movements
Natural-sounding language	[L]	Language	Lack of naturalness, incorrect grammar, unsuitable register or style, lack of flow & cohesion between dialogue exchanges, source calques, lack of clarity/comprehension
Semiotic cohesion	[VS]	Visuals & Sound	Lack of cohesion between words & visuals (such as body language) or between words & sound belonging to the original audio track (such as effects, noise, music or lyrics)
Fidelity to source text	[T]	Translation	Translation issues such as mistranslations, unnecessary loss, omission or addition, awkward rendering, inappropriate translation (undue non-inclusive, offensive or derogatory terms that are not functional to the plot or characterisation)
Phonaesthetics	[PH]	Phonaesthetics	Unpleasant sounding lines, cacophonic utterances, annoying repetition, unintended rhyme
Script functionality	[F]	Functionality	Issues with format, layout or dialogue segmentation, lack of consistency (within script or across scripts: names/nicknames, forms of address, terminology; non-compliance with glossary sheets), missing reactions or sounds, missing dialogue or background walla, wrong or missing dubbing notations or pause markers, misleading punctuation, orthography mistakes, wrong character allocation, missing or wrong time code, non-compliance with guidelines and policies, miscellaneous

Table 4 TP assessment rubric, variant A

Adapted dialogue	Percentage score
FARSHID (OFF-ON) Here it is. / The wonder of wonders. // I remember the first time I heard you play. [S] (gasps) [F] I'd never heard anything like it. Really, it was [S]… When I think about it, I want to cry. [S]/ I… [F] Do you see this tear running? [S] / Just there. Do you see?	Total no. of words: 57 (W) Total number of errors: 6 (E) S% = 100 – (E/W)*100 **89.5%** = 100 – (E/W)*100

Table 5 Text sample adopting model, variant A

Variant A can be further extended to TP assessment rubric, variant B, as illustrated in Table 6. A sample is provided in Table 7.

Quality parameter (Textual parameters)	Generic error code	Error category	Specific error code	Error specifics
Adequate lip synchronisation	[S]	Synchronisation	[…]	Too short
			[--]	Too long
			[R]	Rhythmic issues (mouth flaps mismatch)
			[L]	Labial consonants mismatch
			[V]	Vowels or semivowels mismatch
Natural-sounding language	[L]	Language	[GR]	Incorrect grammar
			[SC]	Source calque
			[REG]	Unsuitable register
			[COMP]	Lack of clarity & comprehension
			[NAT]	Lack of naturalness
			[FLOW]	Lack of flow & cohesion between dialogue exchanges
Semiotic cohesion	[VS]	Visuals & Sound	[VIS]	Lack of cohesion between words & visuals (such as body language)
			[SND]	Lack of cohesion between words & sound belonging to the original audio track (music & effects, lyrics, noise)
Fidelity to source text	[T]	Translation	[MIS]	Mistranslation
			[OM]	Unnecessary omission
			[ADD]	Unnecessary addition
			[LOSS]	Unnecessary loss (semantic)
			[AWK]	Awkward translation or rendering
			[IMP]	Improper translation (such as undue non-inclusive, offensive or derogatory terms that are not functional to the plot or characterisation)
Phonaesthetics	[PH]	Phonaesthetics	[CAC]	Cacophonic utterances
			[REP]	Annoying repetition
			[RHY]	Unintended rhyme
Script functionality	[F]	Functionality	[CON]	Lack of consistency (non-compliance with glossary sheets; inconsistent use of names/nicknames, forms of address & terminology within same script or across serial production scripts)
			[REAC]	Missing or wrong reaction
			[NOT]	Missing or wrong notation
			[/]	Missing pause marker
			[FOR]	Layout or format issues
			[DS]	Unsuitable dialogue segmentation
			[OR]	Orthography mistakes
			[CH]	Wrong character allocation
			[D-?]	Missing dialogue
			[B-?]	Missing or inadequate background walla
			[PUN]	Misleading punctuation
			[TC]	Missing or wrong time code
			[G/P]	Non-compliance with guidelines & policies
			[PRON]	Tricky articulation or pronunciation
			[MISC]	Miscellaneous

Table 6 TP assessment rubric, variant B

Adapted dialogue	Percentage score
FARSHID (OFF-ON) Here it is. / The wonder of wonders. // I remember the first time I heard you play. [...] (gasps) [/] I'd never heard anything like it. Really, it was [...]… When I think about it, I want to cry. [L]/ I… [/] Do you see this tear running? [...] / Just there. Do you see?	Total no. of words: 57 (W) Total number of errors: 6 (E) **S%** = 100 – (E/W)*100 **89.5%** = 100 – (6/57)*100

Table 7 Text sample adopting model, variant B

5. Conclusions

The application of this tentative model has so far been limited to the author's academic and in-house experience as a trainer. The need to come up with a measurable model emerged during an in-house training at a localization company, so it was initially designed for practical purposes in a real case scenario. The training program in question also included the review of several assigned adaptation tasks. A coded feedback system was necessary for a view of the numerous trainees. At the same time, a scoring system was required to evaluate their work based on consistent criteria across the text samples. Although they formed part of the same cohort, the trainees worked with a variety of language pairs. Therefore, a fair marking and ranking system across different texts had to be set up. The trainees had the possibility to choose the texts according to source language and genre, as well as their level of confidence or preparation. Consequently, some texts were intentionally easier than others. Therefore, those who chose the more challenging texts were granted an error offset. An acceptable or 'good' benchmark score was not determined. What usually matters to companies is a ranking system that provides a tentative order in which to recruit translators who are 'ready' for the job, the 'real' test being the actual professional practice. The above scenario provided the foundations of the TP model.

The design of the model was also based on a similar need to grade students in postgraduate multilingual university classrooms. Training in an academic context reveals the need for common, clear, and possibly measurable criteria accompanied by exhaustive feedback. In this context, where progress can be monitored, the scoring system can also help to highlight skills development across a span of time. The model can also act as a self-assessment tool for students, even by simply applying the quality standards taxonomy and associated errors without necessarily rating their own work.

Despite its limited and subjective application so far, the TP model is being shared and proposed in this paper to trainers and stakeholders who may have encountered similar needs. The TP model requires further application and validation. In this context, the proposal is simply a point of departure intended to entice further investigation, possibly based on a collaborative effort and ideally involving the industry. Further research is also required to build an exhaustive and practical taxonomy of errors related to translation for dubbing. The taxonomy of errors does not necessarily have to correspond with a taxonomy of quality standards (which is the case of the TP model), though classifying errors can certainly help to enhance the existing classifications of quality standards in dubbing.

Bibliographic references

Avila, A. (1997). El Doblaje. Madrid: Cátedra.

Baker, M. (2011). In Other Words: a Coursebook on Translation. London & New York: Routledge.

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. [Doctoral dissertation]. Universidad de Granada, Granada.

Baños-Piñero, R., & Chaume, F. (2009). Prefabricated Orality: A challenge in Audiovisual Translation. inTralinea.

Bartrina, F., & Espasa, E. (2003). Traducción de textos audiovisuales. In M. Gonzáles Davies (Ed.), Secuencias. Tareas para el aprendizaje interactivo de la traducción especializada (pp. 19-38). Barcelona: Octaedro.

Bartrina, F., & Espasa, E. (2005). Audiovisual translation. In M. Tennent (Ed.), Training for the New Millennium (pp. 83-100). Amsterdam/Philadelphia: John Benjamins.

Bolaños García-Escribano, A. (2020). The didactics of Audiovisual Translation in the Age of Cloud technologies. [Doctoral dissertation]. University College London, London.
Bosseaux, C. (2015). Dubbing, Film and Performance: Uncanny Encounters. Oxford: Peter Lang.
Brincat, J. (2015). Morphological and semantic simplification in dubbing techniques: Translating the dialogue of the British films Ae Fond Kiss and The Queen. In M. Pavesi, M. Formentelli, & E. Ghia (Eds.), The languages of dubbing (pp. 197-216). Bern: Peter Lang.
Cerezo-Merchán, B. (2012). La didáctica de la traducción audiovisual en España: un estudio de caso empírico-descriptivo. [Doctoral dissertation]. Universitat Jaume I, Castelló.
Cerezo-Merchán, B. (2018). Audiovisual Translator Training. In L. Pérez-González (Ed.), The Routledge Handbook of Audiovisual Translation (pp. 468-482). London: Routledge.
Cerezo Merchán, B., & De Higes Andino, I. (2018). Using evaluation criteria and rubrics as learning tools in subtitling for the D/deaf and the hard of hearing. The Interpreter and Translator Trainer, 12(1), 68-88.
Chaume, F. (2003). Teaching audiovisual translation: Some methodological proposals. In L. Pérez González (Ed.), Speaking in Tongues: Language across Contexts and Users (pp. 271-302). Valencia: Publicacions de la Universitat de Valencia.
Chaume, F. (2004). Film studies and translation studies: Two disciplines at stake in audiovisual translation. Meta, 49(1), 12–24. https://doi.org/10.7202/009016ar.
Chaume, F. (2007). Quality standards in dubbing: a proposal. TradTerm, 13, 71-89.
Chaume, F. (2008). Teaching synchronisation in a dubbing course: Some didactic proposals. In J. Díaz Cintas (Ed.), The Didactics of Audiovisual Translation (pp.129-140). Philadephia/Amsterdam: Benjamins.
Chaume, F. (2012). Audiovisual Translation: Dubbing. Manchester: St. Jerome Publishing.
Chaume, F. (2016). Dubbing a TV Drama Series. The case of The West Wing. inTralinea.
Chaume, F. (2020). Dubbing. In Ł. Bogucki and Mikołaj D. (Eds.), The Palgrave handbook of Audiovisual Translation and Media accessibility (103-132). London: Palgrave Macmillan,
Chaves, M.J. (2000). La Traducción Cinematográfica: El Doblaje. Huelva: Publicaciones de la Universidad de Huelva.
Delabastita, D. (1989). Translation and mass-communication: Film and TV translation as evidence of cultural dynamics. Babel, 35(4), 193-218.
Díaz Cintas, J. (Ed.), (2008). The Didactics of Audiovisual Translation. Amsterdam/Philadelphia: Benjamins.
Fodor, I. (1976). Film Dubbing. Phonetic, Semiotic, Esthetic and Psychological Aspects. Hamburg: Buske Verlag.
Fisher, C. G. (1968). Confusions among visually perceived consonants. Journal of Speech, Language and Hearing Research, 11(4), 796.
Hatim, B., & Mason, I. (1990). Discourse and the Translator. London: Longman.
Hatim, B., & Mason, I. (1997). The Translator and a Communicator. London: Routledge.
Huertas Barros, E., Vandepitte, S., & Iglesias-Fernández, E. (2019). Quality Assurance and Assessment Practices in Translation and Interpreting. IGI Global Publisher.
Huertas Barros, E. & Vine, J. (2019). New Perspectives on Assessment in Translator Education. London: Routledge.

Hurtado-Albir, A. (2007). Competence-Based Curriculum Design for Training Translators. The Interpreter and Translator Trainer, 1(2), 163-195.
Hurtado-Albir, A. (2015). The Acquisition of Translation Competence. Competences, Tasks, and Assessment in Translator Training. Meta, 60(2), 256-280.
Hurtado Albir, A. (2017). Translation and translation competence. A. Hurtado Albir (Ed.), Researching Translation Competence by PACTE Group, 3-34. Amsterdam/Philadelphia: John Benjamins,.
Kelly, D. (2005). A Handbook for Translator Trainers. Manchester: St. Jerome.
Künzli, A. (2006). Translation revision: A study of the performance of ten professional translators revising a technical text. In M. Gotti and S. Sarcevic (Eds.), Insights into Specialised Translation, Bern & Frankfurt: Peter Lang, 193-211.
Künzli, A. (2017). Die Untertitelung — Von der Produktion zur Rezeption. Berlin: Frank & Timme.
Künzli, A. (2021). From inconspicuous to flow — the CIA model of subtitle quality. Perspectives, 29(3), 326-338.
Martínez Sierra, J.J. (2008a). Humor y Traducción: Los Simpson Cruzan la Frontera. Castelló: Universitat Jaume I.
Martínez Sierra, J.J. (2008b). Hacia una enseñanza más completa de la traducción audiovisual. Tonos, 16/1.
Marzà, A., & Chaume, F. (2009). The language of dubbing: Present facts and future perspectives. In M. Pavesi and M. Freddi (Eds.), Analysing Audiovisual Dialogue. Linguistic and Translational Insights (pp. 31-39). Bologna: Clueb.
Mossop, B. (2001). Revising and editing for translators. St Jerome: Manchester.
Mossop, B. (2007). Empirical studies of revision: what we know and need to know. The Journal of Specialised Translation, 8, 5-20.
Mossop, B. (2011). Revision. In Y. Gambier and L. Van Doorslaer (Eds.), Handbook of Translation Studies, 2 (pp. 135-139). Amsterdam/Philadelphia: John Benjamins.
Mossop, B. (2014). Revising and editing for translators. London/New York: Routledge.
Mossop, B. (2016). Translation Revision. Book of Abstracts, 8th EST Congress Aarhus, 15-17 September 2016. Translation Studies: Moving Boundaries, Aarhus University.
Möttönen, R., & Sams, M. (2008). Audiovisual interaction. In D. Havelock, S. Kuwano & M. Vorlander (Eds.), Handbook of Signal Processing in Acoustics, New York: Springer Verlag. 731-745.
Newmark, P. (1998). A Textbook of Translation. New York, London, Sydney, Tokyo: Prentice Hall International.
Nikolić, K. (2021). Quality Control of Subtitles: A Study of Subtitlers, Proofreaders, Quality Controllers, LSPs and Broadcasters/Streaming Services. Journal of Audiovisual Translation, 4(3), 66-88.
Nord, C. (2005). Text Analysis in Translation. Theory, Methodology, and Didactic Application of a Model for Translation-Oriented Text Analysis. Amsterdam/New York: Rodopi.
Pavesi, M. (1996). L'allocuzione nel doppiaggio dall'inglese all'italiano. In R. M. Bollettieri Bosinelli, & Heiss, C. (Eds.), Traduzione multimediale per il cinema, la televisione e la scena, Bologna: Clueb. 117-130.
Pavesi, M. (2016). Formulaicity in and across film dialogue: Clefts as translational routines. Across Languages and Cultures, 17, 99-121.
Pedersen, J. (2017). The FAR model: assessing quality in interlingual subtitling. The Journal of Specialised Translation, 28, 210-229.
Robert, I., & Remael, A. (2016). Quality control in the subtitling industry: An exploratory survey study. Meta, 61(3), 578-605.

Robert, I., & Remael, A. (2017). Assessing quality in live interlingual subtitling: A new challenge. Linguistica Antverpiensia NS: Themes in Translation Studies, 16, 149-167.
Romero-Fresco, P. (2006). The Spanish Dubbese: A case of (un)idiomatic Friends. The Journal of Specialised Translation, 6, 168-195.
Romero-Fresco, P., & Martínez Pérez, J. (2015). Accuracy rate in subtitling: The NER model. In R. Baños Piñero and J. Díaz Cintas (Eds.), Audiovisual Translation in a Global Context: Mapping an Ever-Changing Landscape, Basingstoke: Palgrave Macmillan. 28-50.
Romero-Fresco, P., & Pöchhacker, F. (2017). Quality assessment in live interlingual subtitling: A new challenge. Linguistica Antverpiensia NS: Themes in Translation Studies, 16, 149-167.
Sánchez-Mompeán, S. (2020). The Prosody of Dubbed Speech. Beyond the Character's Words. Basingstoke: Palgrave Macmillan.
Spiteri Miggiani, G. (2019). Dialogue Writing For Dubbing. An Insider's Perspective. London: Palgrave Macmillan.
Spiteri Miggiani, G. (2021a). English-language dubbing: challenges and quality standards of an emerging localisation trend. The Journal of Specialised Translation, 36, 2-25.
Spiteri Miggiani, G. (2021b). Exploring applied strategies for English-language dubbing. Journal of Audiovisual Translation, 4(1), 137-156.
Spiteri Miggiani, G. (forthcoming). Cloud studios and scripts: evolving workspaces and workflows in dubbing.
Tennent, M. (2005). Training for the New Millennium. Amsterdam/Philadelphia: Benjamins.
Toury, G. (2012). Descriptive Translation Studies - and beyond. Amsterdam/Philadephia: John Benjamins.
Whitman-Linsen, C. (1992). Through the Dubbing Glass. Frankfurt am Main: Peter Lang.
Zabalbeascoa, P., Santamaria Guinot, L., & Chaume, F. (2005). La traducción audiovisual: investigación, enseñanza y profesión. Granada: Comares.

Words: 7847
Characters: 52 714 (29,29 standard pages)

Giselle Spiteri Miggiani
Department of Translation
Terminology and Interpreting Studies
University of Malta
Malta
mspit12@um.edu.mt