Subtitling for Social and Language Minorities: Subtitling of Oral Errors and Dialectal Features in the Case of Minoritised Languages

Mercedes Martínez-Lorenzo
University of Vigo, Spain

Abstract

Speakers of any (minoritised or majority) language sometimes make language mistakes. Bilingual speakers may use a hybrid language, mixing languages within a sentence or even within a word, especially when they are formally similar, as Spanish and Galician are. For minoritised languages, language errors may contribute to a negative perception towards the minoritised language. The Galician public broadcaster Televisión de Galicia (TVG) has received criticism for not being a high-quality language model, permitting the intrusion of language mistakes in its content. From an exclusively linguistic viewpoint, these errors should be corrected in subtitling. Conversely, subtitling guides and target users favour a verbatim rendition of the audio, in which oral language mistakes should not be corrected. Dialectal features, even if they are not considered errors, are non-standard language. This paper aims at answering the question of “to correct or not to correct” oral errors and dialectal features in the case of minoritised languages. It presents the most relevant data from a literature review, and an analysis of subtitling guidelines & standards and of current practices at TVG. These results have yielded an original protocol for the correction or reproduction of oral errors according to speech control, target audience and broadcast genre, the effect of a mistake, and the type of language error (vocabulary vs. grammar).

Key words: dialectal features, Galician, hybridism, inclusive subtitling, error correction, error reproduction, minoritised languages, mistakes, oral error, purism.
1. Introduction

This paper studies the viability of linking the areas of media accessibility (MA), with special emphasis on inclusive subtitling, and the promotion of minoritised or minority languages, when encountering oral language errors and dialectal features in the audio of audiovisual content. In order to subtitle for both social and language minorities, should oral language errors and dialectal features be subtitled so as to comply with users’ preference for more literal subtitles and minimise the vulnerability of subtitling? Or should errors and dialects be corrected in inclusive subtitles in favour of learning the standard and to encourage a more positive perception of the language?

The starting point is to analyse a given minority/minoritised language. While a language may be both described as a minority language and a minoritised language at the same time, these two concepts are often wrongly used as synonyms: minority relates to a lower quantity of speakers of a language with respect to another language within a territory (Council of Europe, 1992); minoritised alludes to lower quality, that is, inferior use and perception (Díaz-Fouces, 2005, p. 96). Galician, a minoritised language spoken in Spain, is taken in this article as a case study and analysed quantitatively (knowledge and active use) and qualitatively (speakers’ perception towards the language before and after promotion measures). Its status may resonate with other languages and the points of analysis and further steps could be applied to them.

The following step is to examine how (inclusive) subtitling guidelines (or recommendations) and subtitling standards (ratified by standardisation bodies) (Matamala & Orero, 2018, p. 141) and language experts answer the question of whether or not to reproduce oral language errors and dialectal features. The key instructions that best adapt to Galician, and possibly to other languages, have been extracted from a literature review of 16 subtitling guides (though not all of them included indications on linguistic errors and dialects), academic texts on Galician Linguistics and an empirical observation of 13 hours of subtitling by the Galician public broadcaster TVG (Martínez-Lorenzo, 2019, 2020). This assessment of these sources has led to a protocol for the reproduction of dialecticisms and for the correction or reproduction of oral language errors according to: the effect of a mistake, the type of broadcast (speech control/spontaneity, genre, and target audience), and the type of language error (vocabulary, grammar) (section 4).

The expression “inclusive subtitles/subtitling” is put forward here as an updated term for “subtitling for the D/deaf and hard of hearing” (SDH or SDHH) and “closed captions” (CC), in line with “enriched subtitling” as proposed by Neves (2018, p. 83–84). Because “enriched” mainly connotes financial wealth or nutrient-added food in Spanish and Galician (“enriquecer”, RAE, n.d.; RAG, n.d.), the semantics of the word has culturally shifted in these languages. Initially, the adjectives “integrated” and “integrative” were considered; they are used in the context of theatre accessibility by Fryer
Subtitling for Social and Language Minorities: Subtitling of Oral Errors and Dialectal Features in the Case of Minoritised Languages

(“integrated access”, 2018) and in the Spanish Education Act LOGSE (1990), which ruled students with special education needs (such as persons with disabilities) were to be integrated in and study at regular schools. However, “integration” is understood as providing the individual with special measures to adapt them to a fixed environment and is negatively perceived. Accessibility in this sense would still only pertain to people with disabilities, which corresponds to a particularist approach to access (Greco, 2019, p. 18). The next Education Act LOE (2006) refers to “inclusive education”, in which the school has to embrace diversity and adapt itself to all students. This would correspond to a universalist approach to access, meaning accessibility concerns all human beings (Greco, 2019, p. 18). The meaning of “inclusive” therefore fits with the latest advances in accessibility research, education, and had already been pointed out by Neves herself back in 2008 when she started to question the SDH label:

If subtitles are well devised for the d/Deaf they will be equally useful for hearers. They may not be ideal for each person, but they will be “good enough” for most viewers. So rather than having subtitles for the hearing impaired, at a stage when we cannot have individually tailored subtitles, one should be pursuing subtitles that are reasonably adequate “for All”. Inclusive subtitles should not be labelled; they should not reinforce loss or lack. In stressing “deafness” they are reinforcing discrimination even if positively meant. They could be simply called intralingual subtitles, interlingual subtitles, (stressing the language issue); or prepared and live subtitles (to emphasise production time). Perhaps they could be called “full subtitles” (to include all the extras that now go with SDH), or they could simply continue to be called “subtitles” (Neves, 2008, p. 140, emphasis added).

Inclusive subtitling refers to (potentially interactive) intra- or interlingual, open or closed subtitles involving more information than dialogues and inserts and catering to a wider audience (Neves, 2018, p. 83–84). However, as per current market practices offering one single set of subtitles, until customisable subtitling becomes the norm, inclusive subtitles seem to better adapt to postlingually deaf and hearing viewers, rather than to the prelingually Deaf (Pereira & Lorenzo, 2005).

2. Analysis of a Minoritised Language: Galician

Since 1981, Galician (galego) has been an official language in Galicia, a 29,576 km² region (Autonomous Community) in north-western Spain, home to 2,699,499 inhabitants (IGE, 2019a). The Galician Statistical Office (Instituto Galego de Estatística, IGE) gathers detailed quantitative data on language knowledge and use every five years, the most recent report corresponding to 2018 (IGE, 2019b). As of 2018, 88.1% of the population in Galicia aged 5 and above had some or high knowledge of the language, which equals over 2.3 million speakers of Galician (IGE, 2019b). Galician is also spoken, though without an official status, in border areas of neighbouring regions, like the

1 “That it be non-neutral (creative and/or subjective); that it be collaborative so as to reflect the director’s vision (auteur); that it is considered a priori; and, that it be open and inclusive, available to be heard by all” (Fryer, 2018, emphasis added).
westernmost parts of Asturias (Eo-Navia) and Castilla y León, as well as in Extremadura (northwest Cáceres) (Ramallo, 2018, p. 482; RAG, 2019). These adjoining territories are estimated to add 75,000 more speakers of Galician (RAG, 2019).

Although 11.9% had little or no knowledge of the language, the worrying fact is that the lowest levels of language knowledge were concentrated in the youngest age frame, 5–14 years old (IGE, 2019b). For the population below 30 years old, school has substituted family as the main medium of language transmission (IGE, 2014, 2019b). Even more daunting are percentages for language use according to age, social, educational and geographic distribution. As for general language use, 30.6% always spoke Galician, 21.7% spoke more Galician than Spanish, 23.3% spoke more Spanish than Galician, and 24.4% always spoke Spanish. However, while 73.4% of the population aged 65 and above spoke only or mainly Galician, 73.9% aged 5–14 spoke only or mostly Spanish. In the society, less than 30% at that point in time used only Galician to communicate with someone of a higher socioeconomic status (teachers, doctors, bank or administration staff). In compulsory education, teaching was equally balanced between Galician and Spanish. Geographically, the highest percentage of Galician-speaking families were concentrated in territories below 10,000 inhabitants at 45.7%, whereas in areas with a population of over 50,000, this percentage dropped to a 5.7% of exclusively Galician-speaking families (IGE, 2019b).

As a conclusion, language use is much more reduced than language skills and as such is restricted to the eldest generation (aged 65 and above), informal situations and small territories; Spanish still enjoys higher status, being used in all situations, dominating among the younger generations and in large cities. The following brief account of the sociolinguistic history of Galician (qualitative data) aims to explain the above figures.

Galician had been a socially and culturally prestigious language within the region until it was excluded from written use (16–17th centuries), which was recovered during the Renaissance (19th century) (Ramallo, 2018, p. 464). In the 20th century, a loss of Galician speakers was brought about by Franco’s dictatorship and emigration from rural areas to big cities and abroad: during Francoism, Galician was marginalised from public life, education and media, which were run in Spanish. Alongside its use by the upper class, Spanish was identified with upward social mobility and economic prestige (Cidadanía, 2002, p. 10; Monteagudo, 2002, p. 10, 26–27). In contrast, Galician denoted rural and poor origins, ignorance and uselessness outside family and informal contexts (Subiela, 2002, p. 156; Ramallo, 2018, p. 464). Since education, considered the key to a better life, was in Spanish, Galician-speaking families switched to Spanish to educate their offspring (Cidadanía, 2002, p. 58, 165; Monteagudo, 2002, p. 10). Conversely, in Catalonia, Catalan was spoken by the middle and upper classes, so the language became “a passport to integration and upward social mobility” (Mar-Molinero, 2000, p. 88).

After the end of Franco’s regime, the Spanish Constitution (1978) led to the establishment of 17 self-governed Autonomous Communities (Comunidades Autónomas) and the Statute of Autonomy of Galicia (1981) declared the Galician language co-official with Spanish in the region of Galicia. From
1983 (enactment of the Law of Linguistic Normalisation) to 2009, actions were taken to normalise and standardise Galician. The objective was to recover its prestige, by spreading its use to new areas, notably administration, education and the media, and by improving the population’s language skills, teaching the newly created standard (1982) with a focus on reading and writing. In other words, the aim was to place Galician at the same level as Spanish, i.e., both languages could share the same communicative situations and be equally used, increasing the use of Galician without decreasing the use of Spanish. However, conflict with powerful Spanish-speakers claiming Galician was being imposed on them has put a curb on the Galician promotion process since 2009 (Monteagudo, 2019, p. 20, 24–27).

The media are a fruitful platform for normalisation (thanks to their prestige) and standardisation (thanks to their high coverage, allowing the spread of the standard to a great audience) (Hermida, 2012, p. 47). As for the normalisation and standardisation results in this area, TVG, committed to the promotion of Galician, has been criticised for language quality issues (Xunta de Galicia, 2004; Hermida, 2012, p. 47–48). These issues comprise oral grammar and vocabulary mistakes and pronunciation errors, mainly caused by the many similarities between Galician and Spanish, which are both Romance languages. Due to the fact that the normalisation and standardisation efforts focused on writing skills, some speakers lack oral fluency. According to language experts (Hermida, 2008, p. 72), the problem is not that speakers use converging structures and words, identical in the dominant and minoritised languages, or even mix both languages, as will be detailed in the following section. Concerns arise because the structures not shared with Spanish—the distinctly unique Galician grammar and vocabulary—are not used or even known. Should Galician lose its unique richness, it risks being assimilated by Spanish, becoming one and the same language (Hermida, 2008, p. 72). TVG has also been criticised for not giving visibility to geographic varieties, by only portraying the standard (with the above-mentioned language deficit), which disregards many traditional dialects. The presence of language errors and the absence of dialects have caused prejudice against Galician to be maintained.


Summing up the previous section, the initial actions taken to normalise and standardise Galician were aimed at increasing the number of speakers by teaching the standard, notably reading and writing skills, and by spreading language functions to administration, education and media. As for the latter, TVG has been criticised for not being a high-quality language model, allowing language errors to creep into its content. Consequently, it seems that, from an exclusively linguistic viewpoint, all oral language mistakes should be corrected in inclusive subtitling.
On a seemingly opposing extreme, verbatim (complete and literal) subtitles have been globally demanded by some D/deaf users, Associations for the Deaf and by broadcasters. Edited subtitles in any AV content are regarded by part of the D/deaf audience as censorship, since they do not receive the same quantity and exactness of information as their hearing counterparts (Neves, 2005, p. 142; Ofcom, 2005, p. 17). Broadcasters favour verbatim subtitles on economic grounds: a word-by-word transcript of the audio simplifies the subtitling task (no editing decisions need to be taken), making it faster and cheaper and helping to reach subtitling quotas (Ofcom, 2005, p. 6). Because hearers are also inclusive subtitling users (Ofcom, 2006, p. 13), able to tell the differences between the audio and subtitles, editing heightens the vulnerability of subtitling. Also known as the “gossiping” or “feedback effect”, the simultaneous presentation of both source and target texts makes any divergence in the subtitling vulnerable to comparison and negative judgements by the viewer (Díaz-Cintas & Remael, 2007, p. 55). Therefore, not only must the sense of the message be maintained, but also its form, using similar words and structures in both the audio and subtitles (Díaz-Cintas & Remael, 2007, p. 56). In consequence, considering some users refer to subtitle text reduction and editing as “changing” (Ofcom, 2005, p. 17), in order to fulfil users’ preferences and minimise the vulnerability of subtitling or a negative perception of the subtitles, oral language mistakes should not be corrected at all.

To correct (all) or not to correct (at all), that is the question. This section seeks to unravel the clash between media accessibility (leading to subtitling oral errors) and normalisation and standardisation of minoritised languages (leaning towards correction) by studying key literature and subtitling practices.

Scholars like García del Toro (2004, p. 117), Ivarsson and Carroll (1998), Díaz-Cintas (2003), and Díaz-Cintas and Remael (2007) have maintained that subtitles should be written in standard language, especially standard grammar, for the sake of comprehension and to make subtitling useful for learning. To add a colloquial flavour to the text, vocabulary changes are more favoured than using non-standard grammar.

The Spanish standard UNE 153010:2012 (AENOR, 2012) dedicates three short subsections to language criteria: 10.4, according to which subtitling in Spanish should follow the grammar and spelling rules established by the Royal Spanish Academy (exception 10.6); 10.5, which states that subtitling in the other official national languages must follow equivalent criteria by analogous institutions [for Galician, the Royal Galician Academy]; and 10.6, which specifies that oral errors must be corrected except when mispronounced words and grammar and regional inadequacies carry plot-relevant information (AENOR, 2012, p. 17). TVG, following the UNE, does correct some grammar and vocabulary mistakes from the audio in inclusive subtitling, such as Spanish words (*persona* > *persoa*,

---

2 The Deaf (capital “D”) refers to the cultural and linguistic minority of sign language users, while the deaf (lower case “d”) and hard of hearing correspond to people with a hearing loss who use an oral language (Neves, 2008, p. 129).
Subtitling for Social and Language Minorities: Subtitling of Oral Errors and Dialectal Features in the Case of Minoritised Languages

*costumbre* > *costume* and misplaced pronouns (*Se* *nota* > *Nótase*), but also leaves others uncorrected (e.g., Spanish words *ahora**, perdon*a* and wrong verb conjugations *fora**, instead of *agora*, *perdoa*, *fose*), without responding to any discernible pattern. Multiple misspellings observed in the writing of subtitles from TVG (e.g., *coordinadod**, Pacífiuco**, *envína* for *coordinador*, *Pacífico*, *envían*)³ suggest the correction or reproduction of oral errors is done rather randomly (Martínez-Lorenzo, 2019, 2020).

The subtitling linguistic criteria for Catalan TV (ésAdir, CCMA, n.d.) draw a distinction between formal and informal registers. In formal speech broadcasts, such as the news or factual topics, language errors pertaining to grammar (e.g., verbs, plurals, gender) and vocabulary (e.g., *maravillós*/*meravellós*) are to be corrected. If the correction greatly differs from the audio, the mistake should be reproduced without marking (neither quotes nor italics) or the subtitle text may be edited omitting the error (e.g., *És algo* *molt bonic* > *És algo molt bonic* / *És molt bonic*). In informal speech programmes, such as fiction or talk shows, grammar errors are to be corrected. However, non-normative vocabulary (dialectal features and mistakes influenced by Spanish) relevant for character and context description may be reproduced without marking.

The ISO/IEC standard (2018) indicates the following in section 9.2 Grammar: “Verbal content should be presented accurately as heard, even when grammatically incorrect. NOTE. *In some cases*, it is more important to present a caption/subtitle as close to verbatim as possible than correcting poor use of language in audio content” (emphasis added). However, it does not exemplify in which cases grammar errors should be preserved.

Netflix Timed Text Style Guides in Portuguese, Spanish and English (2020) include the following Special Instructions: “Deliberate misspellings and mispronunciations should not be reproduced in the translation unless plot-pertinent.” The British Broadcasting Company (BBC) (Ford, 2009, p. 23; BBC, 2019) adds that editing language will likely be the case for unscripted speakers, in “factual content, e.g. News and documentaries. Do not tidy up incoherent speech in drama when the incoherence is the desired effect.”

According to these standards and guidelines, correction or reproduction of oral errors predominantly depends on the effect a mistake may have, i.e., whether it is plot-pertinent; the type of broadcast, i.e., language register, speech control/spontaneity or genre (formal/informal register programmes, unscripted speakers, factual content vs. drama); also, certain characteristics of a broadcast may suggest errors are plot-related, as will be discussed later. Correction or reproduction may depend on the type of language error (vocabulary vs. grammar) and on how noticeable the correction is in comparison with the audio (vulnerability of subtitling), which may be linked. Therefore, the extremes of correcting *all* errors or not correcting them *at all* should be ruled out. Firstly, if mistakes have an

---

³ Examples of errors and misspellings retrieved from the interview programme O gato con botas, from the episode featuring a Galician archaeologist, Pablo Novoa Álvarez (shown March 2nd 2018 on tvG2), and the cooking show Larpeiros, episode Empanadas e empanados (shown March 27th 2018 on TVG1).
effect on the narrative, these are to be reproduced in subtitling; mistakes with no apparent effect should be amended.

However, if all errors with no effect are corrected, the vulnerability of subtitling would be increased, especially in the case when spontaneous speech containing abundant linguistic inadequacies is subtitled into standard language. Also, considering that inclusive subtitling users with full or residual hearing may be able to detect the difference between an imperfect audio and linguistically perfect subtitles, would that difference be good for normalisation or would it draw more attention to the mistakes? Do all language errors pose the same threat to the survival and growth of minoritised languages? Which types of language mistakes constitute a real peril? Galician literature (Gugenberger, 2013; Rei-Doval, 2013; Silva, 2013) has been considered the most appropriate source for finding a solution that best adapts to Galician, although standards and guidelines for subtitling other languages have proven useful for taking the first step towards solving the main question of this paper. On the other hand, indications by Galician scholars on which types of errors should be corrected might prove equally suitable for other minoritised (and perhaps also for majority) languages.

Viewpoints on language standardisation, i.e., on establishing a variety deemed correct to be learned, vary between the purist and hybrid approaches. For purists, bilingualism is understood as double monolingualism: a perfect command of two languages independent of one another. This means that, e.g., a Galician learner should not compensate for their lack of knowledge in this language with Spanish forms, as mixing languages is a sign of imperfect command and these influences are considered errors that corrupt the language (Gugenberger, 2013, p. 20). “Hybridism” is put forward as an umbrella term encompassing the varied nuances of, among others, loanwords/borrowings, interferences, transfers, code-switching and code-mixing (Gugenberger, 2013, p. 21–28). This approach allows for the learning of a language to be based on previously learnt languages, and influences are not considered errors but a sign of creativity (Gugenberger, 2013, p. 38). Both purism and hybridism have their pros and cons, and should actually complement each other (Ruiz i San Pascual, Sanz i Ribelles, & Solé i Camardons, 2001, p. 235; Rei-Doval, 2013, p. 265): a purist approach may be necessary at the start of the language promotion process to highlight the value of the language by showing it is rich enough for any communicative situation, without needing to be complemented by any other language. This can boost the speakers’ confidence by strengthening the language-identity link (Gugenberger, 2013, p. 42). However, it may also discourage (traditional and native or non-native new) speakers if the standard is too far from their way of speaking and deviations from the norm are censored, shyng them away from speaking the language as they believe they would be incorrect. On the other hand, because mistakes are not negatively perceived, a hybrid approach may encourage speakers to use the language as they know it; in this case, however, accepting all language errors might hinder learning the standard.

Extremes should again be ruled out and a middle ground sought regarding error types. Silva (2013) also contrasts vocabulary with grammar. Thanks to the success of the standardisation process on written usage, the population schooled in Galician (i.e., from the 1980s onwards) learns and knows
the standard vocabulary. Nonetheless, the lexis learnt at school may not be used orally for being deemed too formal and impersonal. This is because the standard, acquired through education, has been associated with formal communicative settings, while Spanish or converging vocabulary is related to familiarity and proximity. Using unique Galician words vs. Spanish or converging ones would be similar to using “ophthalmologist” vs. “eye doctor” according to the context. Most importantly, the use of non-normative vocabulary is not due to the lack of linguistic competence, but a speakers’ choice and part of their idiolect (Silva, 2013, p. 297–298). Conversely, Silva’s studies (2006, in Silva, 2013) with students in their last year of compulsory secondary education (4th ESO, 15–16 years old) show that grammar errors result from the lack of knowledge: diverging uniquely Galician structures (correct) were unknown and identified as incorrect by students, whereas converging (correct) and Spanish structures (incorrect) were selected as correct.

Consequently, only grammar, being the base of any language, should be corrected. In addition, any change of vocabulary in subtitles is deemed more obvious with respect to the audio, therefore more detrimental to the vulnerability of subtitling and speakers’ language perception (normalisation): e.g., substituting the Spanish word silla (chair) for the Galician cadeira in contrast to amending the incorrect position of a pronoun. Correcting grammar seems to be in line with the previously analysed literature and subtitling guides, as well as with Canadian subtitling standards (CAB, 2012, p. 14):

The art of transcribing spontaneous speech for captions is very different from the creative process of writing dialogue. People involved in real conversations do not necessarily use grammatically correct sentence structure. They may use improper grammar, incomplete sentences, run-on sentences, slang, vernacular expressions, and so on.

An important challenge in captioning is to accurately represent imperfect verbatim speech, while providing text that is grammatically clean [...].

Indeed, the effect of the error and the characteristics of the broadcast (speech control, genre) will have to be linked to the type of error (vocabulary vs. grammar) when deciding whether to correct or reproduce a given mistake.

4. Protocol for the Correction or Reproduction of Oral Language Errors and Dialectal Features in the Inclusive Subtitling of a Minoritised Language

If the subtitler has access to the client, filmmaker or creative team of the AV content, as per the accessible filmmaking model (Romero-Fresco, 2019), the following protocol will be discussed to reach a joint decision on whether to reproduce or correct all dialects and oral errors, which ones to leave and which to amend, and how to convey them in inclusive subtitles (marked/unmarked). This communication and joint agreement will help to preserve the artistic intentions of the creators and to unburden subtitlers from making decisions in isolation. If a given speaker can be contacted, they may also be consulted on whether they agree to their speech being corrected or reproduced as it is.
A fundamental distinction is made between errors and dialectal features. In line with Silva (2013, p. 300), errors refer to non-standard language that does not pertain to any dialect. The words “error” or “mistake” should not be negatively perceived, as these may not hinder communication. However, mistakes in a minoritised language may be owed to the influence by the majority language and have a negative effect on the minoritised one, i.e., recipients may deem the language not rich enough on its own and therefore not useful for all communicative situations (cf. purist approach). Dialectal traits, on the other hand, contribute to language variety and richness and have a positive impact on the minoritised language. If the subtitler works on their own, the following protocol presents the items to consider and the steps to follow.

4.1. Errors

This original protocol for the correction or reproduction of oral errors is divided into five levels for the subtitler to follow in order. The first two characterise the AV content, while the latter three levels classify the error. Figure 1 below visually presents the protocol.

1. Speech control, i.e., the degree of spontaneity a speaker has to express themselves.

Although the protocol works on the extremes of high and low, speech control must be understood as a continuum. An example of low speech control is an interview. Advertising and fiction shows are instances of high speech control, since a script prepared beforehand is memorised and delivered by actors and actresses. However, high/low speech control may not always be equivalent to scripted/unscripted content, e.g., talk shows may follow a script on the topics to discuss, but hosts present them more freely.

2. Audience and genre: a difference is made between content for adults, children, and the news.

Considering that children are in their prime learning age, correct language input should be presented in AV content specific for them to favour learning the standard, as opposed to content for adults or the family, where more non-standard solutions are permitted. The news is considered separate from the content for adults as a special case: news presenters (who may be following a teleprompter) and journalists represent high-control speech. Linguistic rigour is necessary to achieve informational rigour, i.e., a piece of news presented with language errors may lose credibility, and therefore errors should be corrected. Conversely, the more spontaneous speech of eye-witnesses or interviewees corresponds to low control, and thus more errors may be reproduced.

3. Effect, i.e., plot-related errors vs. spontaneous mistakes.

Errors with an effect should always be reproduced, regardless of the type of AV content (speech control, audience or genre). Errors with no effect should all be corrected in programmes for children and content delivered by news presenters and journalists. In content for adults, mistakes with no effect should be amended according to guidelines in level 4 below (type of language error).
Level 3, effect, may be linked to level 1, speech control, since errors in high-control speech are likely to have an effect, e.g., mistakes in a scripted fiction film. However, errors in spontaneous low-control speech are likely to have no effect. Consequently, despite the proposed subtitling actions being identical for both high and low speech control, as can be noted in Figure 1, this previous level 1 determines which subtitling solutions will likely be more frequent. Infrequent solutions appear shaded in Figure 1. The division between formal and informal registers introduced by the Catalian guidelines would not have favoured this correspondence between speech control and effect: formal register may imply more language care, and therefore the unlikelihood of errors with no effect; however, because informal register comprises spontaneous interviews from talk shows and scripted speech from drama, it cannot be linked to the (im)probability of errors with effect.

4. Type of language error: vocabulary vs. grammar.

In line with Silva (2013), as well as AVT scholars and Catalan subtitling guidelines, plot-unrelated vocabulary errors should be reproduced, since these result not from the lack of knowledge but the speaker’s choice, while grammar (e.g., verb conjugations, pronoun placement, gender), being the base of the language, should be corrected.

The exception for vocabulary comprises plot-unrelated toponym errors, since the Galician Law of Linguistic Normalisation (1983) determines the names of places located in Galicia are in Galician. For instance, if someone says Orense* (one of the four Galician provinces), the subtitle should read Ourense.

5. Marking: in italics or, if there are technological limitations, English quotations (“ “), Latin quotation marks are bulkier and as such should be avoided, « »).

Only plot-pertinent errors reproduced in the subtitling should be marked; this is also in line with Catalan guidelines. Mistakes should not be marked if the abundance of plot-related errors is part of a character’s idiolect (e.g., Tarzan) (AENOR, 2012, p. 17), so as not to keep changing to and from italics or plague the text with quotes, making reading more difficult. Marking plot-unrelated errors in any way may humiliate the speaker and draw more attention to the mistake. Fernández, Matamala and Vilaró (2014), in line with authors like McConkie and Yang (2003), found that unmarked non-standard units in a subtitle slightly increase the time spent on the subtitling area; this could be exacerbated if those units are typographically highlighted. Unmarked errors may be attributed to slips by the subtitler, but this is only the case for viewers with no access to the audio; viewers with access are able to compare the audio and subtitles and note that the mistake was also said (for once, the vulnerability of subtitling seems to be on the subtitler’s side). So far, the reasons against marking plot-unrelated errors (italics are more difficult to read; embarrassing for the speaker; in line with current practices; draws even more attention to the non-standard subtitle) seem to surpass the argument for marking (so that errors are not blamed on subtitlers). Future research could explore other possibilities, such as tags, e.g., (LITERAL), and carry out reception studies with viewers (with
and without hearing loss) and even with the subtitled speakers. Corrections should not be marked either.

Figure 1.

**Protocol for the correction or reproduction of oral language errors**

<table>
<thead>
<tr>
<th>1. Speech control</th>
<th>Adult</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes: Reproduce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Audience and genre</th>
<th>Adult</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes: Reproduce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Effect</th>
<th>Adult</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes: Reproduce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No: Correct all</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Type of error</th>
<th>Adult</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary³: Reproduce</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar: Correct</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Marking⁴</th>
<th>Adult</th>
<th>Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

2 Except toponym. Places within the region of Galicia are named in Galician language.
3 In italics or, if not possible, English quotations. Except if abundant; if so, do not mark.

### 4.2. Dialectal features

All dialectal traits should be reproduced in the subtitles, regardless of speech control, audience, genre, or possible effect of the dialect, and they should be unmarked, because dialectal features contribute to language richness. Maintaining them is in line with other subtitling practices recommending subtitling regional, slang and dialectal features (Ford, 2009, p. 4–5; BBC, 2019; Netflix EN, 2020). Reproducing dialectal features tackles one of the main points of criticism against TVG, i.e., the underrepresentation of dialects. Instances of dialects are *muito, pantalóis*, instead of the standard *moito, pantalóns*, or the phonetic phenomena *seseo* and *gheada*. Sseo consists of pronouncing the phoneme /θ/ (th in English) like an /s/. Gheada consists in pronouncing the phoneme /g/ as an English h (RAG, 2019). Examples of the written representation of these oral phenomena are *lus* or *ghato*, with standard forms of *luz* and *gato*.

Further language criteria concern full utterances in Spanish and misspellings. Spanish speakers in any AV content should be subtitled in Spanish and not translated to Galician. It is understood that Galician subtitles are also proficient in the majority language. Not preserving this bilingualism may lead to misunderstanding of the plot, as was the case of episode 17 from the fiction TVG series *Viradeira*, where a main character from outside Galicia makes the following comment on the protagonist’s father-in-law organising a stag for his son-in-law, in Spanish: “So, your father-in-law taking you to see strippers is not typical of Galicia?” This joke, based on the character pretending not to know Galician customs because she is from outside the region, makes no sense for an inclusive subtitling user with no access to the audio, for whom this main character, subtitled in Galician, is therefore Galician and is bizarrely asking about her own “traditions”.

Besides the non-standard language uses discussed above, misspellings must not be present in the subtitles, so a careful revision is strongly advised. While this may seem a truism, several spelling and
typing slips were observed during the quality analysis of subtitling done at TVG (Martínez-Lorenzo, 2019, 2020). Lastly, whenever oral language errors and dialectal features are reproduced in the subtitling, exposure time of the subtitle should be extended (Ford, 2009, p. 22; BBC, 2019).

Future steps for the above protocol concern its testing and spreading. In April 2020, the new Guidelines for Inclusive Subtitling in Galician comprising the protocol were tested by the author with a group of 33 students of Audiovisual Translation. The students already had training in interlingual subtitling (for hearing viewers) and were asked to subtitle 11 clips (13 minutes). Student group 1 subitled 6 videos (6.5 min) following the Spanish standard UNE 153010, and the remaining 5 (6.5 min) were subitled following the Galician Guidelines. Student group 2 subtilted the latter 5 videos following the UNE, and the other 6 clips according to the Galician Guidelines. The main goal of the testing was to assess whether the Guidelines were easily understood and implemented (i.e., formulated clearly and exhaustively), and if the resulting subtitles were more consistent than those produced by the UNE, which could lead to the conclusion that the Galician Guidelines cover and improve the Spanish standard’s weaknesses and represent a step forward in the field of subtitling guidelines.

As for its reception, the door is open for future studies to test the Guidelines with inclusive subtitling users. The Royal Galician Academy has expressed interest in the project and in reviewing the Guidelines. Future efforts are directed towards the Galician Government (Xunta de Galicia) to ratify the Guidelines so they become the region’s official standard and towards TVG to follow the guides.

5. Conclusions

Galician is still a minoritised language. Its linguistic proximity to Spanish means speakers often mix Galician and Spanish within a sentence or within a word. The Galician public broadcaster TVG has been criticised for not showing the linguistic richness of the different Galician dialects and for letting language errors creep into its content. Therefore, dialectal vocabulary, grammar and phonetic traits should always be reproduced in any type of AV content for any audience.

Concerning the correction or reproduction of oral language errors, there seems to be an initial clash between the promotion of minoritised languages, which requires correction, and media accessibility, which leans towards reproducing oral mistakes. Ruling out extremes, a middle-ground solution has been sought as to which language errors to edit and which to maintain. Inclusive subtitling guidelines and standards point to considering the genre and characteristics of the broadcast and the effect of a mistake on the plot. Galician literature, as well as AV scholars, guidelines and standards, suggest differentiating types of language errors (vocabulary vs. grammar).

Errors with an effect on the narrative (plot-pertinent mistakes) should be reproduced, which is most likely the case with programmes on the high spectrum of what the original protocol introduced in this paper calls “speech control.” Errors with no effect (plot-unrelated mistakes) should all be corrected in content for children and when news-presenters and journalists speak in the news, so as
to respectively favour the learning of the standard and information rigour. In AV content for adults, plot-unrelated errors are to be corrected according to the type of language error. Vocabulary errors should be maintained, with the exception of toponymy, as incorrect vocabulary may result from a speaker’s choice and not lack of knowledge. Changes in lexis are also deemed more noticeable than editing grammar, and therefore detrimental to the vulnerability of subtitling. Grammar errors should be corrected.

This protocol for the correction or reproduction of oral language errors and dialectal features in the inclusive subtitling of a minoritised language is believed to provide a balance between media accessibility and normalisation and standardisation of minoritised languages. By applying this protocol, the resulting inclusive subtitling could be used for learning and promoting the standard and dialectal varieties, while remaining close to the audio of the AV material, minimising the vulnerability of subtitling and fulfilling users’ preferences to varying degrees. Indeed, the protocol produces subtitles that seem to better address the needs and preferences of the postlingually deaf or hearers. Future research could assess whether this protocol would work for the needs of the prelingually Deaf for more adapted “simpler subtitles, with basic vocabulary, a slower reading rate, and the incorporation of features of SSL [sign language]” (Pereira, 2010a, p. 100), as indicated by Pereira (2010a, 2010b) and Lorenzo (2010a, 2010b), and considering that: “Any form that deviates from the standard (lexis from a different time period or geographical area, slang, teenage expressions, etc.) should be avoided” (Lorenzo, 2010b, p. 143).

Acknowledgments

I am grateful to Dr. Ana Pereira, Dr. Pablo Romero and Dr. Fernando Ramallo for their constant support and guidance, for all the tutoring hours, recommended bibliography and insightful comments yielding the protocol. Many thanks to Matthew Delicate for the proofreading and relevant language corrections.

References


—. (2012). A contribución da radio e da televisión públicas de Galicia á normalización e normativización da lingua galega [The contribution of Galician public radio and television to the normalisation and standardisation of the Galician language]. In M. Ledo, X. López, & M.


Rei-Doval, G. (2013). Purismo e control normativo na lingua galega: análise crítica dunha proposta actual [Purism and control through the standard in the case of Galician: critical analysis of a present proposal]. In E. Gugenberger, H. Monteagudo, & G. Rei-Doval (Eds.), Contacto de linguas, hibrididade, cambio: contextos, procesos e consecuencias [Contact between languages,


