Speech-to-text Recognition for the Creation of Subtitles in Basque: An Analysis of ADITU Based on the NER Model

Ana Tamayo, University of the Basque Country
Alejandro Ros-Abaurrea, University of the Basque Country

ABSTRACT

This contribution aims at analysing the speech-to-text recognition of news programmes in the regional channel ETB1 for subtitling in Basque using ADITU (2024) (a technology developed by the Elhuyar foundation) applying the NER model of analysis (Romero-Fresco and Martínez 2015). A total of 20 samples of approximately 5 minutes each were recorded from the regional channel ETB1 in May, 2022. A total of 97 minutes and 1737 subtitles were analysed by applying criteria from the NER model. The results show an average accuracy rate of 94.63% if we take all errors into account, and 96.09% if we exclude punctuation errors. A qualitative analysis based on quantitative data foresees some room for improvement regarding language models of the software, punctuation, recognition of proper nouns and speaker identification. From the evidence it may be concluded that, although quantitative data does not reach the threshold to consider the quality of recognition *fair* or comprehensible with regards to the NER model, results seem promising. When presenters speak with clear diction and standard language, accuracy rates are sufficient for a minority language like Basque in which speech recognition software is still in early phases of development.

KEYWORDS

NER model, speech-to-text, speech recognition, subtitling, intralingual subtitling, accessibility, Basque.

1. Introduction

This contribution aims at analysing the speech-to-text recognition in Basque of recorded news programmes in the regional channel ETB1 applying the NER model of analysis (Romero-Fresco and Martínez 2015). In the present study, the analysis focuses on the output (the subtitles) and does not delve into the speech-to-text recognition process. In this sense, this contribution is an evaluation of the quality of subtitles in Basque using speech-to-text technology that might help improving the automation process of subtitling in Basque.

In this section we take into account a brief history of Basque within the regional channels of ETB and the current situation and availability of subtitling in the different ETB channels. In section 2, we explain the NER model, which will serve as the basis for gathering quantitative and qualitative data for the analysis. In section 3, the methodology of how samples were gathered and analysed is explained in detail. Section 4 offers the analysis and discussion of the data. In that section we look specifically at quantitative data on the accuracy rate (AR) of the recognition software (4.1), recognition errors and corrected recognition (4.2), as well as subtitle speed (4.3). We also offer a qualitative analysis of where we believe there is room for improvement (4.4) in the recognition software for the creation of automatic subtitling in Basque. Throughout the analysis section, we will be referring to an interview with

Igor Leturia, Head of speech technologies at Orai Natural Language Processing Technologies, who works in the development of ADITU (2024). This interview took place on December 14, 2022, after the quantitative analysis was carried out and a report sent to the interviewee. To conclude, section 5 offers some final remarks emphasising the contributions of the paper.

1.1. Subtitling in Basque in ETB

The Basque language, or *euskera*, is a minority language spoken in Euskal Herria, a small Basque-speaking area of 20,664 km^2 that spans the banks of the Bay of Biscay across northern Spain and southern France (Zuazo 1995: 5). The Basque Statistics Institute (Eustat) divides Basque speakers in three categories, depending on their proficiency and fluency levels: Basque speakers with full proficiency in Basque; Quasi-Basque speakers who face some difficulties when speaking the language, and non-Basque speakers who can neither speak nor understand Basque (Lasagabaster 2010: 404). Table 1 shows the percentage of speakers from each category in Navarre, the Basque Autonomous Community (BAC) and Iparralde[1] (Euskararen erakunde publikoa, Eusko Jaurlaritza and Nafarroako Gobernua 2023a, 2023b, 2023c):

Category	Navarre	BAC	Iparralde
Basque speakers	14.1%	36.2%	20.1%
Quasi-Basque speakers	10.6%	18.6%	9.4%
Non-Basque speakers	75.3%	45.3%	70.5%

Table 1. Number of Basque, Quasi-Basque and non-Basque speakers in Navarre, BAC and Iparralde in 2021 (Euskararen erakunde publikoa, Eusko Jaurlaritza and Nafarroako Gobernua 2023a, 2023b, 2023c)

In Spain, since Basque acquired co-official status in 1978, a number of efforts have been made in order to revive the language (Cenoz and Perales 1997). The creation of a television channel[2] entirely in Basque, along with opening primary schools that teach in Basque, is closely linked to this desire to revitalise the language (Costa 1986: 343).

On May 20th 1982, the Basque Parliament unanimously adopted an act creating Euskal Irrati Telebista (EiTB). The first year of broadcasting constituted a trial period marked by coverage problems and the difficulty of complying with the initial plan to broadcast four hours a day (Larrañaga and Garitaonandía 2012: 34). Thus, the programming consisted, almost entirely, of foreign programmes (such as soap operas, cartoons or documentaries) dubbed into Basque and subtitled in Spanish (Barambones 2009: 89–91). However, over the following years the number of hours of broadcasting increased, as did the range of programmes offered. Just over

three years after ETB's first broadcast, and after a string of political and sociolinguistic controversies, ETB2 was born with the aim of responding to the needs of a large percentage of the Basque population who did not speak nor understand Basque, but wanted to access content that covered their closest reality (Cabanillas 2016: 57).

Nowadays, EiTB Media owns four broadcasting channels: ETB1, a generalist channel with Basque as the broadcasting language; ETB2, a generalist channel with Spanish as the broadcasting language; ETB3, a channel aimed at broadcasting children's and youth programmes mostly in Basque; and ETB4, a channel that broadcasts reruns in both Basque and Spanish. With regard to the subtitling practices, Table 2 and Figure 1 summarise current practices and percentage of hours subtitled.

	Basque	Spanish
Channels	ETB1, ETB3	ETB2, (most of) ETB4
Live-subtitling	No	Yes. News programmes and weather forecast through respeaking from EiTB headquarters in Bilbao. *En Jake* and *Nos echamos a la calle* through respeaking by a company from Burgos.
Semi-live subtitling	No	No
Pre-recorded subtitling	Yes, most programmes With editing, using speech-to-text programme (Idazle, developed by Vicomtech). Without editing, weather forecast using speech-to-text programme (Idazle, developed by Vicomtech).	Yes, most programmes

Table 2. Current subtitling practices in ETB. Source: Larrinaga (personal communication, October 14, 2022)
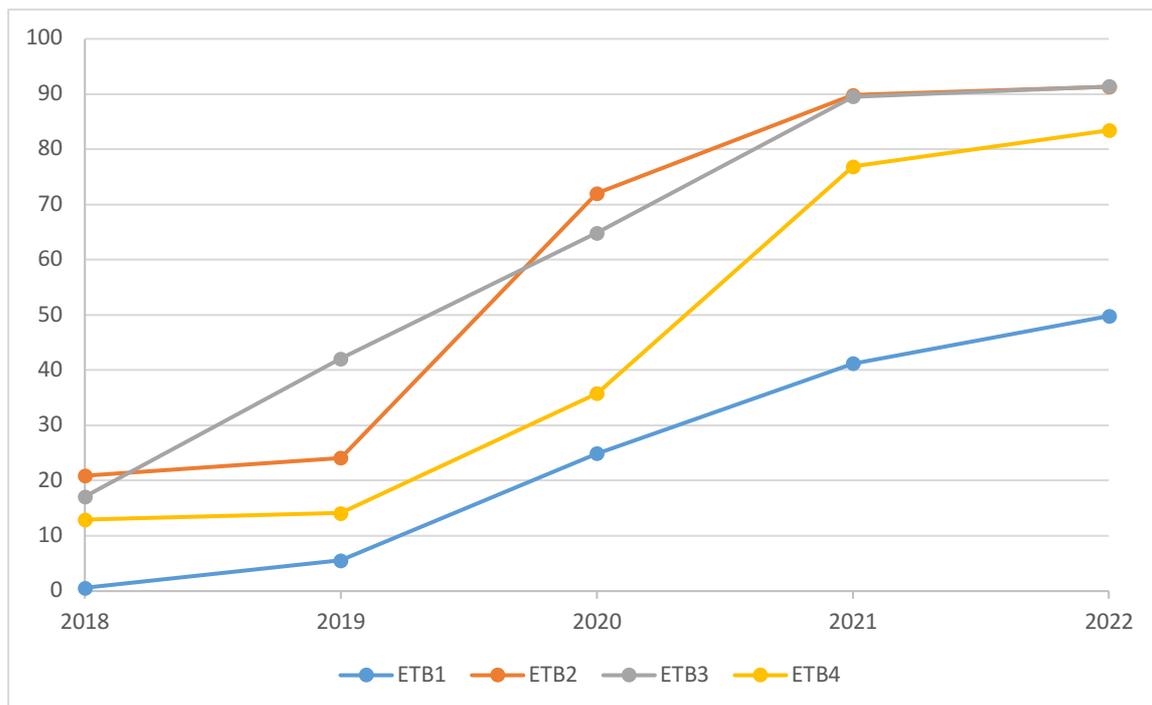
Figure 1. Evolution of percentage of measurable hours subtitled in ETB (2018-2022) (Source: Larrinaga, personal communication, October 14, 2022 and January 9, 2023)

In the matter under discussion, the reality is that no live nor semi-live subtitling is offered in Basque in ETB. All the subtitles included in ETB1 are pre-recorded and have been automatically generated using speech recognition software and edited afterwards. Given the high error rate in speech recognition, it is imperative that they are edited before being broadcast which is why recognition software is only used with recorded programmes, such as *Mihiluze* (Larrinaga, personal communication, October 14, 2022). An exception to this is the weather forecast. The weather broadcast (*Eguraldia* in Basque) is a pre-recorded section of the news programme that includes automatic subtitling without editing (Larrinaga, personal communication, October 14, 2022). Thus, without editing, the quality of the subtitling depends, to a large extent, on the number of interferences registered, the speaker's diction, as well as on thematic limitations (Larrinaga, personal communication, October 14, 2022). After an initial test period, and with a view to improving the quality of the subtitles for the weather forecast, ETB's Basque Service compiled a technical weather glossary, a list of popular sayings, expressions and words related to the weather, as well as a list of Basque toponyms, which was later uploaded into the software and improved the quality of this section's subtitles (Larrinaga, personal communication, October 14, 2022). These subtitles are not analysed in the present contribution, as we deal here with another software.

In conclusion, current broadcasting does not offer automatic live subtitles in Basque generated by speech recognition and without editing. Therefore,

in order to gather samples to analyse the current recognition accuracy rates in Basque, we need to turn to non-broadcast subtitles. How these materials were gathered is explained in detail in the methodology section.

2. The NER model

In recent decades, the use of intralingual subtitles, especially those for the deaf and hard-of-hearing, has increased considerably, driven in large part by the ongoing efforts of some countries in the field of audiovisual accessibility (Pedersen 2017: 212). However, once subtitling quotas were relatively met and quantity had reached acceptable standards, attention was drawn to the quality and accuracy of the subtitles (Pedersen 2017, Romero-Fresco 2012). In this context, the NER model was developed, a product-oriented model aimed at analysing the accuracy of live subtitles (Romero-Fresco and Martínez 2015). By adopting the basic principles of the WER model (see Dumouchel, Boulianne and Brousseau 2011), the NER model, which divides transcripts into *idea units* (word chunks that are recognised as one unit) rather than words, has been particularly tailored to suit respoken and automatic intralingual subtitling (Pedersen 2017, Romero-Fresco and Martínez 2015). It is based on the following formula to determine the quality of the subtitles:

$$NER\ VALUE = \frac{N - E - R}{N} * 100$$

Image 1. The NER model formula

In order to calculate the overall score, the number of editing errors (E) and recognition errors (R) is deducted from total number of words (N) contained in the subtitles. The resulting number is divided by the N value, which is then multiplied by one hundred. An error-free subtitle would score a NER value of 100. It should be noted that the threshold set by the NER model for acceptable accuracies is 98% (Romero-Fresco and Martínez 2015, Romero-Fresco 2016). In contrast to the WER model, which is based on word errors regardless of their typology, the NER model differentiates between errors based on their severity, and some mismatches in recognition may not be considered as errors.

In our particular case study, editing errors are not taken into account, as the automatically generated subtitles in our corpus are not edited. Editing errors can occur in live editing (either by a respeaker or by a person controlling the automatically generated subtitles live) or in semi-live and prerecorded subtitling, as would be the case of some subtitling practices for ETB. For this contribution, only recognition errors (R) are taken into account.

3. Methodology

The analysis presented in this contribution was carried out within the project "QUALISUB: The Quality of Live Subtitling: A regional, national and international study". The methodology originally established for the project could not be followed. The aim of the project was to analyse the speech-to-text recognition on the regional news programmes broadcast by TVE, the Spanish public broadcast. Nevertheless, the subtitles provided by TVE between May and December 2021 contained too few automatic subtitles in Basque[3]. This resulted in having to follow a different methodology for Basque, when compared to other languages analysed in the project. The sociocultural reality in which live events and automatic subtitles take place in the Basque Country (see section 1.1) led to the need for using recorded programmes. How the samples were gathered is explained in detail in the next paragraphs.

News programmes from ETB1 (the regional TV channel broadcasting only in Basque) were recorded in May 2022[4]. The whole news event was recorded over 10 non-consecutive days, between May 2nd and May 15th, from all time slots (morning, afternoon and night) and from all days of the week. From each day, two samples of 5 minutes each were extracted, with the aim of having comparable samples with other languages of the project. Samples contain different types of subgenres (sports news, weather forecast, interviews, reporters, headlines, etc.) and are heterogeneous in the use of language (monolingual samples, bilingual [Basque and Spanish] samples, different dialects, etc.). Our corpus, therefore, consists of 20 samples from the news programmes of ETB1 recorded in May 2022. The samples add up to 1 hour and 41 minutes (101 minutes) and their characteristics are shown in Table 3.

Since no real (broadcast) automatic subtitles in Basque are available for live events in the Basque Country (see section 1.1), we proceeded to obtain automatic subtitles by contacting the Elhuyar Foundation. The Elhuyar Foundation is a private non-profit organisation. It was founded in 1972 as a cultural association with the aim of combining science and the Basque language. In 2002 it became a foundation. Currently it offers services for the application of advanced knowledge and aids companies, social agents and administrations to seek innovative solutions to respond to the challenges of globalisation from a multidisciplinary approach. The Elhuyar Foundation also works with professionals in science, artificial intelligence, lexicography, translation and language management.

ID	Time slot	TCR[5] start	TCR end	Genre	Minutes	Subtitles
ETB1_1	9-11 h	0:00:10	0:05:05	Opening, news, presenters.	0	0
ETB1_2	9-11 h	0:39:58	0:45:00	Interview	0:05:02	90
ETB1_3	14-15 h	0:20:37	0:25:38	News, presenters, reporters	0:05:01	85
ETB1_4	14-15 h	0:54:30	0:59:32	Weather forecast	0:05:02	113
ETB1_5	20-21 h	0:10:09	0:15:08	News, presenters, reporters	0:04:59	89
ETB1_6	20-21 h	0:38:14	0:43:29	Sports	0:05:15	92
ETB1_7	14-15 h	0:07:11	0:12:18	First news: presenter, reporters	0:05:07	95
ETB1_8	14-15 h	0:43:47	0:49:02	Sports	0:05:15	91
ETB1_9	20-21 h	0:31:35	0:36:47	Breaking news: culture, history (presenter, reporters)	0:05:12	88
ETB1_10	20-21 h	0:53:10	0:58:30	Weather forecast	0:05:20	99
ETB1_11	9-11 h	0:00:17	0:05:20	News and presenters	0:05:03	81
ETB1_12	9-11 h	0:27:35	0:32:30	On-set chat show	0:04:55	79
ETB1_13	14-15 h	0:24:10	0:29	News	0:05:04	73
ETB1_14	14-15 h	0:33:19	0:38:20	Breaking news: culture, history (presenter, people)	0:05:01	94
ETB1_15	20-21h	0:00:10	0:05:17	Headlines with reporters	0:05:07	86
ETB1_16	20-21h	0:51:20	0:56:29	Weather forecast	0:05:09	124
ETB1_17	20-21h	0:09:35	0:14:35	News	0:05:00	82
ETB1_18	20-21h	0:42:58	0:48:05	Sports	0:05:07	106
ETB1_19	9:10-9:40 h	0:25:10	0:30:00	Iparraldearen orena	0:04:50	75
ETB1_20	9:10-9:40 h	0:04:48	0:10:04	Iparraldearen orena	0:05:16	95
					97 minutes analysed	1737 subtitles analysed

Table 3. Summary of samples in the corpus

In March 2020, the Elhuyar Foundation launched ADITU, a speech recognition software that automatically transcribes and subtitles audio and video files. Its design was based on artificial intelligence together with neural network technology and works with both pre-recorded and live audio and video files (Aztiria, Jauregi and Leturia 2020). In its early

version, the software recognised Basque and Spanish speech separately, but the latest version can also be used to transcribe and subtitle bilingual files. As with many other programmes, ADITU works with both an acoustic model and a language model. Put simply, the acoustic model identifies phonemes and the language model determines what word sequences are more frequent based on different texts that feed the programme. ADITU works with 4- or 5-word sequences to determine which set of words is more probable. Texts that feed the language model of the programme are mostly journalistic and the language model is updated with more texts at various times each year (Leturia, personal communication, December 14, 2022).

As mentioned above, we contacted the staff of Elhuyar in charge of the programme ADITU. With the aim of having subtitles of as high quality as possible for the Basque segments, the monolingual version of ADITU was chosen. All samples (20) were sent in .ts format[6]. Elhuyar provided *raw* .srt files[7] (with no segmentation), *arranged* .srt files (with segmentation in two lines, maximum, with the maximum of 37 characters per line established by ADITU) and .txt files. In this particular study, arranged non-edited .srt files were analysed. One sample (ETB1_1) was not recognised by the programme ADITU and we have neither data nor subtitle files from that sample[8]. Therefore, in this analysis, only 19 samples could be analysed, making a total of 97 minutes and 1737 subtitles. In this case study, pop-on subtitles containing one or two lines were provided. Throughout the analysis, we refer to the term subtitle as a one- or two-line pop-on block. It should be noted that ADITU can also create roll-up subtitles for live broadcasts. However, although the subtitles analysed in the present article have been generated for programmes broadcast in a live setting, the insufficient availability of live subtitling in Basque has rendered pop-on block subtitles, as if they were for recorded programmes. Consequently, the results presented here may differ from those obtained if broadcast live subtitles were analysed, as delay or editing errors, among other factors, would need to be considered.

For the data to fit into the NER model of analysis, transcription of all samples was performed manually. Only subtitles related to segments entirely in Basque are analysed in the present contribution. Therefore, segments containing audio/subtitles in Spanish or other languages are excluded from the analysis.

The quantitative analysis is based on the NER model, which establishes a threshold of 98% as the minimum AR for subtitles to be considered comprehensible. Following the NER model, in this analysis minor recognition errors (R-M) are those which do not hinder the correct comprehension of the audiovisual content. Standard recognition errors (R-ST) are those that the viewer will be able to recognise as an error but which will not allow for a correct comprehension of the audiovisual content or may create confusion for the viewer. Serious recognition errors (R-SE) are those considered unlikely to be noticed by the viewer and will cause

viewers to infer a different understanding than that heard in the audio. Finally, we consider corrected recognition (CR) the omission or reduction of information that is redundant on the audio, such as markers of oral speech or unintended repetitions of words that the software suppresses autonomously.

Labelling of the errors and CRs was carried out manually in Excel templates used specifically for the analysis using the NER model. Following data provided by the Excel sheet and information provided by Igor Leturia, the next sections offer a quantitative and qualitative analysis of the automatic subtitles in Basque for news events created with the ADITU monolingual (Basque) version by Elhuyar using the developmental stage of May, 2022.

4. Results and discussion

In this section, we include a quantitative and qualitative analysis of the 97 minutes and 1737 automatic subtitles in Basque following the NER model. Firstly, we examine the AR of automatic subtitles in Basque, with and without considering punctuation errors and taking into account the NER threshold for acceptable recognition accuracy. Secondly, we look at the types and severity of errors, and we discuss the possible explanations for errors encountered. Thirdly, we examine the average subtitle speed of each programme and compare it with the maximum recommended by the UNE standard in Spain: 15 characters per second (CPS) (AENOR 2012). Finally, we offer a qualitative analysis of the data bearing in mind the particularity of the Basque language and the programmes analysed, and looking at the possible comprehension of automatically generated subtitles beyond the numbers provided by the NER model.

4.1 Accuracy rates

The average AR for the automatic subtitles of this contribution is 94.63%, if we take all errors into account, and 96.09% if we leave punctuation errors aside. The highest AR is 98.01% (99.12% without punctuation errors) and the lowest AR is 89.12% (90.84% without punctuation errors). Average AR for automatic subtitles in our sample, as well as highest and lowest AR are below any data reported by Romero-Fresco and Fresno (2023) for English. AR per each sample can be seen in Figure 2.
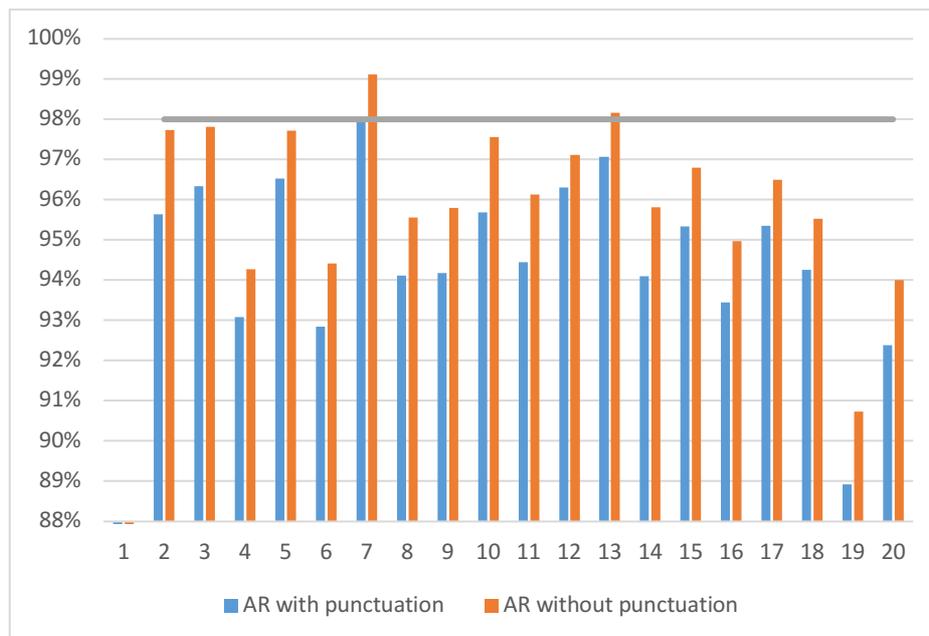
Figure 2. Accuracy rates per sample with and without punctuation

The NER threshold for acceptable recognition accuracy is set at 98% (shown in grey). Only sample 7 (AR with punctuation errors = 98.01%, AR without punctuation errors = 99.12%) and sample 13 (AR without punctuation errors = 98.17%) can be considered *fair* or *good* taking into account only the quantitative measurements by the NER model. Both samples contain the delivery of news by anchors and reporters in which the speech they deliver is prefabricated, very well structured and with standard diction. Special attention should be given to samples 19 and 20, which score considerably lower than the rest, as these are the only two samples that contain news and spontaneous speech (interviews) of people from *Iparralde* (French Basque Country). The AR of sample 19 is especially low due to the fact that most of the sample consists of an interview with a farmer from *Iparralde*, with a non-standard accent and dialect. Excluding samples 7, 19 and 20, all other samples score between 92.84% and 97.06% (94.41% and 98.17% without punctuation errors) which, although it might not *translate* into acceptable data for the NER model, seem satisfactory numbers and are similar to the 95% reported in Aztiria, Jauregi and Leturia (2020) for this same software. These numbers should be read bearing in mind the minoritised situation of Basque; the short trajectory of voice recognition software in this language (and, especially, of the software ADITU), when compared to hegemonic languages such as English or Spanish; and the recognition difficulties that may arise from those accents and dialects that differ from the unified Basque (*euskera batua*)[9] (see Aztiria, Jauregi and Leturia 2020).

If we take the sample with the highest AR (sample 7), we can infer which characteristics of the audiovisual product favour the absence of recognition errors and, therefore, a higher AR and comprehension of

subtitles and the audiovisual text. When videos contain fewer people who are speaking normalised Basque (*euskera batua*), a formal register, clear and standard diction and prefabricated linguistic code, the recognition errors observed are mostly related to punctuation. In fact, sample 7 contains only 17 non-punctuation errors (8 R-M and 9 R-ST) and 30 punctuation errors (29 R-M and 1 R-SE).

4.2 Errors and corrected recognition

In the 19 samples analysed, 1967 errors were found (1102 R-M, 768 R-ST, 97 R-SE), divided into the different samples as shown in Table 4. On average, our sample included 20.2 errors/minute, which is below the rate reported by Romero-Fresco and Fresno (2023) for automatic closed captions in English.

ID	Total Errors	R-M	R-ST	R-SE	CR
ETB1_1	-	-	-	-	-
ETB1_2	90	53	34	3	3
ETB1_3	66	38	23	5	0
ETB1_4	160	67	83	10	7
ETB1_5	65	42	19	4	2
ETB1_6	136	72	59	5	3
ETB1_7	47	37	9	1	2
ETB1_8	119	71	46	2	0
ETB1_9	106	61	38	7	0
ETB1_10	97	67	20	10	6
ETB1_11	84	50	27	7	0
ETB1_12	66	40	24	2	5
ETB1_13	52	35	17	-	1
ETB1_14	117	60	54	3	3
ETB1_15	88	67	18	3	0
ETB1_16	164	100	51	13	5
ETB1_17	70	35	29	6	0
ETB1_18	126	76	47	3	4
ETB1_19	162	48	107	7	0
ETB1_20	152	83	63	6	4
TOTAL	1967	1102	768	97	45

Table 4. Number of errors and CR per sample

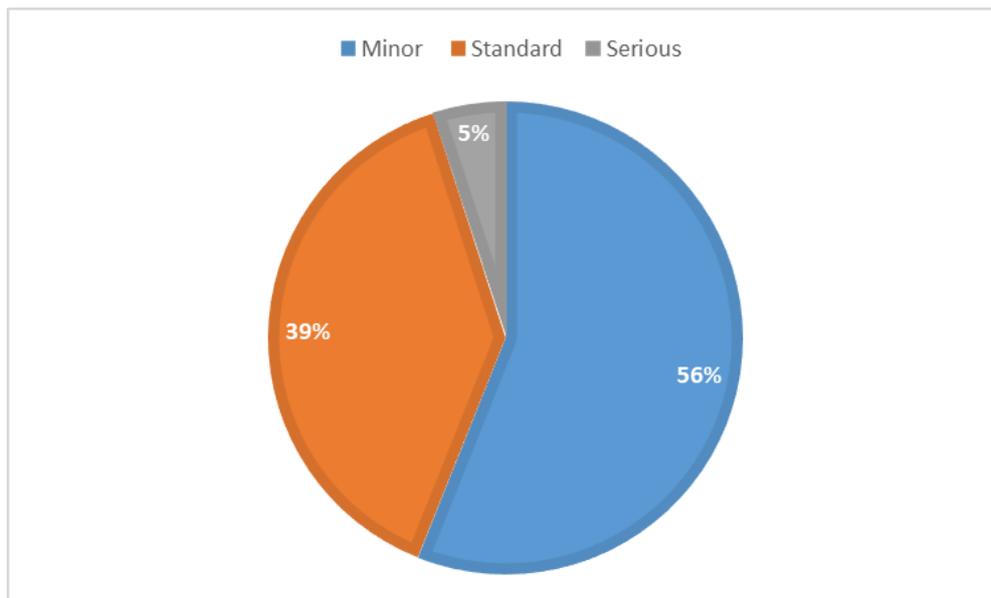The severity of those errors is represented in percentages as shown in Figure 3.

Figure 3. Severity of all errors

As stated previously, minor errors (R-M) are those that the programme has not been able to recognise correctly but that do not hinder the correct comprehension of the subtitle and the audiovisual content, for example:

Sample	Transcription	Subtitle	Meaning of transcription	Meaning of subtitle
ETB1_5	*Bestela, ba suministro**rik** gabe geratuko direlako*	*Bestela, ba suminist ro**ri** gabe geratuko direlako*	*Because otherwise they will run out of supply*	*Because otherwise to supply they run out*

Table 5. Example of a minor error (sample ETB1_5)

Standard errors (R-ST) are those which the viewer will be able to recognise as such, but will not allow understanding the subtitle and the audiovisual content, and will likely trigger confusion in the viewer, for example:

Sample	Transcription	Subtitle	Meaning of transcription	Meaning of subtitle
ETB1_5	***Mendebaldeko** zigorrek salmenta hauei eragiten dietenean,*	***Beno, aldeko** zigorrek salment a hauei eragiten dietenea n,*	*When Western sanctions affect these sales,*	*Well, when favourable sanctions affect these sales,*

Table 6. Example of a standard error (sample ETB1_5)

Serious errors (R-SE) will not be identifiable by the viewer as an error and will make the viewer understand something different from the audio, for example:

Sample	Transcription	Subtitle	Meaning of transcription	Meaning of subtitle
ETB1_4	*euria egin egingo duela, baina denbora gutxiagoan eta azkoz* ***zakarrago***	*euria egin egingo duela, baina denbora gutxiagoan eta azkoz* ***azkarrago***	*that it will rain, but in less time and a lot more heavily*	*that it will rain, but in less time and a lot faster*

Table 7. Example of a serious error (sample ETB1_4)

It is worth noting that a remarkable part of these errors (726 errors; a 36.9%) are punctuation errors. For this analysis in particular, we understand as punctuation errors all those errors related with punctuation itself (a missing full stop at the end of a sentence, missing commas, commas that should be full stops, etc.); ortho-typographic errors that are not a consequence of a poor recognition itself, but are related to the language models inserted in the programme (upper-case and lower-case errors, accents[10], changes of a C by a K in proper names, etc.); and errors related to misidentification of speakers. According to our conversation with Igor Leturia, a new punctuation system has been recently implemented and the number of punctuation errors has allegedly decreased (Leturia, personal communication, December 14, 2022). This new punctuation system has been trained following the ElhBERTeu model (explained in Urbizu *et al*. 2022). Hidalgo (2023) compared the punctuation output in both versions and concluded that while the number of punctuation errors is still high, the use of commas is more prevalent in the newer version and the programme detects the duration of the speaker's pause with greater accuracy. Considering this, we find it useful to look at the severity of errors excluding punctuation errors (a total of 1241 errors) and the severity of the punctuation errors (726 errors). These results are shown in the Figures 4 and 5 below.
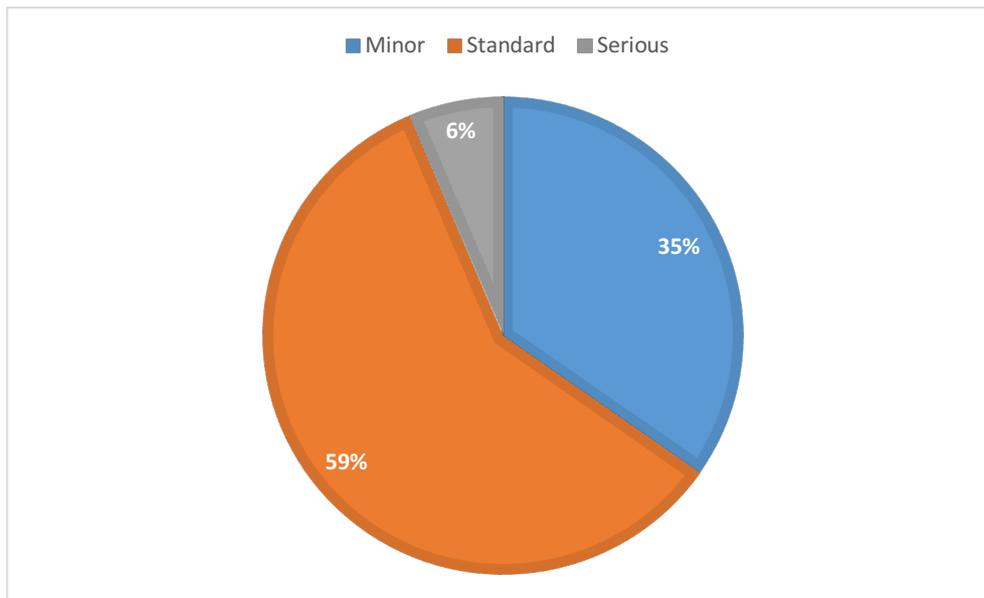
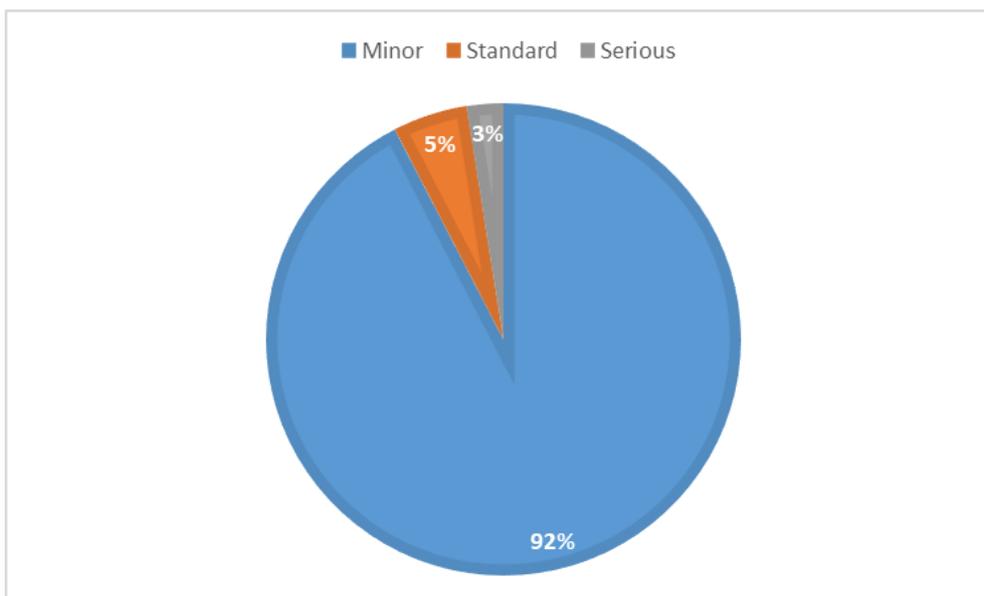**Figure 4. Severity of errors without punctuation**

Figure 5. Severity of punctuation errors

As can be seen, the majority of punctuation errors (92%) are minor errors that do not hinder comprehension of audiovisual content. When comparing these results with the results of other languages using the NER model[11], the high percentage of standard errors is noticeable, both taking into account the severity of all errors (see Figure 3) and the severity of errors without punctuation (see Figure 4).

Some of these errors are due to the number of proper nouns (for instance, regarding the Ukrainian conflict in sample 17 or the Roe versus Wade case in sample 3); others are due to the variety of dialects and dictions that

the programme has difficulty in recognising (such as those in samples 19 and 20 taken from *Iparraldearen orena*). A third explanation for the number of standard errors is the audio that can be heard underneath a voice-over translation (such as in sample 17, where there is voice-over for German). ADITU tries to recognise that audio segment which results in errors that generate confusion in the viewer. When we asked the developers of ADITU about these errors, they confirmed that dialects, proper nouns and bad quality of audio (which may expand to voice-over tracks) are the main reasons for recognition errors in this software (Aztiria, Jauregi and Leturia 2020, Leturia, personal communication, December 14, 2022).

ADITU also applies corrected recognition (CR), the synthesis or omission of information that the software identifies as unnecessary and automatically suppresses from subtitles. In most instances, it corresponds to marks of orality and word repetitions, as shown in Table 8. Nonetheless, it should be pointed out that the software sometimes fails to recognise certain marks of orality as such, resulting in errors at all levels of severity (see Table 9). ADITU was not specifically trained to dismiss repetitions or marks of orality. CRs in ADITU occur when the programme goes from the acoustic model to the language model. As the language model is mostly fed with journalistic texts (in which such marks of orality are not usual), the programme dismisses those words (Leturia, personal communication, December 14, 2022).

Sample	Transcription	Subtitle	Meaning of transcription	Meaning of subtitle
ETB1_10	***Ba*** *Araban, esate baterako,* ***ba*** *Bernedo aldean bi gradu baino ez*	*Araban, esate baterako, Bernedo aldean bi gradu baino ez*	*Well, in Alava, for example, well, in the area of Bernedo only two degrees*	*In Alava, for example, in the area of Bernedo only two degrees*
ETB1_20	*dela* ***aintzina***, *aintzina garatzea.*	*dela aintzina garatzea*	*It is to develop forward, forward.*	*It is to develop forward.*

Table 8. Example of CR

Sample	Transcription	Subtitle	Meaning of transcription	Meaning of subtitle
ETB1_4	***Eh***, *kolore zuria izango litzateke*	***Ez***, *kolore zuria izango litzateke*	*Eh, it would be the white colour*	*No, it would be the white colour*

Table 9. Example of a mark of orality that results in an error

4.3 Subtitle speed[12]

In this study, we analysed subtitle speed in CPS. Figure 6 shows the average CPS per sample:
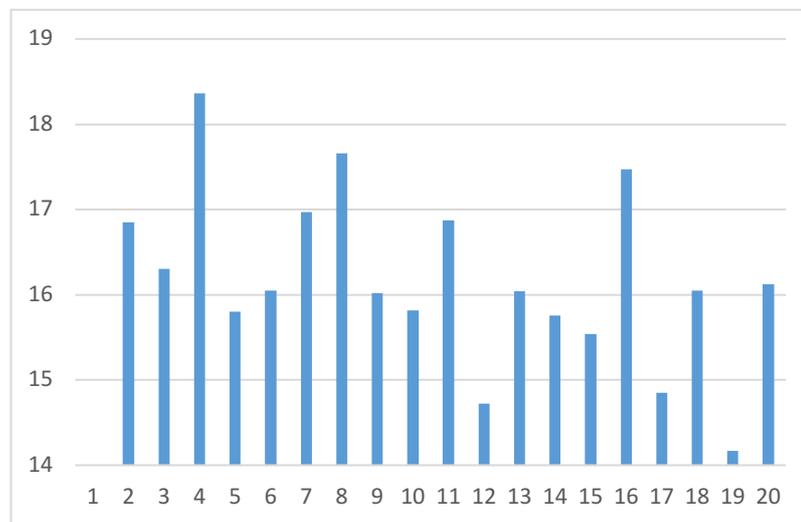

Figure 6. Subtitling speed (in CPS) per sample

ADITU does not set a maximum CPS for its automatically generated subtitles, so subtitle speed will vary depending on the speech rate and the information displayed on the subtitles (Leturia, personal communication, December 14, 2022). The maximum subtitle speed recommended in Spain (for Spanish) by the Standard UNE 153010 (AENOR, 2012) is 15 CPS. Only 3 of the 19 samples analysed are below that recommendation. Nevertheless, it is of note that this is a standard for Spanish and it is difficult to know if that same subtitle speed would be the recommendation for a language such as Basque. Despite this, research shows that setting a maximum recommended speed at a specific number is neither easy nor advisable. Many researchers and studies suggest that the maximum recommended depends on factors such as the information that needs to be inferred from the image, the amount of visual action on a given moment, the complexity of the vocabulary and the syntax used in each subtitle, the audience at which the audiovisual product is aimed, the complexity of the topics tackled in the video, the audiovisual genre, etc. In a recent interview, researcher Agnieszka Szarkowska (episode 32, podcast *En Sincronía*, July 2022) stated that the maximum recommended speed is usually at some point between 12 and 20 CPS, depending on the factors mentioned above. Taking this into consideration, and without a reception study that could actually provide us with data on the comprehension of subtitles and audiovisual text by deaf people, we cannot state that the subtitle speed of these samples is inadequate.

Table 10 shows, per sample, the minimum and maximum values of CPS, the average CPS and the percentage of subtitles exceeding 15 CPS (which would go against recommendations of AENOR), 20 CPS (which would go

against general recommendations for different languages, audiovisual media, genres and audiences, from academics) and 25 CPS.

Sample	Minimum CPS	Maximum CPS	Average	≤ 15 CPS	>15 ≤ 20 CPS	> 20 ≤ 25 CPS	> 25 CPS
ETB1_1	-	-	-	-	-	-	-
ETB1_2	9.02	26.58	16.85	32.26%	49.46%	13.98%	4.30%
ETB1_3	9.92	44.00	16.31	36.78%	55.17%	6.90%	1.15%
ETB1_4	10.714	33.33	18.36	15.93%	55.75%	22.12%	6.20%
ETB1_5	2.20	30.56	15.80	47.31%	38.71%	9.68%	4.30%
ETB1_6	4.27	26.36	16.05	36%	52%	10%	2%
ETB1_7	8.78	28.95	16.97	29.60%	54.08%	13.26%	3.06%
ETB1_8	5.32	34.48	17.66	25%	49.07%	23.15%	2.78%
ETB1_9	7.87	31.43	16.02	38.95%	52.63%	7.37%	1.05%
ETB1_10	8.30	27.88	15.82	41.84%	44.90%	11.22%	2.04%
ETB1_11	9.48	36.54	16.87	27.45%	57.84%	12.75%	1.96%
ETB1_12	5.88	22.69	14.72	53.49%	39.53%	6.98%	0.00%
ETB1_13	7.02	30.00	16.04	37.76%	47.96%	10.20%	4.08%
ETB1_14	7.32	53.85	15.76	49.46%	38.71%	8.60%	3.23%
ETB1_15	4.65	24.56	15.54	43.49%	44.31%	12.5%	0%
ETB1_16	5.45	27.74	17.47	27.42%	45.16%	24.19%	3.23%
ETB1_17	7.74	33.33	14.85	56.63%	36.14%	3.61%	3.61%
ETB1_18	6.22	31.25	16.05	40.54%	45.95%	11.71%	1.80%
ETB1_19	7.69	26.58	14.17	64%	33.34%	1.33%	1.33%
ETB1_20	4.26	30.23	16.12	37.11%	48.45%	11.34%	3.10%

Table 10. Minimum CPS, maximum CPS, average CPS and percentage of subtitles exceeding 15, 20 and 25 CPS

In a more qualitative analysis, and factoring in what has been said for other languages (see, for instance, Romero-Fresco 2016, which contains an analysis of speech rate and subtitle speed in words per minute of UK programmes), the average CPS of these samples does not appear overly high. Only one sample (ETB1_4) is above 18 CPS and 16 out of 19 samples are below 17 CPS. Of interest is that the three samples that are above 17 CPS deal with the weather forecast and sports sections of the news programmes. The speech rate of these sections in news programmes tends to be above other subgenres and therefore the average CPS of these samples is also higher.

When compared with guides other than AENOR (2012) that do not revolve around live subtitling, the average CPS values of the analysed samples are close to the maximum CPS values proposed, for instance, in the style guide for Basque subtitles by the streaming platform Netflix[13], which sets a maximum of 17 CPS for adults and 13 CPS for children.

All the samples register, at some point, subtitles above 20 CPS that would be considered too fast for any type of audience, medium and audiovisual genre. According to Table 10 above, the sample with the fewest subtitles exceeding 20 CPS is EITB1_19, with 1.33% of the subtitles between 20 and 25 CPS and 1.33% above 25. At the other end of the scale is ETB1_4, with 22.12% of the subtitles between 20 and 25 CPS and 6.20% above 25

CPS. Thus, almost one out of three subtitles in ETB1_4 is above what is considered readable for any audience and any audiovisual medium or genre. The peaks reached by all the samples analysed are equally illegible for any audience and medium, the lowest peak being found in sample ETB1_12, reaching 22.69 CPS, and the highest peak in sample ETB1_14, with a maximum of 53.85 CPS.

4.4 Qualitative analysis and possibilities for improvement

This section offers a qualitative analysis and prospects of improvement based on the quantitative data presented above. Moreover, different areas of possible improvement have been identified. Here, we will take a look into language-related issues and more technical issues. This qualitative analysis is not without limitations, as we did not conduct a reception study with deaf people. Nevertheless, these areas of interest were discussed in our online interview with Igor Leturia, in December, 2022.

4.4.1 Language-related issues

The programme ADITU tends to standardise spontaneous speech and speech coming from *euskaldunzaharrak*, native Basque speakers who generally speak in a given *euskalki* or dialect. Table 11 shows some examples of this standardisation found in our corpus:

Transcription	Standardised language in subtitles	Meaning
segundu	segundo	second
dan	den	is (3rd person singular + indirect question declension)
ta	eta	and
azkenian	azkenean	at the end / finally
hoiek	horiek	they / them

Table 11. Examples of speech standardisation

This standardisation has both advantages and disadvantages. On the one hand, and taking into account the minoritised situation of Basque, it benefits language standardisation and normalisation and it facilitates subtitle comprehension by a larger population and the alphabetisation in Basque. On the other hand, it might hinder access to content, as it does not provide information on how speech is delivered. The fact that a person speaking in Basque says "dan" or "azkenian", instead of standardised versions of those words, provides information about their relationship with the language, their origins, the dialect they might prefer using, etc. The continuous standardisation and normalisation when text goes from oral to written language might limit access to some nuances related to the use of a language. In our corpus, nevertheless, this standardisation is not consistent, and excerpts with non-standardised language have been found. Consideration should be given to the future provision of access to subtitles in Basque and the advantages and disadvantages of

normalisation and standardisation. Igor Leturia and the staff working in ADITU and at Elhuyar believe subtitles should be standardised to *batua*, as that is the norm in current subtitles practices (Leturia, personal communication, December 14, 2022).

It has come to our attention, also, that frequently used proper nouns in Basque (such as *Estatu Batuak* [United States] in sample 3 or *Athletic* in sample 8) are often misrecognised in Basque. Other misrecognitions of proper nouns have to do with the name of the reporters, names of places related to the war between Russia and Ukraine, Spanish names related to politics (in which, following the *Libro de estilo de EiTB: información y actualidad* style guide [Martín Sabarís *et al*. 2016], often diacritics are not employed), or proper names related to latest news (Roe versus Wade, for example). Errors with proper nouns in ADITU are one of its main problems but are not easy to overcome. On the one hand, proper nouns related to the latest news will imply updating the language model with a regularity nowadays not possible for the developmental stage of ADITU. Whereas the case and declension system of Basque would imply not only adding the non-declined word (for instance, Járkov), but also all the possibilities of declension for such word (Járkovtik, Járkovera, Járkoven, etc.) happening in different word-sequences for the language model to be able to choose correctly from the different possibilities that the acoustic model might offer (Leturia, personal communication, December 14, 2022). Regarding frequently used proper nouns in Basque (such as *Estatu Batuak* and *Athletic*) Igor Leturia stated that exceptions to the norms might be inserted in the programme. That would be the case of the pronunciation of *Athletic*, the writing of which does not follow Basque conventions and might pose difficulties for the acoustic model, which would have to be trained specifically for that word and its declension cases. After our conversation, ADITU developers stated that they will evaluate the special cases of *Estatu Batuak* and *Athletic* and will try to improve the recognition of those two nouns.

Samples containing a weather forecast, for example, show a considerable number of standard errors. The most noticeable one is sample 4, in which there are a total of 83 standard errors, the second highest number of the corpus. Some standard errors of sample 4 have to do with the topic (weather) and the specific language used in this type of subgenre:

Transcription	Subtitle	Meaning of transcription	Meaning of subtitle
orain, egunez, ba zeru estali horrekin…	orain egunez bazenu eskari horrekin …	now, in daytime, with an overcast sky…	If you now had by day* With that request…
zeru grisa	sailburu grisa	grey sky	Grey councillor
borraska hegoan dagoenez	ordezkari egon dagoenez	as the storm is in the South	As has been representative*

Table 12. Errors related to weather forecast terminology

According to Igor Leturia (personal communication, December 14, 2022), the weather forecast should not pose more problems than any other subgenre, as it has been equally trained. A hypothesis posed by Leturia is that problems with recognition and standard errors in weather forecasts might be due to speech rate and not to the language model (an argument also stated in Aztiria, Jauregi and Leturia 2020). A more in-depth analysis on this subgenre might be useful to confirm such a hypothesis.

Finally, taking into account that Basque is an agglutinating language with a thorough case system, a declension of a word might be determinant to the sense of a sentence, and an error declension (which could be an error of recognition of just one letter) may change the meaning of a sentence completely. In sample 8, "Osasunak" (proper noun in ergative case) is recognised and subtitled as "Osasuna" (which acts either as a proper noun in nominative case or as a proper noun in accusative case), thus turning what should be the subject of the sentence into its direct object. Here it is labelled as a minor error, but a change in just one letter could also lead to standard or serious errors depending on the audiovisual context and the text surrounding that error. Therefore, a future hypothetical training of respeakers in Basque would need to pay special attention to the correct pronunciation of declensions and conjugation of verbs.

4.4.2 Technical issues

Here, we will deal with punctuation, speaker identification and the special case of "ehuneko" (percentage). Punctuation errors are, in our sample, mostly minor errors. On some occasions, lower case letters are used after a full stop or after a colon in the speaker identification. On other occasions, the full stop is missing in sentences. Other errors that have been analysed under the punctuation category such as accents[14] ("Sanchez", for example) or upper-case lettering (some proper nouns appear with all letters capitalised, for instance) are also mostly minor errors. Nevertheless, the high number of punctuation errors leads to the

question of whether there are possibilities for improvement in this matter. A new version of ADITU is now available with a new punctuation system (Leturia, personal communication, December 14, 2022). In a recent comparison of two versions (2022 against 2023), Hidalgo (2023) concluded that, although the total number of punctuation errors decreased, the results and severity of errors regarding punctuation in both versions turn out to be very similar.

Speaker identification is predicted based on acoustic information and silence gaps between interventions (Leturia, personal communication, December 14, 2022). Sometimes, this leads to (1) unnecessary identification (because the tag appears again after a short period of silence even if the same person is speaking); (2) identification in subtitles that may lead to confusion (because the tag appears in the middle of a sentence when a silence gap occurs) and (3) on some occasions serious errors. For example, the same name tag (HZLR1, for instance) is sometimes used in a 5-minute-sample to identify two different speakers. Most of these errors are minor, but current criteria to identify speakers also lead to standard and serious errors. According to Leturia (personal communication, December 14, 2022) some of the errors in speaker identification have been solved in the new version of ADITU. In this regard, Hidalgo (2023) concluded that, despite the fact that the use of tags has improved in the new version, inaccuracies continue to be detected and a significant portion of the speaker identification tags used by ADITU are still incorrect.

Regarding the special case of "ehuneko" (per cent), it came to our attention that this word is not subtitled at all, which results in serious errors (because percentage is understood as an absolute number) and standard errors (because sentences lose meaning and the subtitle creates confusion). In some occasions, as with sample 5, these percentages are shown on a screen behind the presenters and errors are then marked as minor, because they are detectable and correct information can be inferred from the image and audiovisual context. Nevertheless, this is only possible thanks to information redundancy; the subtitle, by itself, does not allow for inferring correct information. According to Leturia (personal communication, December 14, 2022), the problem with "ehuneko" was an exceptional problem during the dates in which our .srt were generated. This problem was solved in the following months and the new version of ADITU does not have those errors (Hidalgo, 2023).

5. Final remarks

This contribution has analysed automatic subtitles generated by the monolingual (Basque) version of ADITU, a speech recognition software developed by Elhuyar and launched in 2020. Although quantitative data does not reach the threshold to establish the quality of recognition as fair or comprehensible taking into account the NER model, both a quantitative and a qualitative analysis show that results are promising although there

are possibilities for improvement regarding language and technical issues. Some of these improvements have allegedly been implemented in a new version of ADITU, partially analysed in Hidalgo (2023). These rapid improvements and their implementation are expected to significantly improve the AR of speech recognition in Basque in the near future.

The analysis permits us to foresee some factors that may have influence on the accuracy rate, such as:

- Batua (standardised language) or the use of euskalkiak (dialects)
- Register
- Diction of speakers
- Prefabricated discourse or spontaneity
- Number of speakers
- Speakers being anchors, reporters or informants
- Background noise and language interference in the audio
- Proper nouns

The interview with Igor Leturia (personal communication, 2022) confirmed that proper nouns, bad quality of audio and non-standardised language pose the most difficulties in speech recognition for ADITU. Our findings based on the NER model and the interview with Igor Leturia also support statements by Pérez Cernuda (2022) on the evaluation of bilingual automatic captions and a previous publication on the launching of ADITU (Aztiria, Jauregi and Leturia 2020). These factors should be researched in depth, and isolated if possible, in order to identify the relationship between the variation of these factors and the AR in the NER model.

Basque society is plurilingual, and so are their audiovisual products (Larrinaga 2019). This contribution has only analysed excerpts in Basque, which give us an idea of the current situation of the recognition software, but does not account for the reality of the society nor the products. Further analysis that takes into account programmes that combine Basque, at least, with Spanish and French, and software that is compatible with that reality would be of great benefit in the future.

In addition, a second round of analysis of these same samples when improvements have been applied to the software will let us analyse the evolution and development of the recognition programme. Also, other technologies (such as the ones developed by Vicomtech or Google) should also be taken into consideration in further analysis.

Funding information

References

- **ADITU — Elhuyar's speech recognition system** (2024). https://aditu.eus/ (consulted 06.09.2023).

- **AENOR** (2012). *Norma UNE 153010: Subtitulado para personas sordas y personas con discapacidad auditiva. Subtitulado a través del teletexto*. AENOR.

- **Aztiria, Josu, Saroi Jauregi and Igor Leturia** (2020). "Elhuyarren Aditu hizketa-ezagutzailea, beste aurrerapauso bat euskararen garapen digitalean." *Senez* 51, 259-266.

- **Barambones, Josu** (2009). *La traducción audiovisual en ETB-1: un estudio descriptivo de la programación infantil y juvenil*. PhD thesis. Universidad del País Vasco/Euskal Herriko Unibertsitatea.

- **Cabanillas, María Candelas** (2016). *La traducción audiovisual en ETB2: estudio descriptivo del género western.* PhD thesis. Universidad del País Vasco/Euskal Herriko Unibertsitatea.

- **Cenoz, Jasone and Josu Perales** (1997). "Minority language learning in the administration: Data from the Basque Country." *Journal of Multilingual and Multicultural Development* 18*,* 261–270.

- **Costa, Pere-Oriol** (1986). *La crisis de la televisión pública.* Barcelona: Paidós.

- **Dumouchel, Pierre, Gilles Boulianne, and Julie Brousseau** (2011). "Measures for quality of closed captioning." Adriana Şerban, Anna Matamala and Jean Marc Lavaur (eds) (2011). *Audiovisual translation in close-up: Practical and theoretical approaches.* Bern: Peter Lang, 161–172.

- **Fresno, Nazaret and Katarzyna Sepielak** (2020). "Subtitling speed in Media Accessibility research: some methodological considerations." *Perspectives* 30(3), 415-431.

- **Euskararen erakunde publikoa, Eusko Jaurlaritza and Nafarroako Gobernua** (2023a). *VII. Inkesta Soziolinguistikoa 2021: Euskal Autonomia Erkidegoa*. Donostia: Euskarabidea.

- **Euskararen erakunde publikoa, Eusko Jaurlaritza and Nafarroako Gobernua (**2023b). *VII. Inkesta Soziolinguistikoa 2021: Iparralde*. Baiona: Euskarabidea.

- **Euskararen erakunde publikoa, Eusko Jaurlaritza and Nafarroako Gobernua (**2023c). *VII. Inkesta Soziolinguistikoa 2021: Nafarroako Foru Komunitatea*. Donostia: Euskarabidea.

- **Hidalgo, Jon** (2023). *Accesibilidad en traducción e interpretación: análisis de la calidad de subtítulos automáticos en euskera generados con reconocimiento del habla*. Unpublished bachelor's thesis. Universidad del País Vasco/Euskal Herriko Unibertsitatea.

- **Larrañaga, Jotxo and Carmelo Garitaonandía** (2012). *Convergencia mediática en Euskal Irrati Telebista (EITB).* Bilbao: Servicio de publicaciones UPV/EHU.

- **Larrinaga Larrazabal, Asier** (2019). *Euskal telebistaren sorrera, garapena eta funtzioa normalizazioaren testuinguruan*. PhD thesis. Universidad del País Vasco/Euskal Herriko Unibertsitatea.

- **Lasagabaster, David** (2010). "Bilingualism, Immersion Programmes and Language Learning in the Basque Country." *Journal of Multilingual and Multicultural Development* 1(5), 401-425.

- **Martín Sabarís, Rosa M., María José Cantalapiedra, Begoña Zalbidea Bengoa, and Andoni Aranburu** (2016). *Libro de estilo de EiTB: información y actualidad*. Bilbao: EiTB.

- **Pedersen, Jan** (2017). "The FAR model. Assessing quality in interlingual subtitling." *The Journal of Specialised Translation* 28, 210-229.

- **Pérez Cernuda, Carmen** (2022). "Subtitulado automático bilingüe: la idea es sencilla, la solución, no tanto." *Panorama Audiovisual.com*. https://www.rtve.es/rtve/20221118/territoriales-tve-subtitulado-automatico-bilingue/2409497.shtml (consulted 06.09.2023).

- **Romero-Fresco, Pablo** (2012). "Quality in Live Subtitling: The Reception of Respoken Subtitles in the UK." Aline Remael, Pilar Orero and Mary Carroll (eds) (2012). *Audiovisual Translation and Media Accessibility at the Crossroads: Media for All 3,* 25-41.

- **Romero-Fresco, Pablo** (2016). "Accessing communication: The quality of live subtitles in the UK." *Language & Communication* 49, 56-69.

- **Romero-Fresco, Pablo and Juan Martínez Pérez** (2015). "Accuracy Rate in Live Subtitling: The NER Model." Jorge Díaz-Cintas and Rocío Baños (eds) (2015). *Audiovisual Translation in a Global Context. Palgrave Studies in Translating and Interpreting*. London: Palgrave Macmillan, 28-50.

- **Romero-Fresco, Pablo and Nazaret Fresno** (2023). "The accuracy of automatic and human live captions in English." *Linguistica Antverpiensia, New Series: Themes in Translation Studies*, 22, 114-133.

- **Tamayo, Ana and Irene de Higes** (2023). "Automatic speech recognition in Spain: the Basque and Catalan case." Paper presented at *The 8th International Symposium on Live Subtitling and Accessibility* (Universitat Autònoma de Barcelona, 19 April 2023).

- **Urbizu, Gorka, Iñaki San Vicente, Xabier Saralegi, Rodrigo Agerri, and Aitor Soroa** (2022). "BasqueGLUE: A Natural Language Understanding Benchmark for Basque". Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Jan Odijk and Stelios Piperidis (eds) (2022). *Proceedings of the Thirteenth Language Resources and Evaluation Conference*. Marseille: European Language Resources Association, 1603–1612.

- **Zuazo, Koldo** (1995). "The Basque Country and the Basque Language: An overview of the external history of the Basque language." José Ignacio Hualde, Joseba A. Lakarra, and R. L. Trask (eds) (1995). *Towards History of the Basque Language.* Amsterdam: John Benjamins Publishing, 5-30.

- **Zuazo, Koldo** (2014). *Euskalkiak.* Donostia-San Sebastián: Elkarlanean.

Biographies

Ana Tamayo is an Associate Professor at the University of the Basque Country (UPV/EHU). Currently, she is a member of the research groups TRALIMA/ITZULIK (GIU21/060, UPV/EHU) and collaborates with TRAMA (Universitat Jaume I) and GALMA (Universidade de Vigo). Her research interests focus on audiovisual translation and accessibility in different modalities. She is especially interested in contributing to the research on media accessibility and minoritized languages, mostly Basque and Sign Language(s).
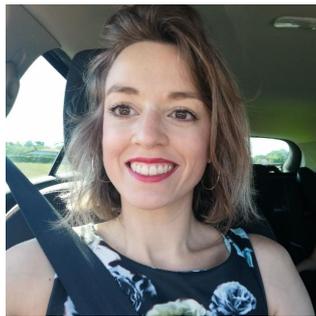
ORCID: 0000-0002-5419-5929

Email: ana.tamayo@ehu.eus

Alejandro Ros Abaurrea holds a Bachelor's Degree in Translation and Interpreting, a Master's Degree in Language Acquisition in Multilingual Settings (both from the UPV/EHU), and a Master's Degree in Compulsory Secondary Education and Baccalaureate, Vocational Training and Language Teaching from the University of La Rioja. Between 2019 and 2022, he worked as a Researcher in Training (PIF) at the UPV/EHU to complete his doctoral thesis on the translations of Leonard Cohen's musical works within the context of Spanish culture. Currently, he teaches at the UPV/EHU and is a member of the research group TRALIMA/ITZULIK. His research interests include corpus-based translation studies, musicology, audiovisual media, and cultural studies.

ORCID: 0000-0003-4142-8794

Email: alejandro.ros@ehu.eus

Notes

[1] The French Basque Country that straddles the western Pyrénées-Atlantiques.

[2] It should be noted that, until 1982, there were only two public state-channels where the use of Basque was limited to *Telenorte,* a weekly report broadcast on the Spanish national broadcast, TVE (Barambones 2009: 72).

[3] From a file of approximately 30 minutes containing approximately 500 subtitles, only around 1 minute (20 subtitles approximately), were in Basque. In addition, these did not report actual live events, but summaries of news or news headlines. This is also supported by statements by Pérez Cernuda (2022).

[4] This was possible thanks to the collaboration of the research group NOR, we are especially thankful to Josu Amezaga, Marijo Deogracias and Hibai Castro for providing us with the requested recordings from the TDT in .ts format.

[5] Time code recording.

[6] Transport stream (.ts) is a digital format that contains the video, audio and subtitle data.

[7] SubRip Subtitle (.srt) files are plain files that contain the text of the subtitles, as well as the start and end timecodes. Unlike .ts files, they do not include any video or audio data.

[8] Igor Leturia informed us that they were unable to identify the problem that hindered the recognition of sample ETB1_1.

[9] *Euskera batua* is the standardised version of the Basque language. It was proposed by Euskaltzaindia (the Academy of the Basque Language) in 1968 and serves as the formal model and written standard of the language. *Euskalkiak* are dialectical varieties of Basque that display differences in grammar, vocabulary and pronunciation. Nowadays, five *euskalki* are recognised: Western (Biscayan), Central (Gipuzkoan), Navarrese, Navarro-Lapurdian, and Souletin (Zuazo 2014: 76).

[10] The *Libro de estilo de EiTB: Informativos y Actualidad* style guide (Martín Sabarís *et al*. 2016) stipulates that, in television broadcasts in the Basque language, proper nouns are to be written without diacritics. For this reason, it was decided to group them together with ortho-typographic errors.

[11] In Catalan, for instance, the following percentages are reported: 70% R-M, 24% R-ST and 6% R-SE (errors with punctuation) and 55% R-M, 35% R-ST and 10% R-SE (errors without punctuation) (Tamayo and de Higes 2023).

[12] The measurements and means of CPS presented throughout this paper should be taken with care. Recent research poses methodological considerations that should be addressed when researching on subtitling speed and using CPS measurements and mean values (see Fresno and Sepielak 2020). Numbers presented in this contribution should be seen as illustrative of the real situation but can vary depending on the methodological model and the programme or criteria used to measure subtitle speed and to consider (or not) mean values per sample as a representative value of that sample.

[13] 　　　　Retrieved　　　　from　　　　https://partnerhelp.netflixstudios.com/hc/en-us/articles/5521178339731-Basque-Timed-Text-Style-Guide (consulted 06.09.2023)

[14] It is worth re-emphasising that EITB's news and current affairs style guide (Martín Sabarís *et al.* 2016) lays down that in Basque proper nouns should be written without diacritics.