One short film, different audio descriptions. Analysing the language of audio descriptions created by students and professionals

Anna Matamala
Universitat Autònoma de Barcelona
España

Anna Matamala: Departament de Traducció i d'Interpretació i d'Estudis de l'Àsia Oriental, Universitat Autòno-
ma de Barcelona, España. | Correo electrónico: anna.matamala@uab.cat

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

186

Abstract

Audio description is an access service that translates visuals into words which are then received auditorily by end users. This article aims to compare audio descriptions created by professionals and by students in Spanish, in order to identify commonalities and divergences, if any. Part of the VIW corpus is used to that end. Although limited in size, the VIW corpus is the first open-access multilingual multimodal corpus of audio descriptions, including various audio descriptions of the same visual input, namely the short film "What happens while". The analysis presented in this paper focuses on quantitative data such as the number of words, number of AD units and sentences, as well as on the distribution of word classes. Where relevant, data are compared to general language corpora so as to highlight the specificities of the language of audio description. The paper also compares the most frequent words used by both groups, and a selection of semantic classes automatically retrieved is analysed. This article provides new insights into the language of audio description and shows the research potential of the VIW corpus.

Keywords: audiovisual translation; media accessibility; audio description; corpus; training.

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

187

Introduction

Audio describing a film means translating images into words that will be read aloud and interspersed in specific film segments where there is usually no dialogue or relevant sounds (Maszerowska et al., 2014; Matamala and Orero, 2016). Translating all visuals is an impossible task, not only due to the complexity of the visual image but also due to the spatial constraints of any audio description (AD). This is why the selection of the visuals to be transferred inter-semiotically is of paramount importance. This is paralleled with the importance of selecting the right words in the process of creating the AD.

In order to better understand both the process of visual selection and the process of word selection, the literature on AD has used theoretical models from narratology (Kruger, 2009) and from Relevance Theory (Braun, 2007), to name just two. Cognitive models have been the basis for experimental research in recent years, but the focus has been mainly on how end users received different versions of the same audio description. Preferences and enjoyment related to different voicing strategies have been investigated (Szarkowska, 2011; Fernández-Torné and Matamala, 2015), along with the impact on comprehension when modifying information explici-tation, intonation or speed (Cabeza-Cáceres, 2013), or the impact on recall when changing the amount of information provided and its segmentation (Fresno, 2014). Experimental research on the creation process has been more limited, analysing for instance the inclusion of new te-chnological components such as machine translation (Fernández-Torné and Matamala, 2016).

Descriptive research has provided new insights into the processes of visual and word se-lection by analysing the final output. Resorting very often to case-studies, be it a single film or a limited number of films, scientific publications have discussed various formats and genres (theatre, opera, art, cinema), and specific components, such as characters (Fresno, 2012) or special effects (Matamala and Remael, 2015), to name just two. Corpus research on the featu-res of the specific language of audio description has been more limited: Salway (2007) led the TIWO (Television Into Words) project, in which 91 scripts in British English were collected and analysed to elucidate the specific features of the language of AD. Part of the TIWO corpus was also analysed by Arma (2011), who focused on adjectives. Another corpus worth mentioning is TRACCE, an exhaustive database including 300 films audio described in Spanish plus 50 films in German, English and French (Jiménez Hurtado and Seibel, 2012). More recently, Reviers has been working with corpus linguistic tools on a corpus of 17 scripts in Dutch to demonstrate that describers use a specialised language (Reviers et al., 2015). However, none of the pre-viously mentioned corpora are publicly available, because copyright restrictions apply.

All the previous research has made it possible to gain a better understanding of audio description, which undoubtedly impacts positively on the training of audio describers. There is a need to provide more access services and, therefore, a need for trained professionals, as

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

188

rightly pointed out by the ongoing European project ADLAB PRO (http://www.adlabproject.eu/). The training of audio describers and their main competences and skills has been discussed by Matamala and Orero (2007), who suggest the design of an AD course. Díaz-Cintas (2006) also puts forward an extensive review of the competences of audio describers, and Matamala (2006) and Badia and Matamala (2007) discuss training challenges in media accessibility. Specific handbooks have also been developed by Fryer (2016) and Snyder (2014).

Against this background, this article aims to contribute to current AD research in an innovative way: by comparing different audio descriptions of the same film. Based on the VIW corpus (Matamala, submitted), it aims to describe the commonalities and divergences found in audio descriptions created by professionals and students, if any, using multimodal and corpus analysis tools. The focus will be on the linguistic features of two subcorpora, both in Spanish: on the one hand, ten audio descriptions by professional AD providers and, on the other, ten audio description by MA students trained in AD. The ultimate aim is to indicate any differences which may lead to improvements in current AD training. Additionally, this research aims to provide data which can contribute to the ongoing debate (Reviers et al., 2015) on whether audio description can be viewed as a specialised language. Although a thorough analysis of this aspect is beyond the scope of this article, defining the features of AD as opposed to general language will contribute not only to gaining a better understanding of this mode of transfer, but also will help to identify the linguistic structures that future audio describers need to master. All in all, the results of this article will not only have a linguistic value, but could ultimately have an impact on AD training.

The article begins with a description of the VIW corpus and of the methodology used to extract the data. Sections 2, 3 and 4 discuss quantitative data linked to general aspects of the subcorpora, with an emphasis on numerical features, word classes and word frequencies. Section 5 uses computational tools to assess the text similarity between the different AD providers, and section 6 offers a semantic analysis based on an automatic tagging. The corpus is limited in size, so descriptive statistics are prioritised over measures of statistical significance. However, the fact that it is the only AD corpus available in open access and the only AD corpus that provides as many as ten versions for a single input justifies the relevance of this analysis, which can be expanded in the future with more data to confirm the preliminary trends found in this research.

1. The VIW corpus

The VIW (Visual Into Words) corpus is a multimodal and multilingual corpus of audio description. A 14-minute film in English, *What happens while…*, was created by film director Núria Nia specifically for the project. Dubbed versions in both Spanish and Catalan were also produced.

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

189

The film portrays how different characters (a student, a businessman, a retiree) envisage time. The original and dubbed versions of the short film are available on the project website (http://pagines.uab.cat/viw/content/film-whw), along with a making-of track by the film director. For further details on the process of film creation and dubbing, see Matamala and Villegas (2016).

Ten professional audio descriptions of the same input were commissioned to professional service providers, totalling 30 audio descriptions (10 in English, 10 in Spanish and 10 in Catalan). Additionally, audio descriptions from students were also collected (10 in Spanish, 7 in Catalan). The corpus currently contains more than 30,000 words, and all materials have been made available in open access to the scientific community through the open access repository of the UAB and the project website (pagines.uab.cat/viw).

Data were annotated using ELAN (Sloetjes and Wittenburg, 2008), since this multimodal analysis tool allows annotations to be linked to the video file. A thorough description of the process is presented in the technical documentation (see project website), but the main process could be summarised as follows: text files were collected and manually corrected as there were problems with certain time codes. Next, transcripts were sent to natural language processing tools, as will be explained later, and a manual error check was carried out. The linguistic annotations were loaded into ELAN, which was then used to query and export data. Graphic visualisations on a web app were produced.

Data were annotated at two levels (or tiers, in ELAN's terminology): at the linguistic level and at the filmic level. At the filmic level, the annotation took into account different items: scene (where the action takes place), shot type, sound elements, character on screen, and text on screen. The rationale behind each of these tiers and the tags used are explained in Matamala and Villegas (2016). As for linguistic tiers, they were split in two top tiers: the AD-unit tier and the Credits tier. Credits were excluded from the data analysis, because it was observed in a preliminary test that their AD in such a short film distorted the final results. They will be analysed separately at a later stage. As for the AD-unit tier, AD units were considered to be units separated by pauses longer than one second, which coincided with the type of segmentation sent by most AD providers. Each AD-unit was split into sentences, chunks, and tokens. ELAN's tokeniser was used for sentences and chunks, while natural language processing tools (NLP) were preferred for tokens as the tokenisation generated by ELAN and our NLP tools did not match. The token-level annotation included parts of speech, lemma and semantic values, which are the basis of the present analysis. Although it has not been developed as yet, an annotation level called AD-focus was established and will allow for the inclusion of future annotations related to the audio description itself.

To tag the audio descriptions with part of speech information, the Stanford parser (nlp.stanford.edu) was chosen for English, and the Freeling parser (nlp.lsi.upc.edu/freeling/node/1) was used for Spanish and Catalan through the web service available at the Competence Cen-

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

190

tre IULA-UPF-CC CLARIN (lod.iula.upf.edu/index-en.html). The Pympi library was used to import and merge the annotations into the ELAN's .eaf files. Semantic tags were taken from the Suggested Upper Merged Ontology (SUMO) (www.ontologyportal.org), and a selective semantic annotation was prioritised. This means that not all words were annotated and, when they were, not all were encoded with the same fine-grained criteria. More specifically, after a preliminary analysis of the corpus and of the bibliography (Salway, 2007; Arma, 2011; Reviers et al., 2015: 179), it was decided to prioritise:

a) verbs linked to spatio-temporal settings, movements, communication, and character descriptions;

b) adjectives linked to the mood of the characters, and adjectives dealing with hearing and sight;

c) nouns linked to spatio-temporal settings and characters;

d) adverbs linked to spatio-temporal settings.

Many units were assigned a top node in the ontology (Process for many verbs, Subjective Assessment Attribute for many adjectives, for instance) and the focus was on the units within the following semantic classes: Animal, Authoring, Body Motion, Body Part, Clothing, Colour Attribute, Communication, Dressing, Drinking, Film Language, Gesture, Hearing, Human, Intentional Motion, Intentional Psychological Process, Location, Motion, Object, Objective Assessment Attribute, Positional Attribute, Possesses, Putting, Quantifier, Radiating Light, Radiating Sound, Seeing, Smelling, State of Mind, Time, Touching, Vegetable and Walking. It must be stressed that the semantic tagging led us to manually identify multiword units and, consequently, re-encode the tokenization produced automatically.

ELAN is a very powerful tool which allows for multiple searches, especially when combining different levels of annotations aligned on a timeline. However, to carry out textual analysis focusing on linguistic aspects, as will be done in this article, certain corpus linguistic tools may provide further functionalities. This is why the corpus data were also prepared to be analysed by CQP, and CQP files are available in the GitHub project repository (github.com/TransmediaCatalonia/viw-scripts).

2. AD units, words and sentences

This section analyses quantitative data in both corpora related to the number of words, number of sentences and AD units. AD units are understood as descriptive units separated by at least a one second pause.

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

191

In terms of AD units, the professional subcorpus contains a higher number (M=49.7; SD=9.27) than the students' subcorpus (M=45.3; SD=12.37), with greater variability in the students' subcorpus, a trend that is confirmed when looking at each specific provider: the maximum number of AD units is 67 (Aptent), and the minimum is 37 (Trágora). In the students' subcorpus, the maximum is higher (74; García) and the minimum is lower (33; Rubio). Concerning sentences, the trend is the same, with a mean of 75.8 in the professional subcorpus and a standard deviation of 15.88, and a mean of 62.4 in the students' subcorpus and a standard deviation of 19.98. When analysing data individually, one finds again that the professional subcorpus ranges from 52 words (EdSol) to 102 (Aptent). Students, on the other hand, feature descriptions ranging from 42 words (Castro, Rubio) to 100 words (Marco). Again, a higher number of sentences is present in the professional subcorpus, which is also more consistent.

The number of words—which has been calculated excluding tokens tagged as punctuation by Freeling—follows the same trend, with a mean of 608.8 (min=456; max=875; SD=139.25) for professionals and 501.2 for students (min=317; max=726; SD=137.90). In both subcorpora there is considerable variation, but overall professionals use a higher number of words. Moreover, it is interesting to observe the mean number of words per AD units: professionals include 12.57 words on average (min=9.37; max=23.54; SD=3.96) whilst students include 11.42 (min=7.47; max=18.15; SD=3.52). The maximum number of words in an AD unit in the professional corpus ranges from 29 to 106 (M=43; SD=22.74), while in the students' subcorpus it ranges from 21 to 49 (M=34.30; SD=9.94). In this case, the professional corpus displays a greater variability, with one provider (Aristia), including less dense AD units and another one (Trágora) favouring longer ones. The students' uniformity may be due to the fact they have been trained in audio description in the same context, while the professionals may use different approaches. Concerning the mean number of units per sentence, the two subcorpora display similar means, around 8.4 words per sentence. A Mann-Whitney test was performed on the previous data, and no statistically significant differences were found between students and professionals under analysis. However, it is interesting to see that, concerning the length of the sentences, audio descriptions have specific features which contrast with general language: for instance, in the corpus Cumbre, Cantos and Sánchez (2011: 29) found that the mean number of words per sentence in Spanish was 20.89, a much higher number than the 8.4 mean found in our corpus. Table 1 summarises some of the main descriptive statistics obtained and discussed in this section.

3. Word classes in the corpus

Analysing the word classes in both subcorpora will allow us to identify whether there are any differences between professionals and students in the VIW corpus. It will also allow us to compare the percentage of frequent word classes in general language and in previous studies

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

192

TABLE 1
AD units, sentences and words countings

#		MEAN	SD	MEDIAN	MIN	MAX
AD units	**Prof**	49.7	9.27	46	37	67
	Students	45.3	12.37	39.5	33	74
Sentences	**Prof**	75.8	15.89	75.5	52	102
	Students	62.4	19.98	56.5	42	100
Words	**Prof**	608.8	139.25	563	456	871
	Students	501.2	137.90	502	317	726

in other language, such as Reviers et al. (2015). Data for general language included in this section have been obtained in June 2017 from the annotated version of the CORPES (Corpus del Español del Siglo XXI) corpus (http://web.frl.es/CORPES/view/inicioExterno.view), created by the Real Academia Española. Different searches for each word class have been made, limiting the results to Spain and written content. Word classes used in the CORPES corpus sometimes do not match tags used in FreeLing, so they have been grouped accordingly. Table 2 presents the relative frequency (i.e. percentages of the total word count) of word classes based on FreeLing's tags. Results tagged as "unknown", interjections (with 0 frequency in the VIW corpus) and punctuation have not been included. This means that the percentage is calculated on the word classes included below and not considering other words which may be included under other categories.

TABLE 2
Relative frequency of word classes

	PROFESSIONALS	STUDENTS	CORPES
Adjectives	4.50	3.75	6.71
Conjunctions	5.57	5.47	6.38
Determiners	17.72	17.87	16.44
Nouns	24.64	25.22	26.82
Pronouns	6.16	6.21	6.39
Adverbs	3.24	2.99	5.61
Verbs	22.09	20.77	17.20
Numbers	0.69	0.98	1.03
Adpositions	15.39	16.74	13.42

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

193

When comparing the distribution of word classes in the two VIW subcorpora, it can be seen that the highest difference is only 1.35 for adpositions, and that most word classes show a difference lower than 1 point. This seems to indicate that students and professionals use the same proportion of word classes. However, professionals seem to use a slightly higher number of adjectives and verbs, while students use a slightly higher number of nouns and adpositions.

When analysing the most frequent word classes, nouns come first (professionals =24.64; students=25.22), followed by verbs (professionals=22.09; students=20.77), a trend also found in Dutch by Reviers, her figures being 22 for nouns and 18 for verbs. Reviers also finds that both nouns and verbs depict a higher presence in an audio description corpus than in a general language corpus. This trend was only partially confirmed in our analysis (see table 2), because there are fewer nouns, but more verbs in the VIW corpus than in the CORPES corpus. However, a specificity that Reviers found is that the AD corpus in Dutch included more nouns than verbs, while in the Dutch general corpus the opposite trend was observed. In our case, both the AD corpus and the CORPES corpus follow the same trend: there are more nouns than verbs. This could be due to specificities related to the language and not to the AD.

Regarding adjectives, professionals tend to include more (4.50) than students (3.75), but surprisingly the percentage is lower than in the general language corpus (6.71).

Overall, there are more open-class words (professionals=54.47; students=52.73) than closed-class words (professionals=45.54; students=47.27), but in the CORPES corpus the difference is slightly larger (open-class=56.34; closed-class=43.66). These figures are very similar to those found by Reviers in her AD corpus (open-class=57; closed-class=43), but different from the data from the general Dutch language corpus (open-class=70; closed-class=30).

When analysing the percentage of open classes (nouns, verbs, adverbs and adjectives) based on the total of open classes (see table 3), similar figures can be observed, although professionals tend to include more adjectives and students tend to include more nouns, as observed previously. Adjectives, as pointed out by Arma (2011), are keyword classes in audio description, which aims to produce a vivid and precise language. Indeed, Igareda and Matamala (2012) found that students trained in AD include a higher number and a wider array of adjectives in their descriptions compared to students with no AD training. Similarly, professionals (who have more experience than students) tend to use more adjectives in our corpus, although the differences are not as remarkable.

The difference in the percentage of word classes between the two subcorpora is not high, but the limited size of the VIW corpus should be considered. However, it is interesting to observe the percentage of unique nouns, i.e., the total number of words minus all word repetitions, and the same with the other open classes. This could be viewed as an instance of language richness or variety. Table 4 provides a summary of the data obtained.

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

194

TABLE 3
Relative frequency of open classes

	PROFESSIONALS	STUDENTS
Adjectives	8.26	7.11
Nouns	45.24	47.82
Adverbs	5.94	5.68
Verbs	40.56	39.39

TABLE 4
Percentage of unique open classes

	PROFESSIONALS	STUDENTS
Adjectives	77.37	78.72
Nouns	61.60	60.76
Adverbs	65.33	66.52
Verbs	58.36	59.56

Again, table 4 shows that there is less than a 1.5 difference between groups, with higher percentages for students in the case of adjectives and verbs, and higher percentages for professionals for nouns and adverbs.

4. Most frequent units

This section will analyse the most frequent lemmas found in the corpus, both when considering all tokens and when only considering open-class words.

- Twenty most frequent lemmas in the professional subcorpus (all word classes): *el, uno, de, y, se, a, en, con, mirar, su, por, qué, James, Rick, caminar, móvil, lo, hacia, estar, playa*.
- Twenty most frequent lemmas in the students' subcorpus (all word classes): *el, uno, de, y, se, a, en, mirar, con, su, por, James, qué, Rick, caminar, Jess, no, hacia, estar, playa*.

The twenty most frequent lemmas in both subcorpora include many shared closed classes: articles (*el, uno*), conjunctions (*y*), prepositions (*a, de, en, con, por, hacia*), determiners (*su*) and pronouns (*qué, se*). The verbs *caminar* (to walk), *mirar* (to look) and *estar* (to be) are also present in both, as well as the proper nouns *James* and *Rick* and the location noun *playa* (beach).

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

195

The only difference is that the professional corpus includes *móvil* (mobile phone) within the twenty most frequent lemmas, while the students' subcorpus does not include it, but features instead the proper noun *Jess*. Another difference is that the professional subcorpus includes the pronoun *lo* whilst the students' subcorpus includes the negation adverb *no*. It is worth highlighting that most lemmas belong to closed classes (65% in both subcorpora), a trend also found in the TIWO and Reviers' AD corpus (2015). When considering only open-word classes, the results are the following:

- Most frequent lemmas in the professional subcorpus (open-class only): *mirar* (to look at), *James*, *Rick*, *caminar* (to walk), *móvil* (mobile phone), *estar* (to be), *playa* (beach), *alrededor* (surroundings), *dejar* (to leave), *hablar* (to talk), *no* (no), *vaso* (cup), *Jess*, *tener* (to have), *hombre* (man), *llevar* (to wear, to bring), *banco* (bench), *haber* (auxiliary verb), *mano* (hand), *sonido* (sound).
- Most frequent lemmas in the students' subcorpus (closed-class only): *mirar* (to look at), *James*, *Rick*, *caminar* (to walk), *Jess*, *no* (no), *estar* (to be), *playa* (beach), *lado* (side), *café* (coffee), *dejar* (to leave), *hombre* (man), *buscar* (to look for), *año* (year), *arena* (sand), *hablar* (to talk), *móvil* (mobile phone), *tener* (to have), *alrededor* (surroundings), *blanco* (white).

The data demonstrate that both subcorpora share adverbs (*no*), verbs (*mirar, caminar, hablar, estar, dejar, tener*), proper nouns (*James, Rick, Jess*), general nouns (*hombre*), nouns related to places (*playa*), objects (*móvil*) and nouns used in locative expressions (*alrededor*). This shows that 70% of the most frequent lemmas are shared by the two subcorpora. When analysing the list of non-coincidences, some parallelism can be drawn: *vaso* is included as an object description in the professional subcorpus, whereas in the students' subcorpus the word that is included in the most frequent list is *café*. The only units that differ are two verbs (*llevar, haber*) and three nouns (*banco, mano, sonido*) in the professional subcorpus, and one verb (*buscar*), three nouns (*lado, año, arena*) and one adjective (*blanco*) in the students' subcorpus.

When considering the ten most frequent lemmas for each open class (see table 5), the percentage of shared units is very high in all of them: 60% of nouns (*James, Rick, móvil, playa, Jess, hombre*), 70% of adjectives (*alto, próximo, pensativo, pelirrojo, largo, canoso*), 80% of adverbs (*no, sí, ahora, después, más, tarde, ya, claro*), and 70% of verbs (*mirar, caminar, estar, dejar, hablar, haber, tener*).

Overall, the percentage of shared tokens in the 10 most frequent units averages 70%, which proves a high degree of consistency. This should come as no surprise, as the visual input is the same, but it suggests the professionals and students do not differ much when approaching the same content. When contrasting this data with other corpus analysis in the AD field, it is worth observing that Reviers et al. (2015: 179) found that "nouns referring to characters and their body parts abound, as do nouns referring to objects, verbs referring to actions of looking and to movement". Similarly, Salway (2007: 156) highlights the presence of nouns and verbs related to characters and

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

196

TABLE 5

Ten most frequent nouns, adjectives, adverbs and verbs

NOUNS		ADJ.		ADV.		VERBS	
PROF.	STUD.	PROF.	STUD.	PROF.	STUD.	PROF.	STUD.
James	James	Blanco (white)	Negro (black)	No (no)	No (no)	Mirar (to look)	Mirar (to look)
Rick	Rick	Negro (black)	Próximo (next)	Después (after)	Ahora (now)	Caminar (to walk)	Caminar (to walk)
Móvil (mobile phone)	Jess	Mismo (same)	Pelirrojo (red-haired)	Claro (sure)	Claro (sure)	Estar (to be)	Estar (to be)
Playa (beach)	Playa (beach)	Pensativo (thoughtful)	Pensativo (thoughtful)	Más (more)	Arriba (above)	Dejar (to leave)	Dejar (to leave)
Alrededor (surroundings)	Lado (side)	Próximo (next)	Nervioso (nervous)	Ahora (now)	Atrás (behind)	Hablar (to talk)	Buscar (to look for)
Vaso (cup)	Café (coffee)	Marítimo (maritime)	Largo (long)	Ya (now, already)	Más (more)	Tener (to have)	Hablar (to talk)
Jess	Hombre (man)	Pelirrojo (red-haired)	Canoso (grey-haired)	También (also)	Sí (yes)	Llevar (to wear, to bring)	Tener (to have)
Hombre (man)	Año (year)	Alto (tall)	Alto (tall)	Sí (yes)	Después (after)	Haber (aux.)	Vestir (to wear)
Banco (bench)	Arena (sand)	Largo (long)	Gris (grey)	Tarde (later)	Tarde (later)	Extrañar (to surprise)	Haber (aux.)
Mano (hand)	Móvil (mobile phone)	Canoso (grey-haired)	Nuevo (new)	De repente (suddenly)	Ya (now, already)	Poner (to put)	Levantar (to raise)

their body parts, actions, objects, and scenes. Our data coincide with the observations made by other researchers. Additionally, Reviers et al. state that the "few adjectives that do occur frequently in both the Dutch and Salway's (2007) and Arma's (2011) English corpus are: black/zwart, white/witte, dark/donker and young/jong". *Blanco* (white) and *negro* (black) are also found in our corpus.

5. Text similarity

As part of the analysis, text similarity was automatically computed using Pedersen's Text-Similarity module (version 0.10, released June 2016). This module measures the similarity of two documents based on the number of overlapping (shared) word tokens scaled by the lengths of

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

197

the files. More specifically, it computes the F-Measure, the Dice Coefficient, the Cosine and the Lesk measure. Although there is extensive literature on how to measure text similarity and the advantages of different measures (Gomaa and Fahmy, 2013), Ted Pedersen's Text-Similarity[1] module was chosen due to its availability as a web service (http://lod.iula.upf.edu/resources/429) and its easy integration in our project. Table 8 provides a summary of the output in percentages when considering all possible combinations. A detailed account of all the data obtained can be found on VIW's web app: http://transmediacatalonia.uab.cat/web/similarity/WHW-ES-Pr.

TABLE 8
Text similarity

SIMILARITY	UNDER 0.30	0.30 TO 0.39	0.40 AND ABOVE
PROF-PROF	0.53	20.00	3.16
PROF-STUDENT	5.79	41.05	5.79
STUDENT-STUDENT	2.11	20.53	1.05
TOTAL	**8.42**	**81.58**	**10.00**

The similarity between all pairs is mostly between 0.30 and 0.39 (81.58%), with 10% or less at a lower or higher range. The least similar texts in the corpus are two professional providers (Kaleidoscope-CEIAF) and a professional and a student (Martín-Aptent), both with a 0.05 score. The most similar texts are those produced by a student (Martín) and a professional provider (SDI-Media), with 0.55. Text similarities below 0.30 are found mostly when comparing professionals and students (5.79% of the all possible combinations), the same percentage that is found in the higher range. It seems that professional texts tend to be more similar, with lower numbers in the lower range and higher numbers in the higher range. However, most combinations are in the middle range, showing inconsiderable differences between the two subcorpora.

6. Semantic classes

This section will compare semantic classes in both subcorpora. Although the relatively small size of the corpus would allow for a manual qualitative analysis, an analysis based on the automatic tagging has been preferred, as proof of concept for future investigations when the corpus is expanded. Annex 1 includes a whole list of tags and their relative frequency when considering the total number of open-class words in the corpus. The most frequent ones are

---

1   http://search.cpan.org/~tpederse/Text-Similarity-0.10/lib/Text/Similarity/Overlaps.pm#DESCRIPTION

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

198

Location (professional=12.09; students=11.80), followed by Human (7.72), Object (6.36) and Motion (5.16) in the professional subcorpus, and followed by Object (5.49), Aux (5.18) and Seeing (4.69) in the students' subcorpus. However, there are only five semantic classes in which the difference is higher than one point (BodyPart, FilmLanguage, Motion, Process, Time), and one in which the difference is higher than 3 points (Human).

A thorough analysis of each semantic class is beyond the scope of this article, hence a brief analysis of the tags related to three relevant topics in the AD literature has been favoured: semantic tags linked to the description of characters (section 6.1), to spatio-temporal settings (section 6.2) and to colours (section 6.3).

6.1. Describing characters

Characters are an "essential part of a film narrative", as stated by Mazur (2015) in the ADLAB guidelines. Various authors have analysed the audio description of characters (Fresno, 2012; Fresno et al., 2016), sometimes focusing on facial expressions (Vercauteren and Orero, 2013; Mazur, 2014), and guidelines generally give specific sections on how to approach their audio description (Rai et al., 2010). Physical appearance is not always viewed as the most relevant aspect when describing characters, as they sometimes can define themselves more easily through their actions and words. Still, as Mazur (2015) suggests, it is advisable to include this type of description to help the end user if there is sufficient time. This is what the describers in the VIW corpus do.

When analysing in our corpus the use of units with the tags Clothing, Human, and Dressing, a subset of the physical description of characters, one can observe the distribution of verbs (professional=28 verbs; students=22 verbs), nouns (professionals=244; students=202) and adjectives (professional=4; students=2) under these semantic tags. This means that 8.32% of the open-class units in the professional subcorpus and a similar 8.55% of the open-class units in the students' subcorpus are given these tags. Focusing on the most numerous class, i.e., nouns, one can observe 16.8% of unique nouns in the professional subcorpus, whilst the students' percentage is slightly higher (19.30%), showing a wider variety. Another commonality is that the two subcorpora share 90% of the 10 most frequent words albeit in a different order: *hombre* (man), *ropa* (clothes), *bolsillo* (pocket), *traje* (suit), *empresario* (businessman), *estudiante* (student), *corbata* (tie), *jubilado* (retiree), *gafas* (glasses).

6.2. Describing time and location

Remael et al. (2015) acknowledge spatio-temporal settings as one of the basic narrative building blocks of films which merit specific attention in the AD. As stated, for instance, by the

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

199

Spanish AD Standard, the "time-space rule must be applied, consisting of clarifying the 'when', 'where', 'who', 'what' and 'how' of each situation which is audio-described" (Rai et al., 2010: 16). Due to the relevance of these items, the analysis in our corpus focuses on the units automatically tagged as Located, Location, Time and TimeLocation. Table 9 presents the distribution of location and time units in both subcorpora, which shows a similar percentage for time units and a higher frequency of location units in the students' subcorpus, even if absolute numbers show a higher presence in the professional subcorpus.

TABLE 9
Distribution of time and location units

	PROFESSIONAL	STUDENTS	PROFESSIONAL	STUDENTS
	LOCATION	LOCATION	TIME	TIME
Adverbs	33	20	53	22
Nouns	280	292	41	40
Verbs	13	3	2	0
Total	326	315	96	62
Relative freq (open classes)	9.83	11.91	2.89	2.34

When focusing specifically on the class with a highest presence, i.e., nouns, it is observed that the percentage of unique location and time nouns in the two subcorpora is very close, showing a similar degree of language richness: location (professionals=18.57; students=18.49), time (professionals=21.95; students=20). When looking at the specific lemmas, 70% of the first 10 most frequent location nouns are shared by both subcorpora: *playa* (beach), *lado* (side), *arena* (sand), *alrededor* (surroundings), *parque* (park), *paseo* (promenade), *cielo* (sky). In fact, when considering all nouns, more than 64.81% are shared by the two subcorpora. Concerning nouns referring to time, two out of the three most frequent ones are shared by the two subcorpora, and 71.42% of the location nouns found in the students' corpus are included in the professional one. Regarding location adverbs, there is a considerable difference in terms of variety: while professionals use 12 different forms, students only use three according to the automatic data extracted.

6.3. Describing colours

It is commonplace that people venturing into the field of AD wonder why colour should be described to blind and visually impaired audiences, but not all blind people are born blind and, additionally, colour often transmits meaning: red, for instance, is often believed to trans-

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

200

mit rage and intensity. This is why some AD guidelines such as the British guidelines by OfCom or the ADI guidelines insist on describing colour where relevant (Rai et al., 2010). Others standards such as the Spanish guidelines do not mention this issue. Table 10 lists all the adjectives referring to colour included in both subcorpus, according to the automatic semantic tagging.

TABLE 10
Adjectives with semantic tag Colour

PROFESSIONAL		STUDENTS	
Blanco (white)	20	Blanco (white)	20
Negro (black)	18	Negro (black)	9
Pelirrojo (red-haired)	7	Pelirrojo (red-haired)	8
Canoso (grey-haired)	6	Canoso (grey-haired)	5
Rubio (blonde)	3	Gris (grey)	4
Verde (green)	3	Verde (green)	3
Azul (blue)	2	Moreno (dark-skinned)	1
Claro (clear)	2	Teñido (dyed)	1
Oscuro (dark)	2	Rosado (rosy)	1
Dorada (Golden)	1	Claro (clear)	1
Marrón (Brown)	1	Rubio (blonde)	1
Entrecano (grayish)	1	Cobrizo (copper-coloured)	1
Verdoso (greenish)	1	Caqui (khaki)	1
Canoso (grey-haired)	1	Marrón (Brown)	1
Plateado (silver)	1	Oscuro (dark)	1
Gris (grey)	1	Amarillo (yellow)	1
Rojo (red)	1	Castaño (Brown)	1
Total	**71**	**Total**	**60**
Rel freq (over open classes)	**2.14%**		**2.27%**

The professional subcorpus includes more units describing colour (71 occurrences) than the students' subcorpus (60 occurrences), but when considering the total number of open-class units it turns out that students show a slightly higher percentage (students=2.27; professionals=2.14). When considering only the adjectives, colour adjectives account for 25.91% of the adjectives in the professional subcorpus and for a higher percentage (31.91%) in the students' subcorpus. The four most frequent units related to colour are common to both subcorpora:

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

201

- *Blanco,* which in different contexts is used to describe the colour of a car, of land, of the screen, of the letters on screen, a jacket, clouds, but never the ethnic origin of the characters.

- *Negro*, used to qualify mostly the last character (Zoe), but also used to describe a film technique, text on screen and the colour of various clothes.

- *Pelirrojo*, to describe characters' hair.

- *Canoso*, to describe a character's hair.

It is interesting to observe that *pelirrojo* (red-haired) is used to describe James by three students and one professional, while one student and three professionals describe him as *rubio* (blonde). It is also interesting to observe that the last character, a black woman, is described with this adjective, while the other characters' skin color is not normally referred to. In this regard, Rai et al. (2010: 6) state that citing "the race only of non-white individuals establishes 'white' as a default and is unacceptable". Further qualitative analysis is beyond the scope of this paper, but these preliminary data show the possibilities of expanding this type of research.

7. Conclusions

This article has approached various linguistic aspects, in order to compare two audio description subcorpora, one made up of professional descriptions and another one made up of students' descriptions of the same input. It has been proven that professionals tend to use a higher number of AD units, sentences and words, with less variability than students. However, it seems that the mean number of words per AD units is more consistent in the students' subcorpus, and overall statistical tests do not show any significant differences between the two subcorpora in all the previous items. Regarding the percentage of word classes, students and professionals show similar percentages, with professionals including slightly more adjectives and verbs and students using a slightly higher number of nouns and adpositions. Language variety is also similar, with students showing somewhat more variety in their choice of adjectives and verbs, and professionals including more variety in nouns and adverbs.

When analysing the most frequent words, it is observed that 70% of the most frequent lemmas are shared by the two subcorpora, which should not come as a surprise as the visual input is the same. When differentiating between open classes, 60% of nouns are shared, 70% of adjectives and verbs are shared, and 80% of adverbs are found in both subcorpora, showing overall a high degree of consistency. Even when the word is not found in both subcorpora, some parallelisms can often be drawn, as demonstrated above.

An additional test to contrast the two subcorpora has been carried out, computing text similarity by implementing Pedersen's Text-Similarity Module. Again, data show no clear-cut

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals
202

differences between professionals and students, although it seems that professionals tend to create more similar texts.

Regarding the semantic classes under analysis, differences are generally below one percent point. There are only five semantic classes in which the difference is above one point (BodyPart, FilmLanguage, Motion, Process, Time), and one with a three-point difference (Human). Focusing on semantic classes related to the description of characters such as Clothing, Human, and Dressing, 90% of the 10 most frequent words are shared, and the percentage of open classes given these semantic tags is very similar in the two subcorpora. Regarding units related to time and location, the percentage is similar for time units, but not for location, where students' subcorpus presents a higher frequency. As far as units referring to colour are concerned, percentages are again similar, although a difference in the usage of adjectives is found, with students tending to use more than professionals. Overall, the previous analysis found small differences between the two subcorpora, but the general trends remain the same. This could perhaps be attributed to the limited size of the subcorpora under analysis, but it could also be considered evidence that the training that MA students have received is adequate and has allowed them to acquire the skills and competences expected from professional describers. Further research with a broader corpus and different describers' profiles are needed. It would also be useful to carry out research not only on the output, but also on the process of both professionals and students, where differences are more likely to appear.

Some interesting differences were found when comparing audio descriptions, be they by professionals or by students, to general language data. In this regard, it is observed that sentences include a lower number of words in audio description than in general language. Concerning word classes, it seems the biggest difference is that the VIW corpus includes more verbs than the general language corpus, probably because the priority has been to transform the visual actions into words in order to enhance the understanding of the short film. This trend was also found by Reviers et al. (2015) in Dutch. However, some specific differences between general language corpora and AD corpora found in Dutch by these authors could not be fully confirmed in our analysis. This could be due to the limited size of our sample or to language-specific issues. In any case, these aspects need to be further explored in a broader corpus. Highlighting the nature of the under-researched language of audio description can prove invaluable when training future professionals.

Acknowledgments

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

203

References

Arma, Saveria, 2011: "The language of filmic audio descriptions: a corpus-based analysis of adjectives", PhD dissertation. Università degli Studi di Napoli Federico II.

Badia, Toni, and Anna Matamala, 2007: "La docencia en accesibilidad en los medios", *Trans. Revista de traductología* 11, 61-72.

Braun, Sabine, 2007: "Audio description from a discourse perspective", *Linguistica Antverpiensia New Series* 6, 357-372.

Cabeza-Cáceres, Cristóbal, 2013: "Audiodescripció i recepció", PhD dissertation. Universitat Autònoma de Barcelona.

Cantos, Pascual, and Aquilino Sánchez, 2011: "El inglés y el español desde una perspectiva cuantitativa y distributiva: equivalencias y contrastes", *Estudios Ingleses de la Universidad Complutense* 19, 15-44.

Díaz-Cintas, Jorge, 2006: *Competencias profesionales del subtitulador y el audiodescriptor*, Madrid: CESYA.

Fernández-Torné, Anna, and Anna Matamala, 2015: "Text-to-speech vs. human voiced audio descriptions: a reception study in films dubbed into Catalan", *Jostrans. Journal of Specialised Translation* 24, 61-88.

Fernández-Torné, Anna, and Anna Matamala, 2016: "Machine translation and audio description? Comparing creation, translation and post-editing efforts", *Skase. Journal of translation and interpretation* 9 (1), 64-87.

Fresno, Nazaret, 2012: "Experimenting with characters. An empirical approach to the audio description of real characters" in Aline Remael, Pilar Orero and Mary Carroll (eds.): *Audiovisual Translation and Media Accessibility at the Crossroads. Media for All 3*, Amsterdam/New York: Rodopi, 147-161.

Fresno, Nazaret, 2014: *La (re)construcción de los personajes fílmicos en la audiodescripción. Efectos de la cantidad de información y de su segmentación en el recuerdo de los receptores*, PhD dissertation. Universitat Autònoma de Barcelona.

Fresno, Nazaret, Judit Castellà and Olga Soler-Vilageliu, 2016: "What should I say? Tentative criteria to prioritise information in the audio description of film characters" in Anna Matamala

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

204

and Pilar Orero (eds.): *Researching Audio Description. New Approaches*, London: Palgrave Macmillan, 143-168.

Fryer, Louise, 2016: *An introduction to audio description*, London: Routledge.

Gomaa, Wael H., and Aly A. Fahmy, 2013: "A survey of text similarity approaches", *International Journal of Computer Applications* (0975-8887) 68 (13), 13-18.

Igareda, Paula, and Anna Matamala, 2012: "Variations on the Pear Tree experiment: different variables, new results?", *Perspectives. Studies in Translatology* 20 (1), 103-123.

Jiménez, Catalina, and Claudia Seibel, 2012: "Multisemiotic and multimodal corpus analysis in audio description: TRACCE" in Aline Remael, Pilar Orero and Mary Carroll (eds.): *Audiovisual translation and media accessibility at the crossroads*, Amsterdam and New York: Rodopi, 409-425.

Kruger, Jan-Louis, 2009: "The translation of narrative fiction: impostulating the narrative origo", *Perspectives. Studies in Translatology* 17 (1), 15-32.

Maszerowska, Anna, Anna Matamala and Pilar Orero (eds.), 2014: *Audio Description. New perspectives illustrated*, Amsterdam: Benjamins.

Matamala, Anna, 2006: "La accesibilidad en los medios: aspectos lingüísticos y retos de formación" in Ricardo Pérez-Amat and Álvaro Pérez-Ugena (eds.): *Sociedad, integración y televisión en España*, Madrid: Laberinto, 293-306.

Matamala, Anna, submitted: "The VIW (Visuals Into Words) project: multimodal corpus linguistics for audio description analysis".

Matamala, Anna, and Pilar Orero, 2007: "Designing a course on audio description: main competences of the future professional", *Linguistica Antverpiensia New Series* 6, 329-344.

Matamala, Anna, and Pilar Orero, 2016: *Researching audio description. New approaches*, London: Palgrave.

Matamala, Anna, and Aline Remael, 2015: "Audio description reloaded: an analysis of visual scenes in *2012* and *Hero*", *Translation Studies* 8 (1), 63-81.

Matamala, Anna, and Marta Villegas, 2016: "Building an audio description multilingual multimodal corpus: the VIW project" in *Proceedings of Multimodal Corpora: Computer vision and language processing (MCC 2016)* [http://www.lrec-conf.org/proceedings/lrec2016/workshops/LREC2016Workshop-MCC-2016-proceedings.pdf, last accessed: date of consultation: 1 March 2017].

Mazur, Iwona, 2014: "Gestures and facial expressions in audio description" in Anna Maszerowska, Anna Matamala and Pilar Orero (eds.): *Audio Description. New Perspectives Illustrated*, Amsterdam: Benjamins, 179-198.

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

205

Mazur, Iwona, 2015: "Characters and actions" in Aline Remael, Nina Reviers and Gert Vercauteren (eds.): *Pictures painted in words: ADLAB Audio Description guidelines*, Trieste: EUT [http://www.adlabproject.eu/Docs/adlab%20book/index.html, date of consultation: 1 March 2017].

Remael, Aline, Nina Reviers and Gert Vercauteren (eds.), 2015: *Pictures painted in words: ADLAB Audio Description guidelines*, Trieste: EUT.

Reviers, Nina, Aline Remael and Walter Daelemans, 2015: "The language of Audio Description in Dutch: Results of a corpus study" in Anna Jankowska and Agnieszka Szarkowska (eds.): *New Points of View on Audiovisual Translation and Accessibility*, Bern: Peter Lang, 167-189.

Salway, Andrew, 2007: "A corpus-based analysis of audio description" in Jorge Díaz-Cintas, Pilar Orero and Aline Remael (eds.): *Media for all. Subtitling for the deaf, audio description, and sign language*, Amsterdam: Rodopi, 151-174.

Sloetjes, Han, and Peter Wittenburg, 2008: "Annotation by category- ELAN and ISO DCR" in *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)* [http://repository.dlsi.ua.es/251/1/pdf/208_paper.pdf, date of consultation: 1 February 2017].

Snyder, Joel, 2014: *The visual made verbal: a comprehensive training manual and guide to the history and application of audio description*, Arlington, VA: American Council of the Blind.

Szarkowska, Agnieszka, 2011: "Text-to-speech audio description: toward wider availability of audio description", *Jostrans. The Journal of Specialised Translation* 15, 142-162.

Vercauteren, Gert, and Pilar Orero, 2013: "Describing facial expressions: much more than meets the eye", *Quaderns: Revista de Traducció* 20, 187-199.

Appendix 1

Relative frequency of all semantic tags taking into account the global number of open classes per subcorpus

	PROFESSIONAL	STUDENTS
A-ColorAttribute	2.14	2.38
A-Intentional Motion	0	0.04
A-ObjectiveAssessmentAttribute	0.12	0
A-SubjectiveAssessmentAttribute	4.10	3.56
Animal	0.42	0.34
Authoring	0.03	0

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

206

	PROFESSIONAL	STUDENTS
Aux	4.89	5.18
BodyMotion	3.89	3.37
BodyPart	4.89	3.63
Clothing	4.16	3.86
A-Clothing	0.12	0.15
Communication	1.75	1.44
Dressing	0.84	0.83
A-Dressing	0.15	0.42
Drinking	0.54	0.79
exists	0.21	0.30
FilmLanguage	1.66	0
A-FilmLanguage	0.51	0.15
Gesture	0.75	1.40
Hearing	0.69	0.42
Human	7.72	3.86
IntentionalPsychologicalProcess	1.81	1.70
A-IntentionalPsychologicalProcess	0.33	0.08
Located	0.39	0.79
Location	12.09	11.80
Motion	5.16	3.90
A-Motion	0	0.04
Object	6.36	5.49
PositionalAttribute	0.27	0.26
A-PositionalAttribute	2.47	2.27
possesses	0.69	0.42
Process	4.16	2.61
A-Process	0	0.08
Putting	1.96	2.65
A-Putting	0	0.04
Quantifier	0.21	0.04
RadiatingLight	0.18	0.26

ONOMÁZEIN 41 (septiembre de 2018): 185 - 207
Anna Matamala
One short film, different audio descriptions. Analysing the language of
audio descriptions created by students and professionals

207

	PROFESSIONAL	STUDENTS
A-RadiatingLight	0.21	0.23
RadiatingSound	1.30	0.76
A-RadiatingSound	0.03	0
Seeing	3.74	4.69
Smelling	0.09	0.38
StateOfMind	0.9	0.49
A-StateOfMind	3.02	2.46
Time	4.22	2.38
A-Time	0	0.04
TimeLocation	0	0.04
Touching	0.97	1.21
Vegetable	0.81	1.48
Walking	2.93	3.52
A-Walking	0	0