

Dealing with Variation in Audio Description Scripts

 **Eva Schaeffer-Lacroix** 

Sorbonne Université, research group STIH
(Sens Texte Informatique Histoire), Paris, France

 **Kirsten Berland** 

Inalco (Institut National des Langues et Civilisations Orientales),
Paris, France

Abstract

Audio description scripts represent a text type structured into several parts helping the speaker produce their recording. The heterogeneous composition and formatting of these parts makes it difficult to describe the linguistic features of audio description (AD) scripts in one go. Hence, it seems useful to implement them into a corpus tool enabling the analysis of the specificities of each AD section. In this paper, the AD scripts of 69 episodes from a German television show serve as a sample to explore a method for dealing with variation when preparing AD scripts for corpus processing. In our article, we offer a short overview of existing research on AD script corpora and on variation, and we present our dataset and the tools we used to prepare and explore the data. We then outline the features of the analysed AD scripts and the treatments applied. In the last section of the article, we discuss our results. Our analysis leads us to conclude that modifying original data for the sake of corpus implementation (e.g., changing formatting features) is a weighty step which may have unforeseen consequences: formal variation in AD scripts conveys more meaning than expected.

Key words: audio description, TV show, German, variation, corpus.

Citation: Schaeffer-Lacroix, E. & Berland, K. (2022). Dealing with Variation in Audio Description Scripts. *Journal of Audiovisual Translation*, 4(1), 150–65. <https://doi.org/10.47476/jat.v5i1.2022.181>

Editor(s): E. di Giovanni

Received: February 21, 2021

Accepted: November 17, 2021

Published: July 6, 2022

Copyright: ©2022 Schaeffer-Lacroix & Berland. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/). This allows for unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

 eva.lacroix@sorbonne-universite.fr, <https://orcid.org/0000-0002-6260-9095>


 kirsten.berland@gmail.com, <https://orcid.org/0000-0001-6532-2199>

1. Introduction

Audio descriptions (AD) are additional audio files which make visual aspects of films available to blind or visually impaired people. AD scripts contain, in addition to the text to be read aloud, a range of text elements helping the speaker produce their recording, mainly composed of time indications and speed instructions, elements that must not be overridden (dialogue prompts and sound events), and indications of scenery changes. This “truly interdisciplinary text type” (Mazur, 2020a, p. 234), described in Mazur (2020b), offers interesting challenges to researchers specialising in linguistics, literature, or translation studies. Ideally, AD creation is right from the beginning part of the film production process (Mazur, 2020a, p. 242). Producers are assisted by a tool which automatically provides them with XML-friendly structures such as time stamps and gaps for audio description. However, if ADs are created for existing films, such structures must be added manually in the case of text-based ADs or, in the case of media-based ADs, with the help of an audio description tool leading to a media-based output. The distinction between text-based and media-based ADs is explained on the website of the company 3Play Media (The Ultimate Guide to Audio Description, n.d.), which is also the source of Figure 1.

Figure 1

Audio Description Formats

AUDIO DESCRIPTION FORMATS				
TEXT BASED		MEDIA BASED		
WebVTT	Word Doc	MP4	OGG	
Merged Doc	Merged Text	MP3	WAV	
Plain Text	Merged Doc			
Stamped Doc	(Stamped)			

Source: The Ultimate Guide to Audio Description (n.d.).

The text-based AD creation method (viewing the film and writing the AD text in a separate text file) is still used in Germany, even if digital tools like Frazier (Pajoncsek & David, 2019) “make it faster and easier to create ADs”, as described on the web page *Audio Description of Visual Information* (Henry, 2019). In contrast to Matamala (2019), who concentrates on the multimodal aspects of the creation of an audio description corpus, our paper focuses on text-based ADs, subsequently referred to as AD scripts, which are less subject to copyright issues than media-based ADs. This choice is also motivated by our focus on understanding the consequences of the use of technology on the formal features of AD scripts during the AD creation process.

This paper starts by outlining research on corpus-based AD scripts and on variation. We then present our corpus data before offering a basic description of the corpus tool TXM (Heiden, 2010) and the text editor ATOM (*ATOM – A Hackable Text Editor for the 21st Century*, 2020). We continue by

describing our observations in the domain of AD script features, and we share our decisions concerning the pre-processing changes to be applied to our dataset. We finally explain how we carried out these changes before summarizing our findings.

2. Corpus-Based AD Script Analysis

Several researchers report that the use of corpus tools can support the description of the language of audio descriptions (Fix, 2005; Perego, 2018; Reviere, 2018; Salway, 2007; Zago, 2019). According to Fix (2005, p. 8), the following aspects of German audio descriptions have been analysed: morphosyntax, lexicology, semantics, discourse features, and non-verbal elements. Reviere (2018) provides a general description of the language of Dutch AD scripts, and she uses the part-of-speech tagger Frog (van den Bosch et al., 2007) and XML tags to annotate her AD data. As far as we know, the language features of the different sections of German AD scripts – time indications, prompts, speaker parts, and stage directions – have not yet been thoroughly analysed, nor described with the help of corpus tools.

3. Dealing with Variation

Variational features make corpus creation a challenging task, as explained in the introduction to a volume of the *CORPUS* journal dedicated to variation in oral corpora (Dugua & Kanaan-Caillo, 2021). Szmrecsanyi (2013, p. 261) defines language variation as follows: “Unter sprachlicher Variation verstehen wir das Vorkommen von zwei oder mehr funktional äquivalenten sprachlichen Ausdrucksformen, sogenannten Varianten, in einer Sprachgemeinschaft, bei einem Sprecher oder sogar in einem Text.” [By linguistic variation we mean the occurrence of two or more functionally equivalent linguistic expressions, so-called variants, used by a speech community, by a speaker or even occurring within one given text.] Applied to audio description, this topic can concern the way an AD is performed: Bardini (2020, p. 285) states that variation is “a deviation of standard language or neutral aural delivery in the AD”. Reviere uses Halliday’s variational concept “context of situation” (Halliday, 1994, p. 271) to study the way “how specific texts can be related to the immediate and specific material and social situation in which they are being used” (Reviere, 2017, p. 21). Broadcasters and AD companies are places where one – perhaps rather implicitly – “knows” how AD scripts should be produced. Therefore, we consider that they represent such “social situations”. AD scripts not only vary by the language used in the different sections, but also on the formatting level, which we consider as a situational feature because it may vary not only from one broadcaster to another but also from one AD author to another.

In our paper, the AD scripts of 69 episodes from the German television show *Neues aus Büttenwarder* (Eberlein, 1997-2016) are used to develop a method to prepare AD script data in text format for corpus processing. Even if one seeks to leave the original data as intact as possible, a posture which Pincemin (2011) calls “faithfulness to the text”, the intention to implement them into a given corpus

tool – in our case TXM – involves the transformation of the Word files into plain text files, as well as a certain number of pre-processing treatments (checking, normalising, and sometimes even modifying the data), failing which the query results will not be reliable enough, especially in the case of “less homogenous language varieties” (Schneider, 2020). Schneider explains why and how he made such changes when creating a corpus containing German song lyrics, called *Songkorpus.de* (Schneider, n.d.). One of the challenges he had to respond to – the presence of non-standard language features – must also be dealt with when working with AD script data.

4. Buettenwarder Corpus

The first edition of the Buettenwarder corpus was created during a one month’s bachelor internship linked to the French and German research project TADS (Translation of Audio Description Scripts). Implemented into TXM, this corpus contains 227,070 tokens, including word forms and punctuation marks. Its creation served as a test of our corpus pre-processing procedures before applying them to a larger body of AD scripts. The content was provided by the German broadcast company Polyphon Film- und Fernsehgesellschaft, Norddeutscher Rundfunk (NDR). We were granted the right to use the AD scripts of 69 episodes (out of 73 in total) of the TV show *Neues aus Büttenwarder* (Eberlein, 1997-2016) for research and teaching purposes and to deposit the annotated dataset in an open corpus repository. In the Internet Movie Database (IMDb), this TV show is described as follows (*Neues aus Büttenwarder (TV Series 1997–) – IMDb, 2001*):

Translating as “News from Büttenwarder”, the series recounts the modern-day trials and tribulations of a small, out-of-the way north German village. It generally focuses on the hair-brained schemes of farmer Kurt Brakelmann and his best friend Adsche Tönnsen to improve their lot. The action usually revolves around the village pub “Unter den Linden” and its cast of regular drinkers. Events are often precipitated by the arrival of outsiders with modern ideas which are sometimes met with suspicion but more often with overenthusiastic interest, especially if it might mean some revenue for the impoverished residents.

Our metadata file contains the episode numbers and titles of episodes 1–73 as well as the names of the AD script authors, the publication date (if available) of each text file and their size (in words). Table 1 lists the metadata of episodes 70–72.

Table 1

Metadata of the Buettenwarder Corpus

Episode number	Episode title	Author	Publication date	Size
70	Groggy	Ovelgönne	2016	1892
71	Laborette	Ovelgönne, Tietz	2016	2177
72	Oh!	Beckmann	2016	1650

Source: Authors' own elaboration.

5. Tools

In this section, we present the tools we used to prepare and to explore our data. The first one is a text editor and the second one a corpus management and exploration tool.

5.1. Text Editor ATOM

ATOM (2020) is an open-source text editor which can be used to check and to normalise data and to export files in different formats. Different users can simultaneously work on a shared file or project, e.g. to find and replace elements. Spelling errors and missing elements are signalled, and symbols like brackets or quotation marks are always written along with their matching symbol. We used the following functions: 1) Find all in the project, 2) Replace all, 3) Directly overwrite typos, insert missing elements, delete superfluous elements, 4) Select XML grammar.

5.2. Corpus Processing Tool TXM

As described on the TXM wiki website (*TXM – TEIWiki*, 2019), TXM is a “free and open-source XML & TEI compatible textual corpus analysis framework and graphical client based on the CQP search engine and the R statistical software”. This tool suite can be used for any language and is available as desktop software for Microsoft Windows, Linux, Mac OS X. When downloading the data into TXM, part-of-speech (POS) tags are added automatically with the help of the associated TreeTagger (Schmid, 1994). TXM offers users the possibility to modify erroneous annotations or typos on the fly. In addition to well-known corpus-tool functions (frequency lists, concordances, word pattern progression graphics), TXM offers functions which help compare the specific features of subcorpora, like FCA (Factorial Correspondence Analysis) and specific word pattern analysis.

6. AD Script Features

AD scripts are composed of text representing oral and written discourse and of numbers and signs. The following sample from episode 21 illustrates the original formatting of the text in the Word file provided by the broadcast company NDR. It contains the numbers of the information gaps, time indications, speaker parts (written in bold in this sample), dialogue prompts surrounded by double plus signs or – if cited within a stage direction part – by double quotation marks, stage directions put between brackets, and the abbreviated speed direction (s), meaning “schnell” [quickly], also put between brackets.

```
44: 10:05:25:09 10:05:27:06 01:22  
(s) Klopstedt setzt sich zum Essen.  
45: 10:05:27:07 10:05:49:11 22:04  
++ Und Sie sind also Arzt? ...
```

leichte Bindehautentzündung. ++
46: 10:05:49:12 10:05:51:22 02:10
Brakelmann lacht Kuno aus.
 47: 10:05:51:23 10:06:04:08 12:10
 ++ Und jetzt machen Sie mal Ferien?
 ... von unserer kleinen, aufstrebenden Gemeinde? ++
48: 10:06:04:09 10:06:05:21 01:12
 ("Ja" übersprechen)
Kuno reibt sich sein Auge.

[English translation:
 44: 10:05:25:09 10:05:27:06 01:22
(s) Kloppstedt sits down for his meal.
 45: 10:05:27:07 10:05:49:11 22:04
 ++ You are a physician, aren't you? ...
 mild conjunctivitis. ++
46: 10:05:49:12 10:05:51:22 02:10
Brakelmann laughs at Kuno.
 47: 10:05:51:23 10:06:04:08 12:10
 ++ And now, you are on holiday? ...
 from our small up-and-coming municipality? ++
48: 10:06:04:09 10:06:05:21 01:12
 (Override "Yes")
Kuno rubs his eye.]

As illustrated in the sample above, our AD scripts are composed of heterogeneous sections and text elements formatted in various ways. What is more, the typographical conventions vary throughout the 69 Buettenwarder files. Some of the AD scripts show advanced text-based formatting options: dialogue prompts and stage directions are put in italics, and speed indications are written in bold. It also happens that capitals (some of them even written in bold) are used in stage directions, probably to attract the attention of the recording person. In Table 2, "id" corresponds to the episode number. For space reasons, the text only appears in its English translation.

Table 2

Varying Dialogue Prompts and Speaker Parts

id	Dialogue prompts	Speaker parts	AD script authors
16	<i>„LIKE ALWAYS“</i> ...	He is going.	Koop
21	++ You are a physician, aren't you? ++	Kuno rubs his eye.	Ovelgönne, Schruhl
18	++ Westerland. Austria. ++	Outside.	Tietz, Schruhl
65	<i>... I will do it.</i>	He is grinning.	Tietz
70	<i>„Uncle Krischan ...“</i>	Krischan has white hair.	Ovelgönne

Source: Authors' own elaboration.

Table 2 presents a selection of varying dialogue prompts and speaker parts. Time indications also vary throughout the files: they only indicate the start time of an information gap, or they also mention its end time, and some of them even include its length; this is the case in the sample above (information gap numbers 46–48). This range of variation is partly due to the fact that the different episodes are written by varying authors or author teams over a period of six years (2011–2016). Our in-depth manual analysis of the 69 files reveals that the thirteen AD script authors or author teams of our dataset do not all have the same formatting habits. Table 3 presents the variations identified in the time indications and the speaker parts. The number provided in the column labelled “frequency” indicates in how many episodes out of 69 a given formatting and/or presentation option was chosen.

Table 3

Variation in Time Indications and Speaker Parts

Time indication	Speaker part	Frequency
beginning	bold	6
beginning	roman	22
beginning	bold	5
beginning	roman	3
beginning, end	bold	1
beginning, length	bold	1
length, beginning, end	bold	1
scenery number, beginning, end	roman	1
scenery number, beginning, end, length	bold	12
scenery number, beginning, end, length	roman	3
scenery number, beginning, end, length	bold	2
scenery number, beginning, length, end	roman	7
scenery number, length, beginning, end	bold	5

Source: Authors' own elaboration.

According to Table 3, these are the most frequent combinations: 22 out of 69 episodes only indicate the start time of an AD and show roman characters in time indications as well as in speaker parts. 12 episodes contain full time indications (scenery number, beginning, end, length), provided in roman characters, and all speaker parts in these 12 episodes are written in bold. In addition, some of the AD authors use italics for dialogue prompts and stage directions. We noted differences between the signs used to surround dialogue prompts (double plus or double quotation marks). Moreover, some conventions change within one text file or authors change their conventions in the different episodes they write. Another explanation for these variations could be the lack of a binding set of AD script writing conventions in German-speaking countries. In Germany, the AD recommendations on which several broadcast companies have agreed (Norddeutscher Rundfunk, 2019) only mention content

requirements. Anke Nicolai, a German audio describer and trainer, provided a list specifying speech direction conventions (which we translated into English). Table 4 juxtaposes her codes to those identified in the Buettenwarder AD scripts.

Table 4

AD Script Conventions for Speech Directions

Speech directions	Nicolai	Buettenwarder
Speak quickly	s	s (s)
Speak very quickly	s+	ss (ss) (ss:) ss+
Change of scenery within the text	*	#
Do not cover the sound within the text.	(sound)	(sound) (<i>SOUND</i>) (SOUND)
Go up with the voice to link the different parts of the sentence.	...	
Keep your voice up.	:	
Information was checked.	(!)	[!]

Source: Authors' own elaboration.

Anke Nicolai uses asterisks instead of hash signs to indicate scenery changes. Three of her conventions (<s+>, <...>, and <:>) do not appear at all in the Buettenwarder dataset; two of them concern prosodic features.

7. Pre-Processing Treatment

During our first upload test of the Buettenwarder dataset into TXM, going along with automatic part-of-speech tagging, a certain number of annotation errors occurred, mainly with dots, double dashes, and other signs or symbols. These errors mischaracterise the distribution of parts of speech in our dataset. As illustrated in Figure 2, the TreeTagger analysed multiple dots as numbers, regular nouns, truncated nouns, or foreign words.

Figure 2

Erroneous Labels for Multiple Dots

22 results found in 1 file for

Result ID	Value	Label	Flag
18932	CARD	[number]-m
39552	N.Reg.Acc.Sg.Neut-n
65201	FM-x
105027	CARD	[number]-m
144295	FM-x
179326	TRUNC.Noun-x
183058	N.Reg.Nom.Sg.Masc-n
238899	N.Reg.Acc.Sg.Neut-n
283245	CARD	[number]-m
295229	CARD	[number]-m
301290	CARD	[number]-m
310827	N.Reg.Acc.Sg.Neut-n
341824	TRUNC.Noun-x
354622	CARD	[number]-m

Source: ATOM screenshot.

To enable consistent and valid TXM queries in our AD script dataset, we normalised single and double quotation marks, commas, multiple dots, dashes, and hyphens. We also deleted unnecessary spaces and line spacing. All typos identified in the data (typo-like commas included) were corrected, e.g. *Parklatz* > *Parkplatz* [parking space], *entführt* > *entführt* [kidnapped], and *vereisen* > *verreisen* [take a trip]. In addition, we normalised time indications, deleted all scenery numbers, and added missing scenery change symbols (see subsection 7.2). In the following subsections, we present in more detail some of the changes applied to the data.

7.1. Time Indications

In our dataset, time indications appear in six different configurations (without considering the scenery number mentioned in Table 3):

- beginning
- beginning, end
- beginning, length
- beginning, end, length
- beginning, length, end
- length, beginning, end

Based on the assumption that the beginning of a time indication corresponds to the end point of the previous one, we decided to only keep the start times. This corresponds to the choice made by the AD authors in the last 20 episodes of our dataset (in total in 22 episodes; see Table 3). The simplification of the time indications (except of the combination “length, beginning, end”) was executed with the Find and replace all function of the text editor ATOM by using the following regular expression:

```
: [0-9]+\t[0-9]+(:[0-9]+){3}\t[0-9]+:[0-9]+
```

For each line with a time indication in the order <beginning, end, length>, this expression looks for whole lines starting from the first milliseconds indication.

7.2. Scenery Changes

Four out of the sixty-nine Buettenwarder episodes were missing the scenery change symbol <#>. Since audio description expert Anke Nicolai finds this information essential, we introduced the hash sign into the files in which it was missing by manually checking the corresponding audio-described videos. They can be accessed for free on the website *Büttenwarder von Anfang an* (NDR, n.d.) until a set date, depending on the episode. In our dataset, the scenery change symbol is never used between two subsequent audio description texts nor within a sentence, even if a scenery change takes place at that time. Thus, we added the hash sign between two subsequent sentences. With the help of the AD time indication, we checked the different audio described parts and inserted the hash sign in the positions of the AD script coinciding with a scenery change.

7.3. Speaker Modes

Speaker modes are used to specify the recording speed of an audio description. The voice talent must read the AD aloud either at normal speed, quickly, or very quickly, so as not to override dialogues nor important sounds. The abbreviations used in our German dataset are respectively <s> (*schnell sprechen*) [speak quickly], <ss> (*sehr schnell sprechen*) [speak very quickly], and <n> (*wieder in normalem Tempo sprechen*) [speak again at normal speed]. The code <n> only appears in the header of episodes 72 and 73. We identified two more codes: the first one, <ss:>, is the only speaker mode used in episodes 40 and 41. The second one, <ss+>, alternates with the codes <s> and <ss> in episodes 69, 70, and 71. In some of the files, all speaker modes are surrounded by round brackets; we deleted them with help of ATOM’s Find and replace all function.

7.4. Citing Text in AD Scripts

Dialogue prompts or speaker parts (i.e. the texts to be recorded as ADs) may contain elements surrounded by double or single quotation marks to indicate written information displayed on the

screen, e.g. the street sign “Kundenparkplatz” [customer car park] in episode 1. They are also used for citing direct speech like in the following dialogue prompt from episode 24:

“Modischer Chic, ... liebe Gäste einladen und sagen können: ‘Dies ist mein Zuhause.’”
[“Modern elegance, ... invite smart guests and be able to say: ‘This is my home.’”]

We also found an example with single quotation marks surrounding on-screen text, more precisely a design printed on a garment mentioned in episode 30:

Adsche trägt eine ‘Erster FC Köln’-Trainingsjacke.
[Adsche is wearing a ‘First FC Köln’ training jacket.]

Although we did not change the double quotation marks used for citations in the Buettenwarder dataset into single ones, we think that single quotation marks would help to better distinguish citations from dialogue prompts or speaker parts.

7.5. Multifunctional Symbols

In the original Buettenwarder AD scripts, stage directions (including sound events) are normally surrounded by brackets, but in places they are surrounded by the same symbols as dialogue prompts, e.g. <++ ... ++> or <“...”>. The first example in Table 5 is a dialogue prompt, the second one represents a stage direction (both are from episode 26). As can be seen in the third line, we also identified triple coding procedures: this stage direction from episode 44 is written in italics and surrounded by double plus signs as well as by round brackets.

Table 5

The Multifunctional Symbol <++>

Section	id	German original	English translation
dialogue prompt	26	++ Tschüss! Gute Reise! ++	++ Bye! Have a nice trip! ++
stage direction	26	++ Türumfallen ++	++ Falling door ++
stage direction	44	++ (<i>Abspann-Musik frei</i>) ++	++(<i>Do not override closing music</i>)++]

Source: Authors' own elaboration.

As explained in section 7.4, quotation marks are used not only to mark dialogue prompts but also to read aloud texts displayed on the screen. We found it necessary to visually distinguish the text parts that must be recorded from the other AD script parts. Consequently, within speaker parts, we surrounded with round brackets all stage directions and dialogue prompts – if necessary: dialogue prompts appearing within speaker texts tend already to be surrounded by round brackets and double quotation marks, as in this example from episode 71:

Die Frau ist Mitte 20 ... („Ich bin total fertig.“) ... und hat rot-braunes Haar.
[This woman is in her mid-twenties ... (“I am totally worn out.”) ... and has red-brown hair.]

Replacing double quotation marks by round brackets to mark stage directions in AD scripts – a formatting choice corresponding to the conventions identified in the AD scripts of another television show (Frach, 2007) – must be done manually. Automatic replacement may lead to errors because the meaning must be considered to decide whether the text to be modified concerns a sound event or a dialogue prompt. In the following example from episode 69, the context confirms that the indication “Stolpern” [Stumbling] is a sound event that must not be overridden:

Peter stolpert. (STOLPERN) Heinz rennt weiter.
[Peter is stumbling. (STUMBLING) Heinz keeps running.]

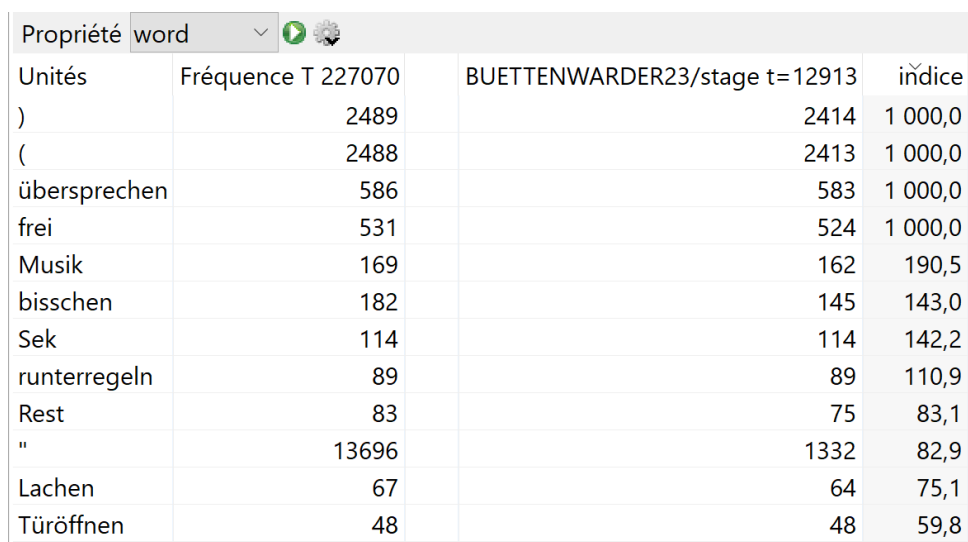
In any case, this sort of procedure is a strong intervention, which requires careful discussion before being applied to a larger dataset.

8. First Annotation Steps

After pre-processing our Buettenwarder data, we annotated four of the AD sections (time, speaker mode, stage directions, and dialogue prompts) with XML-TEI (P5) tags (TEI Consortium, 2020) and imported the dataset into TXM. This facilitates various ways to analyse the data, e.g. by exploring the specific vocabulary of the stage directions subcorpus (see Figure 3).

Figure 3

Specific Word List



Propriété	word	Fréquence T	BUETTENWARDER23/stage	t=	indice
)		2489		2414	1 000,0
(2488		2413	1 000,0
übersprechen		586		583	1 000,0
frei		531		524	1 000,0
Musik		169		162	190,5
bisschen		182		145	143,0
Sek		114		114	142,2
runterregeln		89		89	110,9
Rest		83		75	83,1
"		13696		1332	82,9
Lachen		67		64	75,1
Türöffnen		48		48	59,8

Source: TXM screenshot.

The first lines of the list displayed in Figure 3 contain quotation marks and words like “übersprechen” [override], “frei” [free], and “runterregeln” [turn down], which are part of the core vocabulary of German stage directions.

9. Discussion

Having described the tool-supported treatments applied to the German Buettenwarder dataset, we now present the limits of our interventions, and we share our recommendations for comparable datasets.

9.1. Limits of Automatic Replacement

The text editor ATOM supported most of our pre-processing steps in a satisfying way. However, automatic replacement has its limits: typos or sound event marking errors were better identified when directly scrolling through the text files. In addition, the success of using ATOM for automatic replacement depends on prior normalisation of text structuring signs like commas, dots, exclamation marks, and question marks. Automatic replacement is also conditioned by the quality of the original text: because of missing closing double quotation marks in the original files, our automatic replacement of all ++ symbols surrounding dialogue prompts by double quotation marks led to errors which then had to be manually corrected.

9.2. Limits of Disambiguation Steps

We found out that the formatting options of our AD script dataset do not always help distinguish stage directions from dialogue prompts. From a linguistic point of view, stage directions are quite heterogeneous: some of them are represented by a single noun like **(KNALL)** **[(BANG)]** or by truncated clauses: *(kurz Motor)* **[(Motor short)]**, which makes it difficult to interpret their status (are they stage directions or dialogue prompts?) when they appear within speaker parts. Other stage directions provide more explicit instructions thanks to the use of action verbs, e.g. in the sample **<("Ja" übersprechen)>** **[(Override "Yes")]**.

Finally, we question our deletion of the round brackets surrounding the letters **<s>**, **<ss>**, and **<n>** indicating speed instructions; round brackets would more clearly associate them with stage directions. In addition, during the automatic annotation process, marking such short speed instructions with round brackets would lower the risk of confusing them with hyphenated pronouns, as in episode 1: **<Mach's gut>** [Take care] or with determiners (see episode 18: **<um 10 ist's Licht aus.>** [at 10, the light is switched off]).

9.3. Recommendations

When normalising AD scripts, we recommend using the following safe symbol or sign replacements:

- (ss) > ss
- (s) > s
-)" >)
- "(> (
- ? " > ?"
- ! " > !"

In addition, we advise closing all Word or text files before opening a new dataset with ATOM and saving all changes before reopening the files with the help of another tool. As a matter of fact, some of our requests provided better results with other text editors: ATOM did not provide any result when we searched our files for double quotation marks at the beginning of all lines, but our search for `<\n>` elements was successful when using the text editor Notepad++ (Don Ho, 2010). When processing multiple files, the text editor Geany (Brush & Tröger, 2020) was more helpful than Notepad++ and ATOM in controlling which files were currently open and which changes were being applied to the data through automatic replacement.

10. Conclusion

Preparing text-based AD scripts for import in the TXM corpus tool helped us understand how this complex text type is made. We identified formatting differences and varying choices of content, e.g. regarding the degree of exhaustiveness of time indications, as well as several ways to add stage directions to speaker texts. These variations can be partly explained by the individual writing or editing habits of authors, by their lack of consistency, and by the evolution over time of AD script writing norms applied by an AD company or broadcaster.

Applying the corpus perspective to AD scripts or, more precisely, cleaning the data is one way to reduce formal variation. This sort of intervention has advantages and disadvantages: one must carefully evaluate which of the pre-processing steps are necessary and helpful, weighing the risk that some of such steps lead to simplified, poorly formatted texts which do not contain enough recording instructions. What would happen if a voice talent tried to record the speaker parts using our modified AD scripts? Another way to reduce variance would be to define a binding set of AD script writing conventions for German-speaking countries. However, the strict application of such rules may threaten the creative freedom of writers striving to produce audio descriptions that best fit a given film and a given cultural context.

Acknowledgements

Our warmest thanks go to Ursula Heerdegen-Wessel from NDR (Norddeutscher Rundfunk) for having provided us with the datasets cited in this article. We are also indebted to Anke Nicolai, who kindly answered our questions on AD script conventions.

References

- ATOM – A hackable text editor for the 21st Century (v1.46.0). (2020). [Computer software]. <https://atom.io/>
- Bardini, F. (2020). Audio description and the translation of film language into words. *Ilha Do Desterro A Journal of English Language, Literatures in English and Cultural Studies*, 73(1), 273–296. <https://doi.org/10.5007/2175-8026.2020v73n1p273>
- Brush, M., & Tröger, E. (2020). *Geany* (1.36) [En]. <https://www.geany.org/>
<https://www.br.de/br-fernsehen/sendungen/dahoam-is-dahoam/index.html>
- Don Ho. (2010). *Notepad++* (7.8.9) [Computer software]. <https://notepad-plus-plus.org/>
- Dugua, C., & Kanaan-Caillol, L. (2021). Introduction. *Corpus, Du recueil à l'outillage des corpus oraux : comment accéder à la variation?* (22) [From collection to technical equipment of oral corpora: how to access variation?], Article 22. <https://doi.org/10.4000/corpus.5885>
- Eberlein, N. (Screenplay writer) and Anderson, S. (Executive Producer). (1997-2021). *Neues aus Büttenwarder* [News from Büttenwarder] [TV show]. [Polyphon Film- und Fernsehgesellschaft, Norddeutscher Rundfunk].
https://www.ndr.de/fernsehen/sendungen/neues_aus_buettenwarder/folgen/index.html
- Fix, U. (2005). Hörfilm: Bildkompensation durch Sprache: linguistisch-filmisch-semiotische Untersuchungen zur Leistung der Audiodeskription in Hörfilmen am Beispiel des Films 'Laura, mein Engel' aus der 'Tatort'-Reihe [Hearing Film: Image Compensation through Speech: Linguistic-Filmic-Semiotic Studies on the Performance of Audio Description in Hearing Films Using the Example of the Film 'Laura, my angel' from the 'Tatort' Series]. Erich Schmidt.
- Frach, K. (Executive Producer). *Dahoam is Dahoam* [Home sweet home]. (2007–present). [TV show]. [Constantin Television].
- Halliday, M. A. K. (1994). *An introduction to functional grammar*. Edward Arnold.
- Heiden, S. (2010). The TXM platform: Building open-source textual analysis software compatible with the TEI Encoding Scheme. *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*, 389–398. <https://www.aclweb.org/anthology/Y10-1044>
- Henry, S. L. (2019). *Audio description of visual information*. W3C Web Accessibility Initiative (WAI). <https://www.w3.org/WAI/media/av/description/>
- Matamala, A. (2019). The VIW project: Multimodal corpus linguistics for audio description analysis. *Revista Española de Lingüística Aplicada* [Spanish Journal of Applied Linguistics], 32(2), 515–542. <https://doi.org/10.1075/resla.17001.mat>
- Mazur, I. (2020a). Audio description: Concepts, theories and research approaches. In Ł. Bogucki, & M. Deckert (Eds.), *The Palgrave handbook of audiovisual translation and media accessibility* (pp. 227–247). Springer. https://doi.org/10.1007/978-3-030-42105-2_12
- Mazur, I. (2020b). A functional approach to audio description. *Journal of Audiovisual Translation*, 3(2). <https://doi.org/10.47476/jat.v3i2.2020.139>

- NDR. (nd). *Büttenwarder von Anfang an* [Büttenwarder right from the Start].
/fernsehen/sendungen/neues_aus_buettenwarder/folgen/index.html
Neues aus Büttenwarder (TV Series 1997–)—IMDb. (2001).
https://www.imdb.com/title/tt0276716/?ref=nm_sr_srsrg_0
- Norddeutscher Rundfunk. (2019). *Vorgaben für Audiodeskriptionen* [Audio Description Rules]. NDR.
https://www.ndr.de/fernsehen/barrierefreie_angebote/audiodeskription/Vorgaben-fuer-Audiodeskriptionen,audiodeskription140.html
- Pajoncsek, L., & David, C. (2019). *Frazier*. VIDEO TO VOICE GmbH. <https://accessibility.studio/>
- Perego, E. (2018). Into the language of museum audio descriptions: A corpus-based study.
Perspectives, 27(3), 333–349. <https://doi.org/10.1080/0907676X.2018.1544648>
- Pincemin, B. (2011). Sémantique interprétative et textométrie – Version abrégée [Interpretative semantics and textometry – Short version]. *Corpus*, 10, 259–269.
- Reviere, N. (2017). *Audio description in Dutch: A corpus-based study into the linguistic features of a new, multimodal text type* [Universiteit Antwerpen].
<https://repository.uantwerpen.be/desktop/irua>
- Reviere, N. (2018). Studying the language of Dutch audio description: An example of a corpus-based analysis. *Translation and Translanguaging in Multilingual Contexts*, 4(1), 178–202.
<https://doi.org/10.1075/ttmc.00009.rev>
- Salway, A. (2007). A corpus-based analysis of audio description. In J. Cintas-Díaz, P. Orero, & A. Remael (Eds.), *Media for all: Subtitling for the deaf, audio description, and sign language* (pp. 151–174). Rodopi.
- Schmid, H. (1994). Probabilistic part-of-speech tagging using decision trees. *Proceedings of International Conference on New Methods in Language Processing*. International Conference on New Methods in Language Processing, Manchester, UK. <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf>
- Schneider, R. (n.d.). *Songkorpus—corpus of song lyrics*. <http://songkorpus.de/>
- Schneider, R. (2020). A corpus linguistic perspective on contemporary German pop lyrics with the multi-layer annotated “Songkorpus”. *Proceedings of the 12th Language Resources and Evaluation Conference*, 842–848. <https://www.aclweb.org/anthology/2020.lrec-1.105>
- Szmrecsanyi, B. (2013). *Variation und Wandel*. In P. Auer & H. Behrens (Eds.), *Sprachwissenschaft: Grammatik – Interaktion – Kognition* [Linguistics: Grammar – Interaction - Cognition]. Metzler. 261-284.
- TEI Consortium. (2020, August 19). *TEI P5: Guidelines for electronic text encoding and interchange*. Zenodo. <https://zenodo.org/record/3413524>
- The Ultimate Guide to Audio Description (n.d.). 3Play Media.
<https://www.3playmedia.com/learn/popular-topics/audio-description/>
- TXM – TEIWiki. (2019). <https://wiki.tei-c.org/index.php/TXM>
- van den Bosch, A., Busser, B., Canisius, S., & Daelemans, W. (2007). An efficient memory-based morphosyntactic tagger and parser for Dutch. In P. Dirix, I. Schuurman, V. Vandeghinste, & F. Van Eynde (Eds.), *Proceedings of the 17th Meeting of Computational linguistics in the Netherlands* (pp. 191–206). *Computational Linguistics in the Netherlands Journal*.
https://clinjournal.org/CLIN_proceedings/XVII/vandenbosch.pdf
- Zago, R. (2019). English films vs Italian films: A comparative analysis via the Pavia Corpus of film dialogue and the WordSmith Tools. In I. Ranzato & S. Zanotti (Eds.), *Benjamins translation library* (Vol. 148, pp. 230–241). John Benjamins Publishing.
<https://doi.org/10.1075/btl.148.11zag>