Studying the language of Dutch audio description

An example of a corpus-based analysis

Nina Reviers
TricS Research Group, Department of Translators and Interpreters, University of Antwerp

The present paper aims to combine insights from Applied Linguistics, Corpus Linguistics, Multimodality Research and Audiovisual Translation Studies in order to explore language use in a specific form of audiovisual translation, namely Audio Description (AD) for the blind and visually impaired. It is said that the communicative function of ADs and their multimodal context have a significant influence on the lexical, grammatical and syntactical choices describers make. This article aims to uncover these idiosyncratic linguistic patterns by conducting a quantitative and qualitative analysis of an annotated, audiovisual corpus of 39 Dutch films and series that have been released with AD in Flanders and the Netherlands. The paper analyses frequency lists, keywords, part-of-speech distributions and type-token ratios statistically and subsequently conducts a qualitative analysis taking systemic functional linguistics as a theoretical framework. The results confirm the hypothesis that the language of AD is idiosyncratic and highlight the most salient lexico-grammatical features characterising the language of Dutch AD.

Keywords: Audio Description, Corpus Studies, Audiovisual Translation, Multimodality

1. Introduction

The field of Audiovisual Translation (AVT) has been an increasingly popular topic in the past three decades and it has been studied from a range of interdisciplinary approaches, gradually moving from the periphery of Translation Studies (TS) to its very centre. One of the youngest disciplines in AVT that has been attracting more and more attention is Audio Description (AD), a form of media accessibility

https://doi.org/10.1075/ttmc.00009.rev
Translation and Translanguaging in Multilingual Contexts 4:1 (2018), pp. 178–202.
issn 2352-1805 | e-issn 2352-1813

that renders audiovisual texts, such as films, television series and theatre performances, accessible to blind and visually impaired audiences. AD is a form of intersemiotic translation that transfers images into words that are delivered aurally in between the sound effects and dialogues of the original audiovisual product. The aim is that audiences – not only the blind and visually impaired, but also those lacking access to the images for various reasons – can understand and enjoy the audiovisual text through the audio channel only. As a professional access service, AD took off in the 1990s and while it is well developed in some countries, it is rare or even non-existent in others (Reviers 2016). Research on AD is also recent. The main research focus has been on descriptive studies of specific genres, such as film, theatre, opera or dance, and on the detailed analysis of specific features of one text or a small collection of texts. Prevalent research topics include the question of content selection in AD, i.e., which visual elements should be prioritised (Vercauteren 2012), the degree of interpretation that is acceptable in AD with regard to, for instance, facial expressions (Orero 2012), and the reception of AD by end users (Fryer and Freeman 2012, 2013; Chmiel and Mazur 2012), to name but a few (see also Remael et al. 2016).

Contrary to other forms of AVT such as subtitling and dubbing, AD research has concentrated on the analysis and transfer of the visual aspects of the source text, rather than on the verbal aspects of wording and formulation of the target text. Linguistic aspects of AD, which are the main focus of this study, have been researched only by a handful of scholars, even though the literature frequently highlights the importance of language usage that is concise and comprehensible, yet simultaneously precise and vivid. Bourne and Jiménez (2007) were among the first to study linguistic aspects of AD by contrasting the Spanish and English descriptions of the same film. Kluckhohn (2005) studied cohesion and information order by analysing the AD of the German film *Laura, Mein Engel* (Runze 1994). AD guidelines for professional describers also advise on how to use or not to use language in AD with regard to aspects such as sentence length and complexity, word choice, tense, cohesion, albeit in very general terms (Ofcom 2000; Remael et al. 2014; Matamala et al. 2010). Salway (2007) tackled linguistic issues in English AD in a larger collection of texts with the TIWO corpus project. These studies, however, do not paint a complete picture and cover only a limited set of languages. What is more, today, scholars and practitioners emphasise that the translation of emotional and aesthetic aspects of the visual image – for which wording and style are particularly relevant – are as important as making the narrative accessible (Fryer and Freeman 2013). Finally, the scholarly study of different forms of AVT, including AD, has been criticised over the years, regarding the lack of methodological and theoretical rigour (Gambier 2006, 2009; Pérez-González 2014 and Remael et al. 2016). After all, it has relied largely on small case-studies

and it has not been able to adequately integrate the multimodal aspects of this type of translation – i.e., the fact that different semiotic signs such as dialogue, music and sound effects interact – into its research designs.

In addition, research into the language of Dutch AD, the more specific focus of this study, is virtually non-existent. Whilst at the start of the project AD services in the Low Countries were still in their infancy, in 2016 AD received a new boost under the influence of new regulations, incentives and technological advances.[1] At the time of writing, about 15 Dutch titles were available on DVD with AD, almost 60 titles could be accessed with a newly developed app called Earcatch (earcatch.nl) and at least three prime-time television series were aired (with open and closed AD) by the Flemish public broadcaster VRT per year.[2] These evolutions underline the importance of research projects aimed at supporting AD in the Low Countries.

In particular, Corpus Linguistics research in the field of AD is sparse, with just three relevant projects completed so far. The TIWO project (Television in Words) carried out a detailed analysis of the language of AD in a corpus of 91 British English descriptions (Salway et al. 2004, Salway 2007). The TRACCE project (Jiménez and Seibel 2012) developed what is currently the largest AD corpus, with 300 audio described films in Spanish and 50 in German, English and French. Finally, the recent VIW project (Visuals into Words) conducted at the Autonomous University of Barcelona is the first AD corpus freely available online. It contains 10 English, Spanish and Catalan descriptions of the same short film.[3]

Trying to move beyond the current state of affairs, the research project described here aims to combine insights from Applied Linguistics, Corpus Linguistics, Multimodality Research and AVT in order to obtain a detailed description of the lexico-grammatical features that characterise language use in AD in Flanders and the Netherlands. The hypothesis is that language use in AD is idio-

1. In Flanders, the Flemish Film Fund (*Vlaams Audiovisueel Fonds*; vaf.be) has made AD mandatory for all the films they finance. In the Netherlands, the development of an app called Earcatch (earcatch.nl) through which the descriptions of Dutch films and series can be accessed, has spurred the development of AD services for both film and television. The Dutch Film Fund (Film Fonds; filmfonds.nl) is actively promoting AD and provides the possibility of partial funding.

2. (a) These data were collected in October 2016. (b) The Flemish Public Broadcaster VRT offers AD as an additional audio channel that can be activated via the language menu for digital tv users (closed AD). In addition, the series is aired in parallel on another channel (OP12), where the AD is automatically activated, so that viewers who do not have digital TV, can also access the AD.

3. For more information, visit the project website at http://pagines.uab.cat/viw/ (accessed March 10, 2017).

syncratic and determined by the communicative function the text fulfils for its target audience (Salway 2007; Reviers et al. 2015). In addition, the multimodal context of the text type – the fact that the descriptions of salient visual elements by a narrator have to fit in between the dialogues, music and sound effects of the original audiovisual product – is thought to have a significant influence on the describers' lexical, grammatical and syntactical choices.

The current article presents the first phase of a four-year PhD project, conducted at the University of Antwerp between 2013 and 2017 and provides a first outline of the most salient features of this audiovisual text type. The goal of this first phase was to conduct a quantitative and qualitative lexico-grammatical analysis of a collection of Dutch descriptions. This analysis builds on the results of a pilot study conducted between 2010 and 2011 (Reviers et al. 2015). In Section 2, the article gives a brief description of the corpus and its design. This is followed by a general overview of the corpus data, with a brief outline of the methodology for data extraction and statistical processing. Finally, in Section 3, the data are analysed within the framework of Systemic Functional Linguistics (SFL).

2. The Dutch AD corpus

2.1 Corpus design

The corpus under analysis contains 39 descriptions – films and episodes of television series – totalling 154,570 words, and it was collected between 2011 and 2013. At that time, the corpus included the bulk of the material available in Dutch, except for a handful of products which were no longer available or never recorded, or for which we did not obtain copyright permission. But AD has grown exponentially since then, both in Flanders and the Netherlands. As a result, corpus collection is an ongoing process and it is our goal to keep it up to date at all times. That said, the corpus of 39 scripts that is the basis for the current analysis (henceforth referred to as the Dutch AD corpus) is a representative sample that contains three genres: action, drama and humour (see Table 1). While this may seem a rather limited number of genres, it is an adequate reflection of the Dutch AD market. To begin with, there are few Dutch films or TV series in other genres, such as thriller or sci-fi, and what is more, these are not the type of products that are made accessible to the blind and visually impaired. AD producers in the Low Countries seem to focus on popular mainstream film and TV products for the time being.

The corpus contains text, video and audio material of the 39 descriptions selected for the analysis. First, the voiced descriptions were transcribed from the DVD and provided with timecodes. The mark-up language XML was selected as

Table 1. Genre distribution in the Dutch AD corpus

Type	Genre			
	Action	**Drama**	**Humour**	**Total**
film	13%	61%	1%	76%
episode	22%	1%	1%	24%
Total	35%	62%	3%	100%
				154,570 words

the representation format for these transcriptions and timecodes, so that parts-of-speech (PoS) and lemma annotations could be added automatically with the Frog software.[4] The timecodes, in turn, allowed us to link each descriptive unit to its corresponding position in the video file through a specially developed multimodal concordancer, so that the voiced recording and the surrounding multimodal elements (such as sound effects, music and dialogue) could be consulted together with the transcriptions. This design facilitated the computer-based analysis of word, lemma and PoS frequencies, as well as the calculation of other relevant statistics like standardised type token ratio (TTR), average length of words, sentences and descriptive units, and average AD reading speed.[5] These data were collected in Excel for further statistical processing. The aim of the statistical analysis is to describe the features of a special language or register in terms of statistically significant differences between the corpus studied and a general language sample, as discussed in the next section.

2.2 Statistical processing

To begin with, the frequency counts extracted from the corpus texts were normalised, i.e., converted into a value per thousand words, because the length of the individual texts varied considerably. Means, medians and standard deviations were subsequently calculated. The standard deviation (*sd*) is a measure reflecting the variety of results across individual texts in the corpus. It is an estimate of how much the frequency scores in each individual text deviate from the average frequency score; in other words, how dispersed the data are. The standard deviation in our collection of texts was on the low side overall (maximum *sd* of 15, average

4. (a) The project followed the TEI P5 guidelines for XML developed by the Text Encoding Initiative (http://www.tei-c.org/Guidelines/P5/ (accessed May 23, 2016)). (b) The Frog system is a memory-based morphosyntactic tagger and dependency parser for Dutch developed by the CLiPS research group of the University of Antwerp, Belgium (see Van den Bosch et al. 2007).

5. These data were extracted with Python, a standard programming language in natural language processing (see python.org and nltk.org).

sd of 6.2). This signifies that the results were generally clustered around the average and that the corpus has a high degree of consistency.

Next, the lemma and PoS frequencies were contrasted with frequency data from two Dutch reference corpora: the SoNaR corpus, a written text corpus of about 500 million words from different fields and genres, and Subtlex-nl (Keuleers and Brysbaert 2010), a 44 million word corpus of Dutch subtitles.[6] This comparison investigates to what extent the frequency data in terms of PoS from the Dutch AD corpus deviate from the values one would expect to find in general language (represented by the SoNaR corpus) and in language use in audiovisual texts more in particular (represented by the Subtlex-nl corpus, which contains a language variety determined by the multimodal character of the source and target text as well). A chi-square goodness-of-fit test was used to calculate the probability or p-value for the PoS frequencies. Usually a p-value of 0.05 is taken as the critical value, below which the observed differences in frequency are significant. The obtained p-values were less than 0.001 for both corpora, which means that the observed differences between our corpus and the reference corpora are significant and not due to chance. Usually it is assumed that the lower the p-value, the more significant the differences between the two samples. However, taking the p-value as a measure for the degree of difference has been criticised lately, and some argue that the calculation of an effect size measure is more accurate (Nuzzo 2014). Therefore, we calculated an effect size (w) for the PoS frequencies as well (for SoNaR effect size $w = 0.27$; for Subtlex-nl effect size $w = 1.02$). The effect size was higher in the Subtlex-nl corpus, meaning that language use in our corpus in terms of PoS differs the most from language use in the subtitling corpus.

Finally, a p-value and an effect size were calculated for each PoS individually. This time we used a different critical value for p (alpha = 0.002 for SoNaR and alpha = 0.006 for Subtlex-nl) after taking into account a Bonferroni correction for multiple comparisons (Field 2009). This revealed that all measured PoS categories (adjectives, nouns, adverbs, verbs, articles, conjunctions, pronouns and prepositions) deviate from the reference values ($p < 0.0001$) when comparing with both reference corpora for all categories (see Table 4). However, the effect sizes for each PoS category were rather small (0.1 on average for SoNaR and 0.3 for Subtlex-nl). In other words, the results are statistically significant, but the size of the differences is modest. The results discussed above also corroborate what previous research has suggested. In the pilot study preceding the current project, a similar analysis was conducted, albeit with a smaller corpus and with another refer-

6. For more information, visit the project website: http://lands.let.ru.nl/cgn/ (accessed October 30, 2017).

ence corpus, namely the Corpus of Spoken Dutch (CGN).[7] The results presented in this article overlap with and extend the results from the pilot study, revealing the same overall tendencies (Reviers et al. 2015). In brief, the results of the quantitative analysis support the hypothesis that there is indeed a language of AD, also in Dutch, with idiosyncratic features.[8]

2.3 Overview of the data

Tables 2 to 11 provide a summary of the features of the language of AD in Dutch, which are discussed in greater detail in the next section. Table 2 is an overview of the PoS distribution in our Dutch AD corpus. Note that this distribution is based on PoS tags that follow a formal, rather than a functional approach. This means that the Frog software tags nominal adjectives as adjectives, even if they function as nouns in the sentence at issue. Example: De **oude** liggen daar [*The **old** are there]. Recalculating frequencies to reflect the functional use of PoS is not unproblematic based on the current tags in the corpus, and certain categories overlap partially. For instance, there is a label "free" that applies to words that are used both adverbially or predicatively. In Table 3, therefore, adjectives as well as verbs used predicatively and adverbially are grouped together, which is not ideal for our purpose. However, Table 3 does put the data from Table 2 into perspective. For instance, where adjectives have a relative frequency of around 6% in the formal approach, these numbers drop to 3% in the functional perspective. In other words, adjectives are more often used in other ways than in the traditional prenominal position.

Table 4 summarises the comparison of our corpus to the two reference samples, SoNaR and Subtlex-nl and contains the p-values and effect sizes, which point to statistically significant differences compared to both reference corpora for all PoS categories measured ($p < 0.0001$). The effect sizes, however, point to rather small differences.

Table 5 lists the 50 most frequent lemmas and keywords, taking only the open class words into account. Keywords were sorted based on their effect sizes, taking the lemma frequencies from SoNaR as reference values. In other words, keywords are words that occur more frequently than in the SoNaR corpus.

---

7. The CGN is also part of the analysis in the PhD project that underpins this article, but these data were being processed at the time of writing and are not within the scope of this article. For more information on the CGN corpus, visit the project website: http://lands.let.ru.nl (accessed March 10, 2017).

8. The guidelines are available via the project website www.adlabproject.eu (accessed March 10, 2017).

Table 2. PoS distribution in the Dutch AD corpus (formal approach)

Part of speech category	Proportion
Open class	46.51%
adjectives	5.48%
nouns	19.61%
adverbs	3.22%
verbs	18.21%
Closed class	42.79%
articles	12.35%
interjections	0.02%
numerals	0.52%
conjunctions	4.15%
pronouns	9.89%
prepositions	15.86%
Other (abbreviations, symbols, foreign words)	10.69%

Table 3. PoS distribution in the Dutch AD corpus (functional approach)

Part of speech category	Proportion
Adjectives and verbs used prenominally or postnominally	2.8%
nouns	19.82%
common nouns	3.4%
proper nouns	16.42%
Adverbs and adjectives/verbs used adverbially or predicatively	9.74%
adverbially or predicatively used adjectives	3.05%
adverbially or predicatively used verbs	3.47%
Finite verbs	14.14%

Table 4. Inferential statistics for the Dutch AD corpus

	SoNaR		**Subtlex-nl**	
	goodness of fit (p)	**effect size (w)**	**goodness of fit (p)**	**effect size (w)**
adjectives	0.0000	0.02	0.000	0.15
nouns	0.0000	0.09	0.000	0.53
adverbs	0.0000	0.10	0.000	0.05
verbs	0.0000	0.14	0.000	0.21
articles	0.0000	0.14	0.000	0.62
conjunctions	0.0000	0.01	0.000	0.14
pronouns	0.0000	0.04	0.000	0.07
prepositions	0.0000	0.15	0.000	0.58

Table 5. Most frequent lemmas and keywords of open class words in the Dutch AD corpus

#	Lemmas			Keywords		
	lemma		**pos**	**lemma**		**pos**
1	*kijken*	[to look]	verb	*kijken*	[to look]	verb
2	*zijn*	[to be]	verb	*knikken*	[to nod]	verb
3	*komen*	[to come]	verb	*glimlachen*	[to smile]	verb
4	*staan*	[to stand]	verb	*stappen*	[to walk]	verb
5	*gaan*	[to go]	verb	*lopen*	[to run/walk]	verb
6	*lopen*	[to run]	verb	*rennen*	[to run]	verb
7	*zitten*	[to sit]	verb	*blik*	[gaze]	common noun
8	*Witse*	Witse [proper name]	common noun	*duwen*	[to push]	verb
9	*hand*	[hand]	common noun	*schudden*	[to shake]	verb
10	*weg*	[away]	adverb	*hand*	[hand]	common noun
11	*nemen*	[to take]	verb	*arm*	[arm]	common noun
12	*man*	[man]	common noun	*draaien*	[to turn]	verb
13	*stappen*	[to walk]	verb	*gezicht*	[face]	common noun
14	*hebben*	[to have]	verb	*deur*	[door]	common noun
15	*zien*	[to see]	verb	*raam*	[window]	common noun
16	*worden*	[to become]	verb	*trap*	[stairs]	common noun
17	*vrouw*	[woman]	common noun	*weg*	[away]	adverb
18	*geven*	[to give]	verb	*schouder*	[shoulder]	common noun
19	*auto*	[car]	common noun	*langzaam*	[slow]	predicative or adverbial adjective
20	*nog*	[again/more]	adverb	*staan*	[to stand]	verb
21	*oog*	[eye]	common noun	*pakken*	[to take]	verb
22	*terug*	[again/back]	adverb	*jas*	[jacket]	common noun
23	*hoofd*	[head]	common noun	*hoofd*	[head]	common noun
24	*deur*	[door]	common noun	*bed*	[bed]	common noun
25	*trekken*	[to pull]	verb	*kantoor*	[office]	common noun
26	*liggen*	[to lie]	verb	*openen*	[to open]	verb
27	*niet*	[not]	adverb	*bekijken*	[to look/to examine]	verb
28	*blijven*	[to stay]	verb	*tafel*	[tabel]	common noun
29	*blik*	[gaze]	common noun	*oog*	[eye]	common noun
30	*zitten*	[to sit]	verb	*gooien*	[to throw]	verb

Table 5. *(continued)*

#	Lemmas			Keywords		
	lemma		**pos**	**lemma**		**pos**
31	*dan*	[then]	adverb	*neer*	[down]	adverb
32	*gezicht*	[face]	common noun	*nemen*	[to take]	verb
33	*even*	[momentarily]	adverb	*glas*	[glass]	common noun
34	*houde*	[to hold]	verb	*rijden*	[to drive]	verb
35	*rijden*	[to drive]	verb	*auto*	[car]	common noun
36	*agent*	[police officer]	common noun	*trekken*	[to pull]	verb
37	*weer*	[again]	adverb	*haar*	[hair]	common noun
38	*zetten*	[to put]	verb	*zitten*	[to sit]	verb
39	*tafel*	[table]	common noun	*slaan*	[to hit]	verb
40	*dam*	[dam]	common noun	*stoppen*	[to stop]	verb
41	*bed*	[bed]	common noun	*open*	[open]	predicative or adverbial adjective
42	*draaien*	[to turn]	verb	*gang*	[hallway]	common noun
43	*knikken*	[to nod]	verb	*soldaat*	[soldier]	common noun
44	*halen*	[to get]	verb	*zitten*	[to sit]	infinitive
45	*vader*	[father]	common noun	*wit*	[white]	prenominal adjective
46	*leggen*	[to lay]	verb	*kamer*	[room]	common noun
47	*maken*	[to make]	verb	*agent*	[police officer]	common noun
48	*water*	[water]	common noun	*man*	[man]	common noun
49	*arm*	[arm]	common noun	*steken*	[to put]	verb
50	*moeder*	[mother]	common noun	*zijn*	[to be]	verb

3. The systemic functional linguistic analysis

3.1 Theoretical framework

In the previous section, we have argued that our data confirm the hypothesis that language use in AD is indeed idiosyncratic. What is particularly striking is that the overall language use is not only different to a high degree of statistical significance, but it also differs on many levels, as all PoS categories demonstrate significantly higher or lower frequencies than in our reference samples. The questions that follow then are:

1. How are these linguistic features used and
2. What does this tell us about the communicative function of this text type?

In order to answer these questions, specific PoS categories are studied from a Systemic Functional Linguistics perspective. SFL looks at language in use and relates the usage of lexico-grammatical features in particular texts to both the context in which they occur and the meaning they are meant to convey. Halliday's work in particular has been fundamental to functionalist and text linguistic approaches to language, and has inspired the disciplines to which the present analysis is indebted, such as Corpus Linguistics, Multimodality and AVT (see Sindoni 2011 and Royce 2007 for a combined approach of SFL and Multimodality).

Table 6 presents a concise summary of SFL theory (based on Eggins 1994, 113). A more detailed account is provided by general works on SFL such as Halliday (1994) and Eggins (1994). According to SFL, the aspects that determine how we use language in a particular context are Field, Tenor and Mode. In brief, what the text is about, who is talking to whom and through which specific channels. On the semiotic level, these aspects are expressed by three metafunctions of language: the experiential, the interpersonal and the textual function respectively. Each of these functions is covered by a specific lexico-grammatical system: transitivity, mood and theme organisation. Finally, a text only gains "texture", i.e., it only becomes a coherent piece of discourse, when it is also cohesive (Halliday 1994, 334). Cohesive devices can again be grouped according to the function they support, i.e., lexical relations for the experiential function, conversational structure for the interpersonal function, and reference and conjunction for the textual level.

Table 6. Systemic Functional Linguistics in brief

Context	Function	Lexico-grammar	Discourse semantics/cohesion
Field	experiential	transitivity	lexical relations
Tenor	interpersonal	mood	conversational structure
Mode	textual	theme/rheme	reference and conjunction

For the present article, those aspects of the context of language that are grouped under the umbrella term Field will be discussed. Eggins (1994, 52) describes Field as "what the language is being used to talk about", or which aspects of reality are represented. In other words, with respect to AD, it involves the linguistic choices describers make to translate the reality presented by the images of the audiovisual source text. Tenor, by contrast, refers to the interaction between participants and has to do with conversational structure and dialogue, which are not particularly significant for the descriptive units under analysis here. Finally,

Mode pertains to the way meaning in a text is organised and how the elements of a text are held together. This is especially important for the end product created by AD, as cohesion and coherence are crucial for a proper understanding by the target audience (Braun 2001, Reviers and Remael 2015). In the next sections, we adopt a bottom-up approach to describe what the lexico-grammatical and cohesive features specific to AD tell us about the experiential meaning expressed by AD.

3.2 Analysis

According to Halliday (1994, 106), "Language enables human beings to build a mental picture of reality". This is what he calls the "experiential function of language", which is precisely the function of AD: i.e., describing which reality the images depict, so that blind and visually impaired audiences can form a mental picture of it. This reality can be expressed by a manageable set of process types in the transitivity system. Each process consists of three components: (1) the process itself, expressed by "verbials" (Halliday 1994, 214); (2) the participants, represented by nominals; and (3) the circumstances, translated into adverbials and prepositional phrases. In this context, it is revealing that in the Dutch AD corpus the number of open class words outweighs that of closed class words: 46.5% open class versus 42.8% closed class words (see Table 2). Salway (2007) found the same preponderance of open class words in the TIWO project of English ADs. We see the exact opposite in our reference corpora SoNaR and Subtlex-nl, where closed class words dominate. This finding underlines the importance of the experiential function in AD, as open class words are particularly relevant to such functions. In other words, describers focus on transferring "what they see": who does what to whom and how. In what follows, we will look at each component of the transitivity system separately.

Table 7. Verb forms in the Dutch AD corpus

Part of speech category	Proportion in the corpus	Proportion within the verb category
Verbs	**18.2%**	**100%**
infinitives	1.9%	10.4%
present participles	0.3%	1.6%
finite past tenses	0.2%	1.1%
finite present tenses	14.0%	77%
past participles	1.2%	6.6%
nominal and prenominal verbs	0.6%	3.3%

3.2.1 *Processes – verbials*

Table 7 summarises the verb ratio in the Dutch AD corpus. Column 2 provides the total percentages in the whole corpus; the third column gives the distribution of the verbs separately.

The predominance of finite verb forms, especially in the present tense, stands out (77% of all verbs). Present tense verbs occur significantly more frequently than in both our reference corpora. This seems logical, given that AD tends to describe the action while it is happening, therefore preferring present tenses. The use of such tenses is also advised by European AD guidelines usually consulted by Dutch describers (Ofcom 2000; Remael et al. 2014) and was also highlighted in the TIWO project for English AD. Past tenses are rare and significantly less frequent than in our reference samples. Given that ADs use the present tense as they describe an ongoing narrative, it is interesting to see in which contexts the past tense is appropriate. The pilot study (Reviers et al. 2015) suggested that past tenses are used to link back to or specify characters (or actions) mentioned in previous descriptions. A qualitative analysis of the past tenses in the Dutch AD corpus confirmed this finding. We discovered that most cases consist of constructions with (defining) relative clauses to identify objects, characters or places that have been presented before. The examples below are a selection from the sentences analysed, in order to illustrate the use of tenses and contain the original sentences from the AD scripts and a back-translation.

(1) *Van In zit op het kantoor van de grote, kale man* die met zijn foto op de voorpagina stond. [Van In sits in the office of the tall, bald man who stood on the front page with his picture].

(2) *Ada staat aan de rand van het ravijn waar Esther het verhaal van de diamant vertelde.*
[Ada stands at the edge of the ravine where Esther told her the story of the diamant].

In the Dutch AD corpus, we also found a few past tenses that seem to reflect characters' thoughts; they describe facial expressions (as in Example 3) or flashbacks, with characters remembering previous events (as in Example 4).

(3) *Nu pas ziet Johnny dat Matty weg is. Dit had hij niet verwacht.*
[Only now does Johnny notice that Matty is gone. He did not expect this].

(4) *Hij zoekt zijn gsm en herinnert zich hoe hij hem kwijtspeelde.*
[He looks for his cell and remembers how he lost it].

In some cases, a past tense is used to describe an event after it has happened on screen, when time constraints prevent narrators from describing the action

simultaneously with the images. In most of these cases, the action still tends to be described in the present tense, since it immediately follows the images. But when there is too much time in between the image (and its related sounds) and the description, or when the description is interrupted by other sounds and dialogues, the past tense is used. This allows narrators to make clear to the listener that the description links up with previously heard dialogues or sound effects.

(5) *Gunther die in de huiskamer alles gevolgd had, gaat naar boven.*
[Gunther, who followed it all in the living room, goes upstairs].

(6) *Hij kijkt naar waar het geluid vandaan kwam.*
[He looks to where the sound had come from].

Finally, some of the descriptions in the Dutch AD corpus contain translated dialogue. When a character says a few lines in a foreign language, a voiced version of these subtitles for the benefit of the blind and visually impaired (called Audio Subtitling) is not always provided, but the translation of the dialogue is interwoven in the AD. Some of these lines contain past tenses, too.

(7) *Herinneringen spoken door het hoofd van Nazim. "Ik vraag een rat niet om vergiffenis." "Smeek zijn vergiffenis." "Ik deed wat ik moest doen, ik heb het ongedierte uitgeroeid." Nazim zit nog steeds in de luxueuze villa.*
[Memories haunt Nazim's mind. "I don't ask a rat for forgiveness". "Beg for his forgiveness". "I did what I had to do, I killed the pests". Nazim is still sitting in the luxurious villa].

What stands out most when it comes to the other verb forms is that present participles are more frequent in the Dutch AD corpus as compared to both reference samples. They are mainly used as adverbs, further specifying the action or reflecting facial expressions that accompany the action. The most frequent co-occurrences of verbs and present participles are: *vragend kijken* [*to look questioning], *zoekend kijken* [*to look searching], *glimlachend komen,* [*to come smiling] and *aarzelend komen* [*to come hesitating]. Another function of the present participle is to create reduced subordinate clauses indicating simultaneity (often in sentence-initial position). Simultaneity is indeed quite a challenge for describers: many actions happen simultaneously on screen in a film or a series, making it hard to render them in linear sentences (Braun 2011).

(8) *Haast slaapwandelend stapt Matty verder met haar dochter in de richting van de kassa.*
[Almost sleepwalking, Matty continues to the cash register with her daughter].

(9) *Vera komt aarzelend binnen en gaat tegenover haar moeder staan.*
[Vera comes in hesitantly and faces her mother].

The verbs in the Dutch AD corpus have relatively high frequency scores. The top 50 most frequent lemmas are all closed class words, except for 12 finite verbs. The 7 most frequent open class lemmas are all verbs as well (see Table 5): *kijken* [to look], *zijn* [to be], *komen* [to come], *staan* [to stand], *gaan* [to go], *lopen* [to walk], *zitten* [to sit]. When we look at the most noticeable keywords (keywords were calculated by comparing lemma frequencies between our corpus and SoNaR), we get a slightly different list: *kijken* [to look], *knikken* [to nod], *glimlachen* [to smile], *stappen* [to walk/to step], *lopen* [to walk] and *rennen* [to race]. The verbs in both of these lists, however, describe tangible, physical actions and express how the people performing them move around or what they are looking at. The focus of AD language on such material and behavioural processes is confirmed when we take into account which types of processes represented by verbs are most common in the Dutch AD corpus. Their frequencies are illustrated in Table 8 (which only includes verbs with normalised frequencies per thousand exceeding 10).

Table 8. Process types in the Dutch AD corpus

Process type	Proportion
behavioural	15%
existential	6%
material	75%
mental	1%
relational	2%
verbal	1%

Material and, to a lesser extent, behavioural processes dominate; while mental processes are particularly scarce. As the lemma analysis above has already revealed, these behavioural processes mainly include verbs indicating where people are looking (*kijken* [to look], *zien* [to see], *bekijken* [to look at/to examine], *staren* [to stare]). The most frequent mental processes are v*oelen* [to feel], *denken* [to think], *weten* [to know], *peinzen* [to ponder], *onderzoeken* [to examine], *begrijpen* [to understand] and *ontroeren* [to move emotionally]. It is worth noting that in many cases verbs reflecting mental processes are used to describe facial expressions that are visible on the screen.

(10) *Ze voelt zich betrapt.*
[She feels caught].

(11) *Sofie kijkt diep ontroerd toe.*
[Sofie watches deeply moved].

3.2.2 *Participants – nominals*

Participants are described by nominal groups that range from a single noun to quite complex structures including determiners, adjectives and qualifiers. What do our corpus data reveal about these nominal groups and, hence, the participants in AD texts? First, nouns make up the largest PoS category, with almost 20% of all words. They are significantly more frequent than in our reference samples. This is not surprising, since we have already established that the actions these participants perform (translated into verbs) are also more frequent. What is interesting, though, is that over 80% of the nouns are proper nouns and most of them are character names. For instance, 8 out of the 10 first keywords are proper names. The 20% common nouns in the Dutch AD corpus also refer mainly to human participants. For instance, the words *man* [man] and *vrouw* [woman] are in the top 3 of most frequent common nouns (Table 9). Other high frequency nouns involve the participants' body parts (hand, eye, head, arm). Also note that the nouns in our corpus usually refer to concrete and tangible things; abstract nouns are rare.

Table 9. Top 30 most frequent common nouns in the Dutch AD corpus

#	Lemma	Translation
1	hand	[hand]
2	man	[man]
3	vrouw	[woman]
4	auto	[car]
5	oog	[eye]
6	hoofd	[head]
7	deur	[door]
8	blik	[gaze]
9	gezicht	[face]
10	agent	[police officer]
11	tafel	[table]
12	dam	[dam]
13	bed	[bed]
14	vader	[father]
15	water	[water]
16	arm	[arm]
17	moeder	[mother]
18	kantoor	[office]
19	bureau	[desk]
20	foto	[photo]

A noun can be combined with a number of words to form a noun group, including determiners (articles, demonstratives, possessives), adjectives and qual-

ifiers. First, articles are significantly more frequent in the Dutch AD corpus, which is to be expected given the higher frequency of nouns. We also see that there are almost three times more definite articles than indefinite articles, which holds for our reference corpora too. Demonstratives, though, are significantly less frequent. A possible explanation can be found in the guideline to avoid fuzzy cohesion. Describers must make sure it is very clear which participant a demonstrative is referring to, so in case of doubt they prefer to repeat the noun referring to the participant (Ofcom 2000; Remael et al. 2014).

Next, adjectives have been the focal point of previous research in the field of AD (Arma 2011). They are deemed crucial to presenting a vivid, precise and engaging description. In our analysis, we can see that 5.5% of all words are adjectives. Interestingly, while in our pilot study adjectives occurred more frequently than in the reference corpus (which was CGN in the pilot), they occur less frequently as compared to the SoNaR corpus. This might be due to the nature of the reference corpora, since SoNaR reflects written language and CGN spoken language. It remains to be studied what type of language AD is: it is a spoken form of language, but the texts are prepared in advance in writing. Equally noteworthy is that the frequency of the adjectives in our corpus exceeds the frequencies from the Subtlex-nl corpus, which can be considered to be a 'hybrid' language variety as well, as it is the written version of spoken language. In brief, it is difficult to determine to what extent the frequencies of the adjectives are higher or lower than one would expect from this type of text.

That being said, Table 10 shows the adjective type ratios. What is noticeable is that these adjectives are used adverbially or predicatively rather than occurring in their traditional, prenominal function. In other words, adjectives are used more often to further specify a process (e.g., *Hij komt aarzelend binnen* [*He enters hesitating]) than a participant (e.g., *A tall, dark stranger*). The same observation can be made with regard to our Subtitling reference corpus, but not to our written corpus SoNaR, where prenominal adjectives dominate. We must not forget, however, that the analysis of the verbs has revealed the prenominal use of present and past participles. These account for 0.47% of all words. Previous research has shown a proportion of 1 adjective per 20 nouns in AD (Arma 2011). In our corpus, this proportion is 1 adjective per 7 nouns (also including the prenominal participles).

Contrary to the verbs, the adjectives in the Dutch AD corpus seem to have lower overall frequencies. This may be surprising, given the importance of adjectives in the guidelines as a means to give precise and vivid descriptions (Ofcom 2000; Remael et al. 2014). On the other hand, time constraints can explain why there is not always room for elaborate descriptions including adjectives. The first adjective appears on place 129 in the list of most frequent lemmas. The most frequent adjectives include: *groot* [tall/big], *ander* [other], *wit* [white], *zwart* [black],

Table 10. Proportions of adjectives in the Dutch AD corpus

Part of speech category	Proportion in the corpus	Proportion within the adjectives category
Adjectives	**5.95%**	**100%**
nominal adjectives	0.09%	1.5%
postnominal adjectives	0.02%	0.3%
prenominal adjectives	2.32%	39%
prenominal present participles	0.25%	4.2%
prenominal past participles	0.22%	3.7%
adjectives as adverbs of predicates	3.05%	51.3%

klein [small]. Other prenominal elements are infrequent and have normalised frequencies per thousand words <10. The most frequent adjectives in our corpus overlap to some extent with the findings from previous research (Reviers et al. 2015; Arma 2011; Salway 2007), which mention 'black', 'white', 'dark' and 'young' as the most frequently used adjectives.

Finally, another important aspect of nominal groups are what Halliday (1994) labels "qualifiers", i.e., defining relative clauses or prepositional phrases that qualify the noun in the group. We have already given a few such examples when discussing the past tenses, as they seem to occur in this type of construction (see Examples 1 and 2). These types of constructions are not annotated in our corpus and are therefore not quantifiable at this point in the study. However, the frequency of relative pronouns and prepositions might give us an indication. It is difficult to say anything meaningful about the frequency of relative pronouns, as their use differs according to the reference sample (more compared to CGN, less compared to SoNaR and considerably more compared to Subtlex-nl). In the AD corpus, however, relative pronouns are commonly used to construct defining relative clauses within nominal groups, as the examples below illustrate:

(12) *Ben bekijkt de beelden <u>die zijn klasgenoot maakte</u>.*
[Ben looks at the images <u>his classmate made</u>].

(13) *(…) de pil <u>die hij kreeg.</u>*
[(…) the pill <u>that he received</u>].

Prepositions are more frequent in our corpus, as compared to our reference corpora. The most frequent are: *in* [in], *van* [of], *op* [on], *naar* [towards], *met* [with], *aan* [on/to]. An analysis of their occurrences reveals that they are used with verbs (*met* [with], *aan* [on/to]), to denote locations (particularly *in* [in], *op* [on], *naar*

[towards]) (see discussion of the circumstances in Section 3.2.3) and in noun groups (particularly the prepositions *met* [with] and *van* [of]):

(14) *De auto van Werner*
[Werner's car].

(15) *Een man met lang, rossig haar*
[A man with long, reddish hair].

Finally, participants can be described by means of pronouns. Pronouns occur more often in the Dutch AD corpus than in SoNaR, but less often than in Subtlex-nl. About 76% of these pronouns are personal or possessive pronouns; in other words, pronouns referring to human participants. It is clear, though, that participants are more frequently described by nouns than by pronouns. Indeed, guidelines (such as Ofcom 2000) advise careful use of pronouns to avoid fuzzy coherence: when two descriptive units are interrupted by sound, music or dialogue, the referent might not be clear. The repetition of the noun or the name of a character or an object is preferred, as in the examples below.

(16) *Scalione heeft een wapen op zijn rug. Het zit tussen zijn broeksriem. Hij twijfelt even, maar laat de twee dan toch binnen. Hij wijst hen de weg, laat ze voorgaan.*
[Scalione holds a gun behind his back. It sits behind his belt. He hesitates for a brief moment, but lets both men in anyway. He shows them the way, lets them go first].

3.2.3 *Circumstances – adverbials and prepositional phrases*

The most straightforward category for expressing circumstances is that of adverbs. 3.22% of the words in our corpus are adverbs. They are less frequent than in both our reference corpora. Only 12 adverbs have normalised frequencies per thousand words >10 and they all refer to time (*nog* [still], *even* [for a while], *nu* [now], *dan* [then], *weer* [again]) or to characters' movement (*weg* [away] as in *weggaan* [to go away], *terug* [back] as in *terugkomen* [to come back], *neer* [down] as in *neerzitten* [to sit down]). Salway (2007) explained the lower frequency of adverbs in the TIWO project as follows: temporal information in AD is expressed through other means than adverbs of time, namely verb tense and the order of speaking. The events described are assumed to follow chronologically and do not require an adverb of time. Only in the rare case of a significant time lapse does this require specific mention.

Although AD guidelines emphasise the importance of adverbs for describing facial expressions and emotions, this does not seem to be reflected in our corpus. In this context, however, we need to underline the role of present participles that are used frequently as adverbs in AD (see 3.2.1) and often reflect characters' emo-

tions and/or facial expressions. A frequently occurring verb-adverb combination, for instance, is *glimlachend kijken* [*to look smiling].

Another important means to describe circumstances are prepositional phrases. As we have seen, prepositional phrases are used to qualify nouns. However, they also serve to express circumstances. Our quantitative analysis indicated the more frequent use of prepositions in AD compared to the reference corpora, while a qualitative analysis showed that they are mainly used to indicate place and time with the prepositions *in* [in], *op* [on], *naar* [to/towards].

(17) *Ze laadt de boodschappen in de koffer.*
[She puts the groceries in the car].

(18) *Er staat een muzikant in het zonnetje te spelen.*
[A musician is playing in the sunshine].

Even though a large part of the cases that come under the heading of "circumstances" were only analysed qualitatively, it becomes clear that the most common circumstances expressed are those of time, location and manner. Adverbs contribute somewhat to this function, however the use of present participles and prepositional phrases to express this level of meaning is what stands out most in our corpus.

3.2.4 *Lexical cohesion*

The discussion so far has shown that the vocabulary in the Dutch AD corpus is concentrated at the top of the frequency lists, i.e., it is characterised by a number of high frequency words. The first 128 most frequent words account for 50% of the vocabulary, and 58 of them are open class words. This suggests a high degree of word repetition, which is confirmed by an analysis of the Type Token Ratio (TTR) of our corpus. The TTR reflects the ratio between the number of unique words in the corpus (types) and the total number of words (tokens). In other words, a corpus with a low TTR will contain a considerable amount of repetition. Table 11 contains the Standardised Type Token Ratios (the TTR calculated per thousand words to minimise the effect of different text lengths) for the corpus in general and for the open class words individually. Vocabulary variety in the Dutch AD corpus is relatively low, with an overall STTR of 0.38, i.e., 0.38% unique words per 1.000 words. Variety is still considerably lower for verbs (0.27), pointing to an even higher degree of repetition in this category.

This high degree of repetition is revealing in terms of lexical cohesion, the aspect of discourse semantics that can be related to the experiential function (see Table 6). According to Halliday (1994, 330), lexical cohesion "comes about through the selection of items that are related in some way to those that have

Table 11. Standardised Type Token Ratios in the Dutch AD corpus

STTR Dutch AD corpus	0.38
STTR adjectives (all prenominal items)	0.41
STTR nouns (all nominal items)	0.49
STTR verbs	0.27
STTR adverbs	0.30

gone before". The first category is simple repetition. The STTR of the Dutch AD corpus indicates that this seems to be the most popular strategy in AD. A second important category is synonymy. An initial qualitative analysis of the most frequent words (with relative frequencies per thousand words >10), however, reveals that the list includes very few synonyms, namely: *kijken/zien/bekijken* [to look/to see/to examine], *lopen/gaan/stappen/rennen* [to run/to go/to walk/to race], *pakken/nemen* [to take/to get], *wagen/auto* [car/automobile] and *vrouw/meisje* [woman/girl].

Another subcategory of synonymy discussed by Halliday is that of hyponymy and meronymy, which apply to lexical items with a specific-general and part-whole relation respectively. The analysis of the nouns above, for instance, revealed that such relations are common in the Dutch AD corpus when describing (human) participants: among the frequent nouns we found *man* [man] and *vrouw* [woman], and subsequently lexical items referring to their body parts, such as hand, arm, eyes (meronymy). However, these types of semantic relations between words were not annotated in the Dutch AD corpus, so a quantitative analysis of this kind is beyond the scope of the present study.

4. Concluding remarks

The description of the data extracted from the Dutch AD corpus has first and foremost confirmed our hypothesis that there is a distinct language of audio description, and that this is also true for Dutch. The language of AD is characterised by a set of salient lexico-grammatical features. In addition, we found that AD language is idiosyncratic on all analysed levels, to a high degree of statistical significance, which is in line with the findings of our pilot study (Reviers et al. 2015). Low standard deviations also revealed consistency across scripts, at least in terms of PoS. This means that the choice of PoS seems to be mainly influenced by the inherent constraints imposed by the AD text type, rather than by genre or experience, for instance. It is also quite notable that AD language is characterised by extremely high or low values for certain grammatical categories: it contains very few past

tenses, nearly exclusively finite verbs, and a lot of proper names, to name just a few. Finally, AD texts are marked by a high degree of lexical repetition.

The discussion of the most salient lexico-grammatical features of Dutch AD from an SFL perspective has revealed what type of experiential meaning is expressed in this text type. In other words, we discovered which processes are mainly dealt with and, more importantly, how these processes are expressed linguistically. Material and, to a lesser extent, behavioural processes are preponderant. Actions, most often involving movement or gaze, are described in the present tense, as if the action was happening at that moment. Past tenses only occur in very specific contexts. In particular, the use of present participles to further specify the action stands out in this language variety. The participants, then, are mostly humans, identified through the use of proper names or, to a lesser extent, by concrete nouns pointing to their physique or body parts. Pronouns seem to be second choice when it comes to identifying participants. Prenominal adjectives are relatively frequent to further qualify participants, but the role of identifying relative clauses and prepositional phrases in this context is evident. Finally, the circumstances most commonly expressed are those of time, place and manner. Adverbs play a minor role in the expression of this type of information, but the use of present participles in this context is particularly noteworthy.

The focus of this article was on the experiential function of language, but the corpus provides similar insights for other metafunctions as well (for instance regarding reference and conjunction, sentence length and complexity, word length and reading speed), which however were not analysed. Moreover, the present article only covers the textual analysis of the verbal aspects of AD. Non-verbal aspects, such as sound effects and music, have not been taken into account. However, the hypothesis is, of course, that the findings, even on the linguistic level studied so far, are dictated by the interaction with these different sign systems, which the logic behind the data seems to indicate.

To confirm this hypothesis, a larger project currently underway will cover the other metafunctions as well. In addition, we have developed a multimodal concordancer to analyse the (verbal) corpus data in relation to the sound effects and music that accompany them. These multimodal aspects are particularly relevant to the analysis of textual cohesion and coherence. However, Multimodality is a relatively new discipline that is still in its infancy, and AVT is only now integrating some of its insights into its theoretical and methodological frameworks. Much work remains to be done to create a coherent framework for the analysis of audiovisual translations. In particular, methodological and technical challenges remain with regard to multimodal corpus development, a development to which we aim to contribute.

A final issue worth mentioning is that AD in the Low Countries is still developing. Current practice could therefore be expected to be rather heterogeneous, since professional describers have so far followed no standard, even though ADLAB guidelines (which have been recently translated into Dutch) are now being followed. Our data, however, reveal significant consistencies across the corpus, which arguably point to the impact of the intersemiotic functioning of the text. Nevertheless, analyses such as the one presented here should be rerun as new material is published and added to the corpus.

To conclude, the current project is one of few corpus projects in the field of AD and the first of its kind in the Low Countries. The above insights can be of particular value to, on the one hand, the fledgling professional field – providing concrete input for the development of guidelines, for instance – and, on the other hand, to the university and vocational trainings of describers. The results yielded by corpus studies in the field of AVT/AD are promising and it is envisaged that in the future, this research approach will further gain in importance, as advances are made in multimodal corpus development.

Acknowledgments

This paper reports on the results of a PhD project funded by the BOF fund of the University of Antwerp, under the supervision of Prof. Dr Aline Remael and Prof. Dr Reinhild Vandekerckhove (2012–2017). The author would like to thank a number of people who have contributed to the project: the CLiPS Research Group of the University of Antwerp for implementing an xml-interface for the Frog Software. Ron van den Brande of the "Centrum voor Teksteditie en Bronstudie" for his help with the development of an xml-tagset following the Text Encoding Initiative. The author is also indebted to Javier Serrano from the CAIAC Research Centre of the Autonomous University of Barcelona, for the development of the multimodal concordancer. Finally, thanks are due to all the professionals, companies and (volunteer) organisations who have supplied the material for the AD corpus.

References

Arma, Saveria. 2011. *The Language of Filmic Audio Description: A Corpus-Based Analysis of Adjectives*. PhD diss., Università degli Studi di Napoli Federico II,).

Bourne, Julian, and Catalina Jiménez. 2007. "From the Visual to the Verbal in Two Languages: A Contrastive Analysis of the Audio Description of *The Hours* in English and Spanish." In *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*, ed. by Jorge Díaz Cintas, Pilar Orero and Aline Remael, 175–187. Amsterdam: Rodopi.

Braun, Sabine. 2011. "Creating Coherence in Audio Description." *Meta: Translators' Journal* 56 (3): 645–662. https://doi.org/10.7202/1008338ar

Chmiel, Agnieszka, and Iwona Mazur. 2012. "AD Reception Research: Some Methodological Considerations." In *Emerging Topics in Translation: Audio Description*, ed. by Elisa Perego, 57–80. Trieste: EUT Edizioni Università di Trieste.

Eggins, Susanne. 1994. *An Introduction to Systemic Functional Linguistics*. London: Pinter Publishers.

Field, Andy. 2009. *Discovering Statistics Using SPSS*. London: Sage.

Fryer, Louise, and Jonathan Freeman. 2012. "Presence in Those with and without Sight: Implications for Virtual Reality and Audio Description." *Journal of Cyber Therapy and Rehabilitation* 5 (1): 15–23.

Fryer, Louise, and Jonathan Freeman. 2013. "Cinematic Language and the Description of Film: Keeping AD Users in the Frame." *Perspectives: Studies in Translatology* 21 (3): 412–426. https://doi.org/10.1080/0907676X.2012.693108

Gambier, Yves. 2006. "Multimodality and Audiovisual Translation." In *Audiovisual Translation Scenarios: Proceedings of the Marie Curie Euroconferences MuTra*, ed. by Mary Carroll, Heidrun Gerzymisch-Arbogast and Sandra Nauert.

Gambier, Yves. 2009. "Challenges in Research on Audiovisual Translation." In *Translation Research Projects*, ed. by Anthony Pym and Alexandra Perekrestenko, 17–15. Tarragona: Intercultural Studies Group.

Halliday, Michael A. 1994. *Introduction to Functional Grammar*. New York: Routledge.

Jiménez, Catalina, and Claudia Seibel. 2012. "Multisemiotic and Multimodal Corpus Analysis in Audio Description: TRACCE." In *Media for All: Audiovisual Translation and Media Accessibility at the Crossroads*, ed. by Aline Remael, Mary Carroll, and Pilar Orero, 409–425. Amsterdam: Rodopi.

Keuleers, Emmanuel, and Marc Brysbaert. 2010. "Subtlex-nl: A New Measure for Dutch Word Frequency Based on Film Subtitles." *Research Methods* 42 (3): 643–650. https://doi.org/10.3758/BRM.42.3.643

Kluckhöhn, K. 2005. "Informationsstrukturierung als Kompensationsstrategie: Audiodekription und Syntax." In *Horfilm. Bildkompensation durch Sprache*, ed. by Ulla Fix, 49–66. Berlin: Erich Schmidt Verlag.

Matamala, Anna, Pilar Orero, and Laura Puigdomenech. 2010. "Audio Description of Films: State of the Art and a Protocol Proposal." In *Perspectives on Audiovisual Translation*, ed. by Lukasz Bogucki and Krzysztof Kredens, 27–43. Frankfurt am Main: Peter Lang.

Nuzzo, Regina. 2014. "Statistical Errors: P Values, the 'Gold Standard' of Statistical Validity, Are Not As Reliable As Many Scientists Assume." *Nature* 506: 150–152. https://doi.org/10.1038/506150a

Ofcom. 2000. *ITC Guidance on Standards for Audio Description*. London: The Independent Television Commission.

Orero, Pilar. 2012. "Film Reading for Writing Audio Descriptions: A Word is Worth a Thousand images?" In *Emerging Topics in Translation: Audio Description*, ed. by Elisa Perego, 13–28. Trieste: EUT Edizioni Università di Trieste.

Pérez-González, Luis. 2014. *Audiovisual Translation: Theories, Methods, Issues*. London/New York: Routledge.

Remael, Aline, Gert Vercauteren, and Nina Reviers, eds. 2014. *Pictures Painted in Words: ADLAB Audio Description Guidelines*. Trieste: ADLAB.

Remael, Aline and Nina Reviers, forthcoming. "Multimodal Cohesion in Accessible Films: A First Inventory". In *The Routledge Handbook of Audiovisual Translation Studies*, ed. by Luis Perez Gonzales. London: Routledge.

Remael, Aline, Nina Reviers, and Reinhild Vandekerckhove. 2016. "From Translation Studies and Audiovisual Translation to Media Accessibility." *Target* 28 (2): 248–260. https://doi.org/10.1075/target.28.2.06rem

Reviers, Nina, Aline Remael, and Walter Daelemans. 2015. "The Language of Audio Description in Dutch: Results of a Corpus Study." In *New Points of View on Audiovisual Translation and Accessibility*, ed. by Anna Jankowska et al., 167–189. Bern: Peter Lang.

Reviers, Nina, and Aline Remael. 2015. "Recreating Multimodal Cohesion in Audio Description: A Case Study of Audio Subtitling in Dutch Multilingual Films." *New Voices in Translation Studies* 13: 50–75.

Reviers, Nina. 2016. "Audio Description in Europe: An Update." *The Journal of Specialised Translation* 26: 232–247.

Royce, Terry D. 2007. "Intersemiotic Complementarity: A Framework for Multimodal Discourse Analysis." In *New Directions in the Analysis of Multimodal Discourse*, ed. by Terry D. Royce and W. Bowcher, 63–109. Mahwah, NJ: Lawrence Earlbaum.

Salway, Andrew, Elia Tomadaki, and Andrew Vassiliou. 2004. "Building and Analysing a Corpus of Audio Description Scripts." (Research report). University of Surrey.

Salway, Andrew. 2007. "A Corpus-Based Analysis of Audio Description." In *Media for All: Subtitling for the Deaf, Audio Description and Sign Language*, ed. by Jorge Díaz Cintas, Pilar Orero, and Aline Remael, 151–174. Amsterdam: Rodopi.

Sindoni, Maria G. 2011. *Systemic-Functional Grammar and Multimodal Studies: An Introduction with Text Analysis*. Pavia: Ibis.

Van den Bosch, Busser Bertjan Antal, Sander Canisius, and Walter Daelemans. 2007. "An Efficient Memory-Based Morpho-Syntactic Tagger and Parser for Dutch." In *Computational Linguistics in the Netherlands: Selected Papers from the Seventeenth CLIN Meeting*, ed. by P. Dirix, I. Schuurman, V. Vandeghinste, and F. Van Eynde, 99–114. Leuven, Belgium: KUL.

Vercauteren, Gert. 2012. "Narratological Approach to Content Selection in Audio Description. Towards a Strategy for the Description of Narratological Time." *Monti: Multidisciplinary in Audiovisual Translation* 4: 207–231. https://doi.org/10.6035/MonTI.2012.4.9

Address for correspondence

Nina Reviers
Voetbalstraat 123
B-2800 Mechelen
Belgium

Nina.reviers@uantwerpen.be