

**Integrating corpus-based audio description tasks into an
intermediate-level German course**

Journal:	<i>International Journal of Applied Linguistics</i>
Manuscript ID	InJAL-19-030-Art.R1
Manuscript Type:	Original Article
Keywords:	audio description, corpus linguistics, German, foreign language teaching methodology / teacher education, tasks

Integration korpusbasierter Audiodeskriptionsaufgaben in einen Deutschkurs auf mittlerem Niveau

Dieser Artikel stellt eine Methode zur Entwicklung korpusbasierter Aufgaben vor. Er zeigt, wie diese in einen Deutschkurs (B2) für Lehramtsstudierende integriert wurden. Die Hauptaufgabe des Kurses war die Erstellung der Audiodeskription (AD) eines Filmvorspanns. ADs sind zusätzliche Tonspuren, die sehbeeinträchtigte Personen über visuelle Ereignisse informieren, die ihnen helfen sollen, die Handlung eines Films zu verstehen. Das Korpus *Audiodeskription*, per Sketch Engine zugänglich, enthält die AD-Skripte zweier deutscher Fernsehserien. Das Durchforsten dieses Korpus sollte den Lernenden helfen, ihre AD-Skripte sprachlich zu verbessern und ein Bewusstsein dafür zu entwickeln, wie räumliche Phänomene und visuelle Wahrnehmungen sprachlich ausgedrückt werden können. Der Artikel endet mit dem Vorstellen zweier korpusbasierter Lernaufgaben, die gezielt Fragen behandeln, die beim Durchführen des Kurses an der Sorbonne Université aufkamen.

Schlagwörter: Audiodeskription, Korpuslinguistik, Aufgaben, Deutsch als Fremdsprache, Lehrmethode

For Peer Review

Integrating corpus-based audio description tasks into an intermediate-level German course

This paper presents a method for designing corpus exploration tasks and integrating them into an intermediate-level German course during which learners collectively produce the audio description (AD) of a film trailer. ADs are additional soundtracks that inform blind people about the visual events that are considered necessary to understand a film. The Audiodeskription corpus, available on Sketch Engine, contains the AD scripts produced for two German television series. Exploring such a corpus to (re)write an AD script is an activity that is likely to foster language awareness of features of linguistic items needed for describing spatial events and visual perceptions. The tasks designed tackle learning issues identified during the course implemented at the teacher training department yyy (France).

Key words: audio description, corpus linguistics, tasks, German, foreign language teaching methodology

Introduction

Corpora are data collections that can be explored and represented in a variety of manners with the help of digital tools. Over the past thirty years, their potential for foreign language learning (FLL) has been widely discussed (Cobb & Boulton, 2015; Johns, 1988). However, one of the conclusions of Pérez-Paredes' metastudy (2019) is the fact that data-driven learning (DDL) and corpus exploration are not yet mainstream in the classroom. He considers the integration of DDL tasks into a course syllabus as a key factor for the normalisation of corpus use. This paper presents one method that could contribute to meeting this challenge.

When designing a foreign language learning scenario for groups in which learners have heterogeneous levels, one should strive to offer a range of flexible learning opportunities. Creating an audio description (AD) for a film trailer fits such a comprehensive learning frame. This task requires the development of competencies in the domain of oral comprehension (when watching the video), pragmatic skills (when selecting the parts to be audiodescribed), writing and text revision (when producing the AD script), oral reproduction (when reading it aloud), and technical skills (when recording the manuscript and aligning it with the video). The use of a corpus containing AD scripts is likely to support two of the competencies cited above: AD script production and text revision. Corpus exploration has also proven effective for fostering L2 language awareness (xxx, 2016).

This paper presents the way in which corpus-based AD tasks have been designed for a German course for which the exit level was B2 (*Common European Framework of Reference for Languages*, 2009). The students, eight French trainee-teachers of a non-linguistic discipline, were asked to collectively produce an audio description for the trailer of *Gundermann* (2018), a film about a famous singer in Eastern Germany. Inspired by Euler's design principles for learning environments (Euler, 2017), I developed corpus-based language awareness tasks on the basis of the learning needs identified prior to and during the course.

To better understand the types of errors that are likely to be made by learners expected to reach level B2 on an intermediate German course, I examined what learners of another course struggled with when executing a comparable writing task. I invited future German teachers, with a B2/C1 level, to write an AD script for a film trailer of their own choice. The first drafts of their texts contained language errors (marked by an asterisk or in bold) in the following areas:

- Word order: "Der Junge hält ein Referat über seine Mutter *im Klassenzimmer.>" [The boy is doing a paper on his mother *in the classroom.].

- Semantic choice of prepositions: "Ein Auto fährt *an Nacht in einer Stadt." [A car drives through the city at night.]

- Word endings (due to gender, number, or case errors), some of which concerned motion events: "Zwei junge Männer schieben ein Sofa in *einem *Schwimmbad." [Two young men push a sofa *in a pool.]

In addition, the students made discourse and/or pragmatic errors, such as providing irrelevant time-consuming details or lacking precision. They also selected elements which, in fact, did not need to be audiodescribed because they could be understood thanks to oral cues. It can be expected that learners majoring in disciplines other than German would make comparable errors, in particular in the field of morphology.

The first section of this paper presents the theoretical framework: corpora for FLL, audio descriptions for FLL, linguistic features of AD scripts in general, and, in particular, features of the Audiodeskription corpus (xxx, 2018). The second section describes the practical framework (AD script creation). The third section is devoted to the design of the study (sample, main task, tools, resources, research instruments, ethical frame, research focus). The fourth section lists the procedures adopted. In the fifth section, the collected data are presented and in the sixth one, they are analysed. The seventh section offers ready-to-use corpus exploration tasks likely to support morphology learning motivated by AD script writing tasks. The conclusion mentions the results and the limits of the study and offers suggestions for further research.

Theoretical framework

Corpora for foreign language learning

According to Teubert (2009), a corpus is "a principled collection of real language data, of texts (or parts of texts) sampling and thus representing a given discourse". Thanks to text analysis tools (e.g. concordancers), the corpus data can be explored and represented in various ways (e.g. thanks to graphics or tables). Also called data-driven learning (Johns, 1991), corpus use for FLL is slowly but steadily growing and now well documented by research (Boulton & Cobb, 2017; Chen & Flowerdew, 2018; Luo & Zhou, 2017). The main pedagogical concepts linked to corpus exploration are discovery learning (Bakker, 2018; Cheng, Warren, & Xun-feng, 2003) and learner autonomy (Ciekanski, 2014). Corpora provide students with information about language in use (Cheng et al., 2003; Ellis, 2017). Several studies suggest that corpus exploration enhances the students' knowledge of collocational behaviour (Li, 2017; Vyatkina, 2016) and that it supports language awareness processes (Boulton & Landure, 2016; xxx, 2016). Although none of the studies analysed by Crosthwaite (2017) provides evidence for a better mastery of morpho-syntactical features thanks to corpus exploration within the context of English for academic purposes, a study conducted with English-speaking learners of German L2 showed that data-driven learning improves the morpho-syntactic knowledge of learners (Vyatkina, 2016). This issue demands further research.

Audio description for foreign language learning

Audio descriptions (AD) are additional soundtracks informing blind or visually impaired people of the visual events considered necessary to understand a story told through a film. ADs link written texts to their spoken counterpart, which makes them relevant for research into embodied language processing (Buccino and Mezzadri 2015; Bottineau 2013). This multimodality helpfully dramatises motion events or static events and the description of people and objects. In her overview of research on this topic, Lertola (2019, p. 51) lists nine experimental studies published between 2013 and 2018. I added three studies to her sample. The target languages of the students concerned are English, German, Italian or Spanish. Five studies out of twelve focused on speaking skills (Ibáñez Moreno & Vermeulen, 2015a, 2015b, 2015c; Navarrete, 2018; Talaván & Lertola, 2016). Caldach and Talaván (2017) use AD tasks for the development of writing skills. Ibáñez Moreno and Vermeulen (2017) explore the link between AD production and the development of a better (meta)linguistic competence, and Vermeulen and Ibáñez Moreno (2017) focused on the topic of intercultural competence. Vermeulen and Ibáñez Moreno (2013) concentrated on the improvement of lexical and phraseological competence. Cenni and Izzo (2017) consider that AD could help to tackle challenges linked to morpho-syntactic features (pronouns, prepositions, verbs) or to lexical aspects (encouraging precision and variety), and they mention the (inter)cultural dimensions of AD tasks. Vermeulen and Ibáñez Moreno (2014) integrate AD into an English L2 course. Their results support the idea that "AD is an excellent didactic tool to promote the four language skills (reading, writing, listening, talking) in an integrative way". Burger (2016) details a method for the integration of this highly standardised text genre into a German L2 course. The scenario developed below builds on procedures and suggestions made in the latter paper.

Features of AD scripts

Audio descriptions share several features with movie storyboards and with stage directions used in drama scripts. They inform about essential aspects of visual or audiovisual events which can't be conceived solely by oral cues. The linguistic features of English AD scripts have been analysed by Salway (2007), who presents a preliminary linguistic study using corpus linguistics tools. He stresses the high degree of regularity of this text genre. Arma's paper (2011) focuses on adjectival functions in English AD scripts. Burger (2016) cites existing research about spatial verbs and perception verbs, adjectives, prepositional groups, and compound words (composed of nouns and/or adjectives), identified in English or Spanish AD scripts. In a corpus-based study on Dutch AD scripts, Reviere, Remael, Daelemans, Jankowska, and Szarkowska (2015) compare AD features to those of a general corpus (parts of the Corpus Spoken Dutch). They conclude that "there is a 'language of audio description' that differs considerably from general language", e.g., they are made of simple sentences, most of the verbs are in the present tense, and they contain more coordinating than subordinating conjunctions (see Reviere, 2018 for more details about lexico-grammatical features of Dutch AD scripts).

Features of the Audiodeskription corpus

The Audiodeskription corpus (xxx, 2018) contains data from two German television series. I was granted the right to use these data for teaching and research purposes. The first part of the corpus comprises the audio description manuscripts of 69 episodes of *Neues aus Büttenwarder* (Eberlein 1997-2017), a series that recounts rural stories and is set in Northern Germany. The second part of the corpus comprises 606 episodes from *Dahoam is Dahoam* (2007), a Bavarian soap opera. Film script prompts, as well as time and speech indications (e.g. "Szenenwechsel innerhalb des Textes, weiterlesen, durch Betonung deutlich machen" [Change of scenery within the text, read on, make it explicit through tone], are included in the corpus. Both data sets (including time and speech indications and film script prompts) were uploaded to Sketch Engine (Kilgarriff, Rychlý, & Pomikálek, n.d.), a corpus management system which specialises in lexicography. The data were automatically annotated with the RFTagger for German (Schmid & Laws, 2008), a tool that provides fine-grained part-of-speech tags, including case, gender and number. The files can be searched as a whole or separately.

--insert Figure 1– The Audiodeskription corpus on Sketch Engine.--

--

The sample below comes from the first part of the corpus and, more precisely, from episode 73 of the series *Neues aus Büttenwarder* (Eberlein, 1997). This script was created by an AD expert for the German broadcast channel NDR (Norddeutscher Rundfunk). It contains time indications, film script prompts surrounded by quotation marks, and speech indications highlighted in bold (e.g. "ss": speak very quickly; "s": speak quickly).

10:01:46

"Ich bin beim Essen!" ["I am eating!"]

ss Adsche nimmt Griem den Teller weg. [Adsche takes the plate away from Griem.]

10:01:51

"mal warm stellen?" – "Also, das ..." ["keep it warm?" – "Well, that ..."]

s Shorty geht mit dem Essen. [Shorty leaves with the food.]

An audio describer must strive to be brief, precise, and to show empathy towards visually impaired people. These requirements can be met by using specific forms, which can be identified in the Audiodeskription corpus. The corpus contains a high proportion of short sentences, and subordinates are rare. Salient features are preverbs, prepositions, compound adjectives and present participles. Verbs (except those used in the prompts) appear nearly exclusively in the third person of the present indicative tense, and some of the utterances do not contain any verb: "Jürgen in die Kamera: (...)" [Jürgen towards the camera: (...)].

--insert Figure 2– Frequent Preposition-noun groups in the Audiodeskription corpus.--

As displayed in figure 2ⁱ, in the Audiodeskription corpus, the most frequent collocations that start with a preposition and end with a noun are used for static location or motion; a means to set the scene. Two of the samples are in the accusative case, three in the dative case.

Practical framework: Creating an audio description

AD scripts are generally created by professional authors and/or translators, and they may be recorded either by humans (ideally actors or dubbing specialists) or by text-to-speech tools. The AD production process is structured by national and international norms, such as those described by Clark (2001). The world-wide web consortium (W3C) develops and publishes international standards for the Internet. One of their actions is the W3C Web Accessibility Initiative (WAI), which specialises in understanding and implementing accessibility. The WAI has produced a tutorial containing a basic workflow for creating pre-produced audio descriptionsⁱⁱ. This workflow, copied below, was used to help structure the German course put into place within the framework of this study. The corresponding AD creation tasks are mentioned in brackets. Tasks 3 and 4 are not part of the professional workflow. I added them for pedagogical reasons (for more details, see section [Procedures](#) and the complete AD project outlineⁱⁱⁱ).

- 1. Watch the video (tasks 1 and 5)**
Note opportunities for descriptive narration (i.e. pauses in narration or dialogue into which descriptions can be inserted).
- 2. Write the description script (tasks 1, 2, and 6)**
Use a description-authoring tool, a word processor or any text editor to create the script that will be recorded by the narrator.
- 3. Record the descriptions (task 7)**
Using audio-editing software, record the narrator reading the description script. (...)
- 4. Create an open-described audio track (task 8)**
Using audio-editing software, duplicate the original program audio track, then paste each audio description into this track at the appropriate time intervals. Add this new, open-described audio track to the presentation.

Design

Participants and institutional setting

Participants in the study reported upon in this paper were enrolled in the teacher training department of zzz (France). They were majoring in non-linguistic subjects: literature, music, primary education, or librarian studies. No participant was older than 30 years.

--insert table 1: Group composition.--

The students were offered a 20-hour course, held into ten 2-hour sessions, to obtain or confirm level B2 (*Common European Framework of Reference for Languages*, 2009) in German. The starting level of this group was quite heterogenous.

Main task, online tools, and resources

The participants were invited to collectively produce the audio description of the trailer of *Gundermann* (2018), a film about a successful singer from Eastern Germany, who worked for several years as a spy for the Stasi. The trailer was divided into eight parts. Each student was responsible for describing one part. The learning platform (*Edmodo*, 2008) was chosen for peer-to-peer and peer-to-tutor communication throughout the course and to share text or multimedia files. *VideoAnt* (n.d.) is a video annotation tool was used to write, comment, and revise the AD script drafts, as well as to later link them to the film trailer. YouDescribe (Miele, 2017) is a "free, web-based platform for adding audio description to YouTube content" (Miele, 2019). The participants relied on it to publish online their audio description part of the *Gundermann* trailer. The group members were granted access to the Audiodeskription corpus stored on Sketch Engine (see section [Audiodeskription corpus](#)) and to (*DeSkell*, 2019), a pedagogical version of Sketch Engine (see [Task 4](#)).

Research instruments and ethical frame

The participants were required to complete two pretest-like tasks. The online course tools provided digital traces of the learners' written and oral productions. The students were informed about the research focus of the course. All of them gave oral consent, but two out of eight students did not sign the ethical

consent form provided during the last session. One of the group members who had previously signed the ethical consent form dropped out of university before the end of the course. That is the reason why only five out of eight datasets are complete.

Research focus on morphology

Morphological features have not yet received much attention from AD researchers, perhaps because of the nature of the languages observed. German nouns, pronouns, and other items, e.g. adjectives, are morphologically marked to indicate gender, number, and case. Cases are triggered by adjectives, prepositions, or by transitive verbs (for more details, see the chapter "Morphologie" in Schäfer (2018)). They inform, inter alia, of the spatial features of the linguistic item. To give an example, a distinction is made between static location (dative case) and motion, i.e. moving from one place to another (accusative case). Adjectives and prepositions are salient and constituent elements of audio description manuscripts.

Thanks to a longitudinal study in French-speaking Switzerland, Diehl (2000, p. 198) found out that German adjective endings are generally acquired very late: only after an average of five years' study. This result is in line with evidence provided in the learner corpus *Merlin* (Wisniewski et al., 2013). Weiß (2017) studied the distribution of case use in assignments uploaded to this corpus and classified these according to five CEFR scores. According to her results, the case distribution displayed in figure 3 is a valid parameter for evaluating a learner's existing FLL level or his/her level obtained after language training. The dative case is much more often used by C1/C2-level learners than by A1-level learners. The latter tend to overuse the nominative case.

--insert Figure 3-- Case use and proficiency scores.--

Thus, morphological flexion (particularly adjective and noun declination) was chosen as one of the criteria for validating the B2 level of the study participants (see more details in section [Procedures](#)).

Procedures

Eight tasks were offered to the participants to help them create their AD script part. In order to identify their level in two of the competencies needed (selecting relevant information and writing), I invited them to start with two unprepared controlled tasks. During task 1, the learners were exposed three times to an excerpt (30 seconds) of the *Gundermann* movie (Gundermann, 2018), showing a group of coal miners discussing with a political leader. Gundermann, an excavator operator and a committed singer, criticizes openly the changing working conditions. Afterwards, the students had to write down a description of what they had seen. Task 2 asked the group to describe a screenshot of the same movie, showing Gundermann and his pregnant wife, reconciling after an argument.

The next two tasks offered pragmatic training to the group. The objective for both was to help learners put themselves in the shoes of a visually impaired or a blind person: task 3 consisted of listening to the *Gundermann* film trailer without seeing it. Then, the participants were asked to orally share the information gained through this reduced input. Afterwards, they were confronted with the audiovisual version of the trailer and they had to tell the group what they had previously missed. The second pragmatic task (task 4) was completed in pairs. The first student had to orally describe a screenshot taken from the film trailer. The second student had to draw a picture from the description provided.

Task 5 consisted of watching the *Gundermann* film trailer several times and taking notes about the parts which needed audio description. The students had to decide together which parts of the film trailer should be audiodescribed and choose the ones for which they wanted to be responsible. This segmentation was executed with the help of the online tool *VideoAnt* (n.d.). The learners put a marker at the beginning of the selected film segment (see figure 7). This marker was linked to a note-taking space, which was used to draft the audio description. In order to support the collective aspect of the audio description project, this *VideoAnt* was then shared with the group. For training purposes, before individually writing their audio description parts, the students were invited to describe one of the trailer parts not assigned to anyone in the group.

The participants were subsequently invited to produce their first audio description draft linked to their *VideoAnt* segment (task 6). To do so, they were explicitly allowed to use online dictionaries, e.g. *PONS* (n.d.) or *Linguee* | *Dictionnaire allemand-français* (2019) but no translator. They rewrote their text on

the basis of the comments I gave in the text file exported from *VideoAnt*. In a next step (task 7), the students recorded their AD script part with *Online Voice Recorder* (2012) and shared it on *Edmodo* (2008). Afterwards, I gave them advice, and they rehearsed the scripts thanks to the text-to-speech tool *Naturalreaders* (n.d.).

The tenth (and last) session of the German course was dedicated to individual text revision, to the technical mastery of YouDescribe (Miele, 2017), and to the recording of the AD script part on the students' private section of this AD adding platform (task 8).

Morphology was introduced to the students thanks to the present-practice-produce method (Carless, 2009), enhanced by language observation activities applied to AD scripts. On the basis of the morphological needs observed during the Gundermann course, I created corpus-based language awareness tasks which can be used during future audio script writing projects (see section [Understanding the function of word endings](#)).

During the course, the students were offered basic training for the use of *DeSkell* (2019) and Sketch Engine (Gilgarriff et al., n.d.). Corpus tools were applied, and search results discussed whenever it seemed helpful for language awareness and/or text revision.

Data presentation

In this section, the data stemming from the tasks that led to written productions (tasks 1, 2, 6, and 8) and to drawings (task 4) are presented.

Task 1

The group was invited to watch three times a 30-second long excerpt of the Gundermann movie and to take notes about places and the characters' actions.

--insert Figure 4– Louis' task 1.--

[Translation: There are a lot of people *in a construction site. A woman gives a man flowers. The man is happy. Gundermann has a question. He does not understand why the man drives a Rolls and not a Trabi.]

In the texts obtained, I counted the number of words, determiners (including quantifiers like "alle" [all]), and adjectives, and the number of correct endings and ending errors. As displayed in figure 4 and in table 2^{iv}, created with R ('R: The R Project for Statistical Computing', n.d.), this quite open task led to relatively few ending errors (median = 2.0). Louis made only one error in his short description, containing nine determiners (see figure 4).

--insert table 2: Endings in pre-task 1.--

Task 2

The second pre-task asked the students to describe a Gundermann movie screenshot. The only writing aid allowed was a word list translated into French: schwanger - enceinte [pregnant] | die Latzhose – la salopette [the dungarees] | der Fleck – la tache [the stain] | die Wand – le mur [the wall] | der Schwamm – l'éponge [the sponge] | streiten (stritt, gestritten) – se disputer [to argue, argued] | sich versöhnen – se réconcilier [to reconcile with each other].

Clara wrote a text without any adjective. She used the accusative case four times and the dative case three times. Seven of the word endings need revision.

--insert Figure 5– Clara's task 2.--

[Translation: There are two persons who share a love story. The woman is pregnant and the man looks preoccupied with this baby. Perhaps he is melancholic. It looks like a reconciliation. *In the wall is a *<spot of coffee>. Perhaps this photo was taken after the argument. The man has thrown coffee *in the wall.]

--insert table 3: Endings in pre-task 2.--

Despite the support provided for gender, the median value and the mean value of the number of the ending errors in task 2 are higher than in the texts corresponding to task 1. This can be explained by the

fact that the description of the screenshot is constraining and that it requires students to go into details, including spatial features.

Task 4

Task 4 was a describe-and-draw activity executed in pairs. When orally describing Gundermann's clothes represented on a screenshot, Inga, a student with a German family background, confused the word "darunter" ["under it"] with "drunten" ["down there"]: "*Drunten trägt er ein Hemd" [*Down there, he is wearing a shirt.] To study the use of the word "drunten", I let her explore this word on (*DeSkell*, 2019), which is a pedagogical version of Sketch Engine (Kilgarrieff et al., n.d.). Its content is based on so-called "good dictionary examples". The concordance lines obtained provide 40 examples, which display the spatial, even geographical, character of the word "drunten".

--insert Figure 6– Searching DeSkell for "drunten" [down there].--

[Translation: 1 Even without satellites, I like it down here. | Down in the valley, a small lake will form.]

Inga was not able to immediately exploit these lines. We discussed her semantic problem together, and afterwards, she went back to DeSkell to choose the right solution for her text. For space reasons, this sample only appears in its translation from German into English.

Teacher: Did you find something about "drunten"?

Inga: "Drunten" is actually, I think, hem, let's say, orally accepted, but not in writing.

Teacher: Orally? I'm not sure. (...) It doesn't work for me.

Inga: So, I should have said "unten rechts" [down on the right] and not "drunten" [down there].

Teacher: Ah, I thought you meant "unter der Jacke" [under the jacket].

Inga: I don't remember which word it was. It was...

Teacher: Wasn't it this one?

Inga: (...) He's got a blue jacket and under it... he's got a grey shirt.

Teacher: That means under the jacket, doesn't it?

Inga: Yes, yes, well.

Teacher: Well, I thought now that you would find out with DeSkell, that it does not work with clothes...

Inga: ...but rather with places.

Teacher: Yes, exactly.

Teacher: I'm on the mountain, yes, and down there is the valley or something. Okay?

Inga: Yes.

Task 6

Task 6 invited the participants to use *VideoAnt* to write the first draft of their AD script part. They decided together how to audiodescribe the segment entitled "Der Traum des Kommunismus [The dream of communism]".

--insert Figure 7– Using *VideoAnt* for multimedia text production.--

Gundermann und ein Stasispitzel sitzen in einem Restaurant. Sie haben viel Bier getrunken. Gundermann schlägt *(<den Tisch mit dem Glas>.

[Gundermann and a Stasi spy are sitting in a restaurant. They have drunk a lot of beer.

Gundermann hits the table with the glass].

As can be seen in the sample above, the students struggled with describing the gesture made by Gundermann, who, during an argument with a Stasi spy, violently slams his beer glass down on the table. I invited the group to revise Guillaume's text with the help of formulations identified in the Audiodeskription corpus. The query "schlägt" [hits] followed by "Tisch" [table] at a distance with the parameter between 1 and 4 words brought up six occurrences, each containing the preposition "auf" [on] commanding a noun in the accusative case. The group adapted the fifth sample, "Mit der Faust schlägt er auf den Tisch" [He slams the table with his fist.], to the current situation by replacing "Faust" [fist] with "Bierglas" [beer glass]. The fourth and the sixth concordance line contain an adverb which could have been taken over to highlight the emotional character of the situation: "verärgert" [out of anger] and "empört" [outraged].

--insert Figure 8– Transitive constructions with "schlagen" [to hit] and "Tisch" [table].--

[Translation of the samples: He hits the table and gets up. | Roland hits the table. | Out of anger, Benedikt slams the table. | He slams the table with his fist. | Outraged, she slams the table.]

The first AD script drafts^v obtained during the German course needed, inter alia, revision of word endings. The ending errors in Guillaume's first AD script draft are marked in bold; "xxx" indicates a missing ending.

Gundermann und seine Gruppe sind in der Bar, nach ihren **Konzert**. (...) Gundermann ist in **die Kabin** seines Baggers. Wir sehen **der** Bagger, der dreht und grabt in ein**xxx** riesig**xxx** Standort.

[Gundermann and his group are in the bar after their concert. (...) Gundermann is in the cabin of his excavator. We see the excavator which is turning and digging in a huge site.]

According to the data presented in table 4, gender is not the reason for the error endings; even with other genders, the endings would not match.

--insert table 4: Case errors in Guillaume's AD script (first draft).--

Some of the first AD drafts were not relevant from a pragmatic point of view: they contained verbs such as "sehen" [to see], which, in order to respect the condition and the feelings of blind people, had to be replaced by other ones. Other texts were too long. Shortening can be achieved by using compound words instead of longer word combinations or simply by deleting irrelevant text parts. In Guillaume's first draft, the elements put in bold had to be modified with relation to this.

Es ist spät. **Gundermann und seine Gruppe** sind in der Bar, nach ihren **Konzert**. **Die sechs andere Musiker sind gesetzt**. Sie trinken und rauchen. Gundermann nimmt seinen Gitarrenkoffern und geht **raus**. Morgen, Gundermann ist in **die Kabin seines Baggers**. **Wir sehen der Bagger, der dreht und grabt in ein riesig Standort**.

[It's late. **Gundermann and his group** are in the bar after their concert. **The six other musicians are sitting**. They drink and smoke. Gundermann takes his guitar *cases and goes **out**. Tomorrow, Gundermann is in **the cabin of his excavator**. **We see the excavator which is turning and digging in a huge site**.]

Guillaume's final draft contains 218 characters; that is one third less than the first draft, which is 317 characters long. Two compound words are used instead of longer word combinations ("Musikgruppe" [band] and "Baggerkabine" [excavator cabin]), and some words or sequences were deleted without being replaced, e.g. "raus" [out] and "sich drehen" [to turn].

(Es ist spät.) Die Musikgruppe ist in einer Bar, nach ihrem Konzert. Sie trinken und rauchen. Gundermann nimmt seinen Gitarrenkoffer und geht. Am nächsten tag ist er in seiner Baggerkabine und gräbt ein riesiges Loch.

[(It's late.) The music group is in a bar after their concert. They drink and smoke. Gundermann takes his guitar case and leaves. The next day, he is in his excavator cabin and digs a huge hole.]

During the AD writing period, Christine provides a first nearly perfect draft containing only one case error (she used the dative case instead of the accusative case to refer to a dynamic event), and the verb describing the gesture made by one of the protagonists in her scene is not precise enough: "Sie frisiert sich hinter ihm" [Standing behind him, she does her hair]. These words are finally replaced by "Sie steckt hinter ihm ihre Haare hoch" [Standing behind him, she puts her hair up].

[Translation of the third draft: Gundermann and Conny are in a blue bathroom. His upper body is bare. He is looking into the mirror. Conny wears a pullover and a coat. Behind him, she puts her hair up. He is making a grimace.]

--insert Figure 9– Gesture described in Christine's AD script.--

The Audiodeskription corpus was checked to know if Christine should have inserted the reflexive pronoun "sich" [herself] after "steckt" [puts up]. The corpus contains 19 relevant occurrences of the verb "hochstecken" [to put up], but none of them mentions "sich". The corpus data confirmed Christine's choice of verb construction.

Ihr langes, dunkles Haar hat sie zu einem Knoten hochgesteckt.

[She has put up her long dark hair to a knot.]

Task 8

Clara and Guillaume used the *YouDescribe* note-taking space to insert their final drafts, modified by the pronunciation and word stress aids provided.

--insert Figure 10– Excerpt of Clara's final draft on YouDescribe.--

[Translation: Pandora Film. Gundermann in the concert hall. Behind him is blue light. The camera is moving and points to his red braces.]

In Clara's text, upper-case letters indicate that a syllable must be stressed: "KonZertsaal" [concert hall]. Long vowels are repeated: "IIIhm" [him]. In the script copied in his note-taking space, Guillaume chose a capital letter at the end of "eiN" [a] so that he would not be tempted to add a final schwa.

Data analysis

According to the data presented in the previous section, corpus exploration helped the students understand the meaning of a word in a given context ("drunten" [down there] versus "darunter" [under]). They found out how to more precisely describe a spatial situation by using a prepositional phrase instead of an intransitive verb (see the beer glass example), and how to check verb constructions (presence or absence of "sich" [him- or herself]).

Morphology was one of the grammar contents selected for the Gundermann course. The offered teaching method was data-driven in the sense that the students were invited to observe word endings in printed AD scripts. It did not imply direct corpus consultation by the learners.

--insert table 5: Endings in the first AD script drafts.--

--insert table 6: Endings in the final AD script drafts.--

According to the results displayed in tables 5 and 6, the mastery of word endings is less satisfactory in the first AD drafts than in the final drafts. The median values corresponding to the ending errors in the data of the final drafts (see table 6) are lower than the median values data that correspond to the previous tasks. However, this last data set cannot be used to provide evidence for morphology learning: the pre-task lines are based on six data sets and the AD script lines only on five, and both AD drafts have been produced in an uncontrolled situation.

--insert table 7: Endings in four AD tasks (median values)."

Corpus-based morphology awareness tasks

The language awareness tasks presented in this section focus on morphology. They have been designed on the basis of the AD script writing needs that were observed.

The dative case

The dative case, combined or not with an adjective, is a frequent error source in the learners' texts. This case could be studied with the help of the Audiodeskription corpus by running a query retrieving adjectives tagged with the dative case and followed by any regular noun.

Instruction: Run the CQL query [tag="ADJ.*Dat.*"] [tag="N.Reg.*"]. In which contexts does the dative case appear? Which adjective endings can you identify in the nodes of these concordance lines?

--insert Figure 11– The dative case in adjective-noun groups.--

[Translation of the first five items: (...) suits with **red ties** | Waiters are standing in front of a **long table**. | Kuno chews with his **cheeks full**. | On a **big poster** in the background is written: (...) | With a **serious face** (...)]

Result: The contexts in which the dative case appears are static events or events implying a material or an immaterial instrument. The dative endings of the adjectives have three possible realisations: "-en", "-em", and "-er". Adjectives referring to plural nouns, like "Krawatten" [ties], all have the "-en" ending.

The functions of "-en"

The ending "-en" is often overused by German learners within the context of plural nouns. The following language awareness raising task, applied to the Audiodeskription corpus, can help students study its different functions.

Instruction 1: To find out the function of the ending "-en", run the CQL query "[word="*.en"]" – Frequency – Part-of-speech tags". Which elements are on the top of the list? Do plural nouns frequently end in "-en"?

--insert Figure 12– Elements ending with "-en".--

Result: The frequency list presenting words ending with "-en" starts with verbs and determiners. Feminine accusative plural nouns occupy rank 5 (N.Reg.Acc.Pl.Fem). The other noun plural forms occupy even lower ranks (12, 16, 17, etc.). As a by-product of this query, it can be observed that most of the elements tagged as neutral plural nouns ending in "-en" are deverbal nouns, like "das Lachen" [the laughing] and "das Klopfen" [the knocking].

Instruction 2: Run the CQL query "[tag="N.Reg.Nom.Pl.*"]" to retrieve plural nouns. Which are the most frequent endings?

Result: The most frequent endings are "-er" (Männer [men]), "-rn" (Bauern [farmers]), and "-en" (Frauen [women]). It can be observed that "-en" is not the only possible plural form and not the most frequent one in the Audiodeskription corpus.

Concluding remarks

This paper reports on the learning needs of a group of eight students that were identified during a course dedicated to AD creation. Corpus-based tasks were created on the fly and completed by the learners whenever possible without being a thread to the institutional goal, which was to obtain or confirm B2 level in the following five competencies: writing comprehension, oral comprehension, writing, speaking, oral interaction. The discussions I had with the students after their initial use of *DeSkell* (2019) or Sketch Engine (Kilgarrieff et al., n.d.) revealed that corpus exploration needs guiding to be fruitful.

The adopted morphology teaching method ("present-observe-practice-produce") has shown its limits during the Gundermann course. In order to help learners to better understand the role of the German word endings, morphology-oriented corpus-based tasks were designed on the basis of the results of this study. They are available for future editions of comparable German courses.

Another limit of this study is the unstable situation experienced with regard to course participation: two of the participants left university during the course and others were absent for one or more lessons. This made it impossible to collect a complete data set from all eight students. It would also have been beneficial to more precisely control how the last AD script drafts have been revised and to conclude by a post-test corresponding to task 2 dedicated to written screenshot description.

In the future, the Audiodeskription corpus data will be structured to distinguish audio descriptions from speech indications ("speak quickly") and from prompts, and the automatically created annotations will be checked in order to improve their reliability and to prepare the corpus for thorough linguistic description. On the basis of the improvements obtained, new editions of AD-oriented German courses can be implemented, including a range of more or less guided corpus exploration tasks. These will be best exploited if combined with metalinguistic discussions.

References

- Arma, S. (2011, November 29). The Language of Filmic Audio Description: A Corpus-Based Analysis of Adjectives [Tesi di dottorato]. <https://doi.org/10.6092/UNINA/FEDOA/8740>
- Bakker, A. (2018). Discovery learning: Zombie, phoenix, or elephant? *Instructional Science*, 46(1), 169–183. <https://doi.org/10.1007/s11251-018-9450-8>
- Bottineau, D. (2013). Pour une approche enactive de la parole dans les langues. *Langages*, N° 192(4), 11–27.
- Boulton, A., & Cobb, T. (2017). Corpus Use in Language Learning: A Meta-Analysis: Meta-Analysis of Corpus Use in Language Learning. *Language Learning*, 67(2), 348–393. <https://doi.org/10.1111/lang.12224>
- Boulton, A., & Landure, C. (2016). Using Corpora in Language Teaching, Learning and Use. *Recherche et Pratiques Pédagogiques En Langues de Spécialité - Cahiers de l'APLIUT*, (Vol. 35 N° 2). <https://doi.org/10.4000/apliut.5433>
- Buccino, G., & Mezzadri, M. (2015). Embodied language and the process of language learning and teaching. In U. M. Lüdtke (Ed.), *Consciousness & Emotion Book Series* (Vol. 10, pp. 191–208). <https://doi.org/10.1075/ceb.10.10buc>
- Burger, G. (2016). Audiodeskriptionen anfertigen – ein neues Verfahren für die Arbeit mit Filmen. *Informationen Deutsch Als Fremdsprache*, 43(1), 44–54. <https://doi.org/10.1515/infodaf-2016-0105>
- Calduch, C., & Talaván, N. (2017). Traducción audiovisual y aprendizaje del español como L2: El uso de la audiodescripción. *Journal of Spanish Language Teaching*, 4(2), 168–180. <https://doi.org/10.1080/23247797.2017.1407173>
- Carless, D. (2009). Revisiting the TBLT versus P-P-P Debate: Voices from Hong Kong. *Asian Journal of English Language Teaching*, 19, 49–66.
- Cenni, I., & Izzo, G. (2017). Audiodescrizione nella classe di italiano L2: Un esperimento didattico. *Incontri. Rivista Europea Di Studi Italiani*, 31(2), 9. <https://doi.org/10.18352/incontri.10169>
- Chen, M., & Flowerdew, J. (2018). A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom. *International Journal of Corpus Linguistics*, 23(3), 335–369. <https://doi.org/10.1075/ijcl.16130.che>

- Cheng, W., Warren, M., & Xun-feng, X. (2003). The language learner as language researcher: Putting corpus linguistics on the timetable. *System*, 31(2), 173–186. [https://doi.org/10.1016/S0346-251X\(03\)00019-8](https://doi.org/10.1016/S0346-251X(03)00019-8)
- Ciekanski, M. (2014). Les corpus: De nouvelles perspectives pour l'apprentissage des langues en autonomie ? *Recherches En Didactique Des Langues et Des Cultures*, 11(1). <https://doi.org/10.4000/rdlc.1710>
- Clark, J. (2001, 2007). Standard techniques in audio description. Retrieved 29 May 2018, from <https://joelclark.org/access/description/ad-principles.html>
- Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge Handbook of English Corpus Linguistics* (pp. 478–497). <https://doi.org/10.1017/CBO9781139764377.027>
- Common European Framework of Reference for Languages: Learning, teaching, assessment* (10. printing). (2009). Cambridge: Cambridge Univ. Press [u.a.].
- Crosthwaite, P. (2017). Retesting the limits of data-driven learning: Feedback and error correction. *Computer Assisted Language Learning*, 30(6), 447–473. <https://doi.org/10.1080/09588221.2017.1312462>
- Dahoam is Dahoam* [Soap opera]. (2007, today). Retrieved from <https://www.br.de/br-fernsehen/sendungen/dahoam-is-dahoam/index.html>
- DeSkell* [En]. (2019). Retrieved from <https://www.sketchengine.eu/deskell-german-corpus/>
- Diehl, E. (Ed.). (2000). *Grammatikunterricht, alles für der Katz? Untersuchungen zum Zweitspracherwerb Deutsch*. Tübingen: Niemeyer.
- Eberlein, N. (1997). *Neues aus Büttenwarder*. Retrieved from https://www.ndr.de/mediathek/mediatheksuche105_broadcast-129.html
- Edmodo*. (2008). Retrieved from <https://www.edmodo.com/>
- Ellis, N. C. (2017). Cognition, Corpora, and Computing: Triangulating Research in Usage-Based Language Learning: Cognition, Corpora, and Computing. *Language Learning*, 67(S1), 40–65. <https://doi.org/10.1111/lang.12215>
- Euler, D. (2017). Design principles as bridge between scientific knowledge production and practice design. *EDeR. Educational Design Research*, 1(1), 1–17. <https://doi.org/10.15460/eder.1.1.1024>
- Gundermann* [Biography, Drama, Music]. (2018). Retrieved from <https://www.imdb.com/title/tt7205942/>
- Ibáñez Moreno, A., & Vermeulen, A. (2015a). Profiling a MALL App for English Oral Practice A Case Study. *Journal of Universal Computer Science*, 21(10), 1339–1361.

- Ibáñez Moreno, A., & Vermeulen, A. (2015b). Using VISP (VIdeos for SPeaking), a Mobile App Based on Audio Description, to Promote English Language Learning among Spanish Students: A Case Study. *Procedia - Social and Behavioral Sciences*, 178, 132–138. <https://doi.org/10.1016/j.sbspro.2015.03.169>
- Ibáñez Moreno, A., & Vermeulen, A. (2015c). VISP 2.0: Methodological considerations for the design and implementation of an audio-description based app to improve oral skills. *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy*, 249–253. <https://doi.org/10.14705/rpnet.2015.000341>
- Ibáñez Moreno, A., & Vermeulen, A. (2017). *The ARDELE project: Audio description as a didactic tool to improve (meta)linguistic competence in foreign language teaching and learning*. Retrieved from <http://hdl.handle.net/1854/LU-8537626>
- Johns, T. (1988). Whence and Whither Classroom Concordancing? In T. Bongaerts, P. de Haan, S. Lobbe, & H. Wekker (Eds.), *Computer Applications in Language Learning* (pp. 9–27). Dordrecht: Foris.
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research, University of Birmingham*, 1–16.
- Kilgarriff, A., Rychlý, P., & Pomikálek, J. (n.d.). *Sketch Engine | language corpus management and query system*. Retrieved from <https://www.sketchengine.eu/>
- Lertola, J. (2019). *Audiovisual translation in the foreign language classroom: Applications in the teaching of English and other foreign languages*. <https://doi.org/10.14705/rpnet.2019.27.9782490057252>
- Li, S. (2017). Using corpora to develop learners' collocational competence. *Language Learning & Technology*, 21(3), 153–171.
- Linguee | Dictionnaire allemand-français. (2019). Retrieved from <https://www.linguee.fr/francais-anglais>
- Luo, Q., & Zhou, J. (2017). Data-driven Learning in Second Language Writing Class: A Survey of Empirical Studies. *International Journal of Emerging Technologies in Learning (IJET)*, 12(03), 182. <https://doi.org/10.3991/ijet.v12i03.6523>
- Miele, J. (2017). *YouDescribe*. Retrieved from <https://youdescribe.org/>
- Miele, J. (2019). YouDescribe [SKERI]. Retrieved 4 October 2019, from The Smith Kettlewell Eye Research Institut website: <https://www.ski.org/project/youdescribe#targetText=YouDescribe%20is%20a%20free%2C%20web,audio%20description%20to%20YouTube%20content>.
- Naturalreaders. (n.d.). Retrieved from <https://www.naturalreaders.com/online/>

- Navarrete, M. (2018). The use of audio description in foreign language education: A preliminary approach. *Translation and Translanguaging in Multilingual Contexts*, 4(1), 129–150.
<https://doi.org/10.1075/ttmc.00007.nav>
- Online Voice Recorder. (2012). Retrieved from <https://online-voice-recorder.com>
- PONS - Le dictionnaire gratuit pour les langues étrangères, l'orthographe allemande et la traduction de textes. (n.d.). Retrieved from <https://fr.pons.com/traduction>
- R: The R Project for Statistical Computing. (n.d.). Retrieved 31 July 2018, from <https://www.r-project.org/>
- 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>
- Reviere, N., Rémel, A., Daelemans, W., Jankowska, A., & Szarkowska, A. (2015). The Language of Audio Description in Dutch: Results of a Corpus Study. In *New Points of View on Audiovisual Translation and Media Accessibility* (pp. 167–190). Retrieved from <http://www.clips.ua.ac.be/~walter/papers/2015/rrd15.pdf>
- Salway, A. (2007). A Corpus-based Analysis of Audio Description. In J. Cintas Díaz, P. Orero, & A. Rémel (Eds.), *Media for All: Subtitling for the Deaf, Audio Description, and Sign Language* (pp. 151–174). Amsterdam: Rodopi.
- xxx. (2018). *Audiodeskription corpus*.
- xxx. (2016). Talking about German verb particles identified in concordance lines – from spontaneous to expert-like metatalk. *Language Awareness*, 25(1–2), 127–143.
<https://doi.org/10.1080/09658416.2015.1122023>
- Schäfer, R. (2018). *Einführung in die grammatische Beschreibung des Deutschen*. Retrieved from <http://langsci-press.org/catalog/book/224>
- Schmid, H., & Laws, F. (2008). Estimation of Conditional Probabilities with Decision Trees and an Application to Fine-grained POS Tagging. *Proceedings of the 22Nd International Conference on Computational Linguistics - Volume 1*, 777–784. Retrieved from <http://dl.acm.org/citation.cfm?id=1599081.1599179>
- Talaván, N., & Lertola, J. (2016). Active audiodescription to promote speaking skills in online environments. *Sintagma: Revista de Lingüística*, 28, 59–74. <https://doi.org/10.21001/sintagma.2016.28.04>
- Teubert, W. (2009). Corpus Linguistics: An Alternative. [Texte anglais original]. *Semen. Revue de Sémiotique Linguistique Des Textes et Discours*, (27). Retrieved from <http://journals.openedition.org/semen/8912>

- Vermeulen, A., & Ibáñez Moreno, A. (2013). Audio description as a tool to improve lexical and phraseological competence in foreign language learning. In D. Tsagari & G. Floros (Eds.), *Translation in language teaching and assessment* (pp. 45–61). Retrieved from <http://hdl.handle.net/1854/LU-4211313>
- Vermeulen, A., & Ibáñez Moreno, A. (2014). La audiodescripción como recurso didáctico en el aula de ELE para promover el desarrollo integrado de competencias. In *New directions in hispanic linguistics* (pp. 263–292). Retrieved from <http://hdl.handle.net/1854/LU-4211345>
- Vermeulen, A., & Ibáñez Moreno, A. (2017). Audio Description as a Tool to Enhance Intercultural Competence. In *Towards Transcultural Awareness in Translation Pedagogy* (pp. 133–156). Retrieved from https://www.researchgate.net/publication/321124982_Audio_Description_as_a_Tool_to_Enhance_Intercultural_Competence
- VideoANT*. (n.d.). Retrieved from <http://ant.umn.edu/>
- Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL*, 28(2), 207–226. <https://doi.org/10.1017/S0958344015000269>
- Weiß, Z. L. (2017). *Using Measures of Linguistic Complexity to Assess German L2 Proficiency in Learner Corpora under Consideration of Task-Effects* (MA Thesis). Universität Tübingen, Tübingen.
- Wisniewski, K., Schöne, K., Nicolas, L., Vettori, C., Boyd, A., Meurers, D., ... Hana, J. (2013). *MERLIN: An Online Trilingual Learner Corpus Empirically Grounding the European Reference Levels in Authentic Learner Data*. Retrieved from http://conference.pixel-online.net/ICT4LL2013/common/download/Paper_pdf/322-CEF03-FP-Wisniewski-ICT2013.pdf

Tables

Anonymized first name	Gender	First language	Intended occupation	CEFR starting level in writing	AD script provided	Ethical form signed
Didier	male	French	music teacher	A1+	yes	yes
Hugo	male	French	music teacher	A2	yes	no
Aldina	female	French	primary school teacher	A2-	no	no
Clara	female	French	French teacher	B1	yes	yes
Guillaume	male	French	French teacher	B1	yes	yes
Louis	male	French	French teacher	B1	yes	yes
Christine	female	French	school librarian	B2+	yes	yes
Inga	female	French	primary school teacher	C1-	no	yes

Table 1–Group composition.

Words	Determinants	Adjectives	Correct endings	Ending errors
Min. : 29.00	Min. : 4.00	Min. : 0.00	Min. : 1.000	Min. : 1.000
1st Qu.: 46.00	1st Qu.: 9.25	1st Qu.: 1.25	1st Qu.: 8.000	1st Qu.: 2.000
Median : 76.50	Median : 11.50	Median : 2.50	Median : 9.500	Median : 2.000
Mean : 71.33	Mean : 10.83	Mean : 2.50	Mean : 9.333	Mean : 3.167
3rd Qu.: 94.25	3rd Qu.: 13.00	3rd Qu.: 3.75	3rd Qu.: 13.250	3rd Qu.: 2.750
Max. : 110.00	Max. : 16.00	Max. : 5.00	Max. : 14.000	Max. : 9.000

Table 2–Endings in pre-task 1.

Words	Determinants	Adjectives	Correct endings	Ending errors
Min. : 38.00	Min. : 7.00	Min. : 1.000	Min. : 1.000	Min. : 1.00
1st Qu.: 45.50	1st Qu.: 9.25	1st Qu.: 1.000	1st Qu.: 6.500	1st Qu.: 1.25
Median : 57.50	Median : 10.00	Median : 1.500	Median : 8.500	Median : 4.00
Mean : 54.83	Mean : 10.17	Mean : 2.167	Mean : 8.167	Mean : 4.00
3rd Qu.: 62.00	3rd Qu.: 11.50	3rd Qu.: 3.500	3rd Qu.: 9.750	3rd Qu.: 6.75
Max. : 71.00	Max. : 13.00	Max. : 4.000	Max. : 15.000	Max. : 7.00

Table 3–Endings in pre-task 2.

Sample	Used case	Needed case	Preposition	Adjective
nach ihren Konzert > nach ihrem Konzert	accusative	dative	x	
ist in die Kabin > ist in der Kabine	nominative	dative	x	
Wir sehen der Bagger > Wir sehen den Bagger	nominative	accusative		
in ein riesig Standort > an einem riesigen Standort	nominative? accusative?	dative	x	x

Table 4–Case errors in Guillaume's AD script (first draft).

Words	Determinants	Adjectives	Correct endings	Ending errors
Min. : 19	Min. : 2.0	Min. : 0.0	Min. : 1.0	Min. : 0.0
1st Qu.: 33	1st Qu.: 5.0	1st Qu.: 2.0	1st Qu.: 2.0	1st Qu.: 1.0
Median : 34	Median : 5.0	Median : 2.0	Median : 3.0	Median : 5.0
Mean : 40	Mean : 5.2	Mean : 2.2	Mean : 2.8	Mean : 3.4
3rd Qu.: 51	3rd Qu.: 6.0	3rd Qu.: 2.0	3rd Qu.: 3.0	3rd Qu.: 5.0
Max. : 63	Max. : 8.0	Max. : 5.0	Max. : 5.0	Max. : 6.0

Table 5–Endings in the first AD script drafts.

Words	Determiners	Adjectives	Correct endings	Ending errors
Min. : 33	Min. : 5	Min. : 2.0	Min. : 2	Min. : 0.0
1st Qu.: 34	1st Qu.: 5	1st Qu.: 2.0	1st Qu.: 5	1st Qu.: 0.0
Median : 35	Median : 6	Median : 2.0	Median : 9	Median : 0.0
Mean : 38	Mean : 6	Mean : 2.8	Mean : 7	Mean : 1.4
3rd Qu.: 35	3rd Qu.: 7	3rd Qu.: 3.0	3rd Qu.: 9	3rd Qu.: 2.0
Max. : 53	Max. : 7	Max. : 5.0	Max. : 10	Max. : 5.0

Table 6–Endings in the final AD script drafts.

task	number of words	number of correct endings	number of ending errors
pre-task 1	76.5	9.5	2
pre-task 2	57.5	8.5	4
AD script first draft	34	3	5
AD script final draft	35	9	0

Table 7–Endings in four AD tasks (median values).

Figures

Figure 1–The Audiodeskription corpus on Sketch Engine.

Figure 2–Frequent Preposition-noun groups in the Audiodeskription corpus.

Figure 3–Case use and proficiency scores.

Figure 4–Louis' task 1.

Figure 5–Clara's task 2.

Figure 6–Searching deSkell for "drunten" [down there].

Figure 7–Using VideoAnt for multimedia text production.

Figure 8–Transitive constructions with "schlagen" [to hit] and "Tisch" [table].

Figure 9–Gesture described in Christine's AD script.

Figure 10–Excerpt of Clara's final draft on YouDescribe.

Figure 11–The dative case in adjective-noun groups.

Figure 12–Elements ending with "-en".

ⁱ The translations into English on this screenshot are mine.

ⁱⁱ Source: https://www.w3.org/2017/04/xsmedia-tute/page6.html#tts_dx

ⁱⁱⁱ <https://drive.google.com/file/d/1yqpjMTf68bFcFpsIAgPIxfuQ0MT6K820/view?usp=sharing>

^{iv} Table 2 and 3 represent the data corresponding to the six students having signed the ethical consent form.

^v The data discussed (see table 5 and table 6) are those of the five students having signed the ethical consent form and having produced an AD script.

#	Original file	Plain text	Vertical	Words ?
1	buettenwarder.vert	✓	✓	131,566
2	dahoom.vert	✓	✓	941,456

Figure 1–The Audiodeskription corpus on Sketch Engine.

104x18mm (240 x 240 DPI)

	Word	↓ Frequency	Frequency per million		
1	in die Kamera [towards the camera]	608	321.97	<div></div>	...
2	in der Wohnküche [in the kitchen living]	475	251.54	<div></div>	...
3	auf den Tisch [on the table]	471	249.42	<div></div>	...
4	In der Metzgerei [In the butchery]	470	248.89	<div></div>	...
5	in der Hand [in his her hand]	451	238.83	<div></div>	...

Figure 2–Frequent Preposition-noun groups in the Audiodeskription corpus.

106x41mm (220 x 220 DPI)

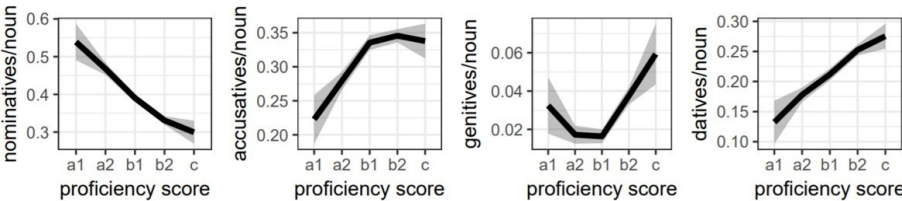


Figure 3–Case use and proficiency scores.

218x54mm (150 x 150 DPI)

1. Filmbeschreibung

Sie werden drei Mal einen kurzen Filmausschnitt sehen. Notieren Sie alles, was Sie sehen, so detailliert wie möglich, z.B.: Wer ist wo? Wer macht was?

Es gibt viele Leute in einer Baustelle.
Eine Frau gibt einem Mann Blumen.
Der Mann ist frohlich.
Gunderson hat eine Frage.
Er versteht nicht warum der Mann mit einem Ball und nicht mit einem Traktor fährt.

Figure 4-Louis' task 1.

145x61mm (220 x 220 DPI)

Hier gibt es zwei Personen die ein Liebes Geschicht
teilen. Die Frau ist Schwanger und der Mann
sieht preoccupiert von dieses Baby aus. Vielleicht ist
er Melancholisch. Es sieht wie ein Versöhnung aus.
In der Hall gibt es einen Fleck von Kafe.
Vielleicht dieses Foto ist nach der Streitung
genommen.
Der Mann hat der Kafe in der Hall werfen.

Figure 5-Clara's task 2.

88x51mm (220 x 220 DPI)

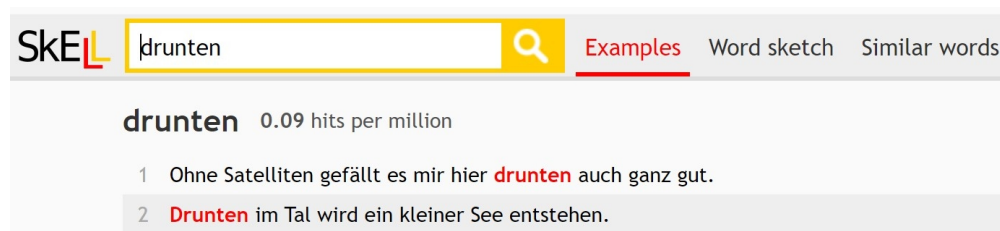


Figure 6–Searching DeSkell for "drunten" [down there].

146x34mm (240 x 240 DPI)

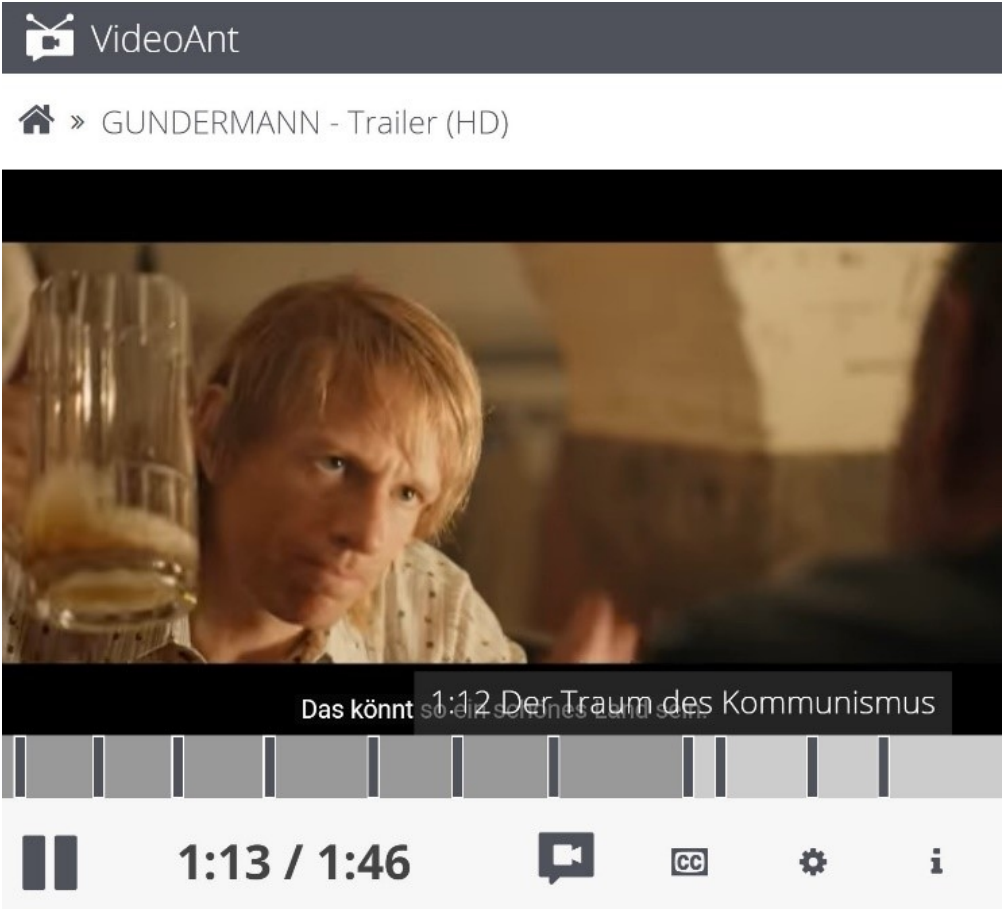


Figure 7–Using VideoAnt for multimedia text production.
85x77mm (220 x 220 DPI)

Left context	KWIC ↓	Right context
) "es langt, Moni." Er	schlägt auf den Tisch	und erhebt sich. 08:28 "ni
7 (Klopfgeräusch) Er	schlägt auf den Tisch	. 22:21 "Super hibracht." (
1 langweilig." Roland	schlägt auf den Tisch	. 09:47 "a ohne dei Spui."
1?" – "Na." Verärgert	schlägt Benedikt auf den Tisch	. 10:42 (Schlaggeräusch)
schlag) Mit der Faust	schlägt er auf den Tisch	. 10:06 "Das gibst doch ne
(über "oiso") Empört	schlägt Moni auf den Tisch	. 06:46 "a auf meim Zimm

Figure 8–Transitive constructions with "schlagen" [to hit] and "Tisch" [table].

81x37mm (220 x 220 DPI)



Figure 9–Gesture described in Christine's AD script.

96x42mm (220 x 220 DPI)

Pandora Film. Gundermann Im KonZertsaal.
Hinter IIIhm ist blaues Licht. Die Kamera
bewegt sich und zeigt auf seine rOten
HOsenträger.

Figure 10–Excerpt of Clara's final draft on YouDescribe.

76x20mm (220 x 220 DPI)

schwarze Haare und tragen schwarze Anzüge mit	roten Krawatten	. 112: 10:14:05:16 10:14:23:13 17:22 ++
> An der Terrassenseite stehen Kellner vor einem	langen Tisch	. </s><s> Darauf Töpfe mit Kartoffeln. </
:offel. ++ Schmeckt gut, Schorty. ++ Kuno kaut mit	vollen Backen	. </s><s> Brakelmann spießt mit der Ga
iem Kreis aus Strohballen auf. </s><s> Auf einem	großen Plakat	im Hintergrund steht: "Moin Moin China"
en. ++ 147: 10:19:01:21 10:19:06:07 04:11 (s) Mit	ernster Miene	steigen die Chinesen aus. </s><s> Brak
ht von außen an der Tür. </s><s> Sie fällt aus der	oberen Angel	. </s><s> Er betritt die Scheune und kni
d fotografieren die drei Chinesen in dem schwach	beleuchteten Raum	. </s><s> Hühner hocken auf Gerümpel
uto. </s><s> Grimmig sieht Brakelmann ihnen mit	gesenktem Kopf	durchs Fenster nach. 170: 10:22:07:23 '

Figure 11–The dative case in adjective-noun groups.

121x45mm (220 x 220 DPI)

	Tag	↓ Frequency
1	VFIN.Full.3.Pl.Pres.Ind	12,420
2	ART.Def.Acc.Sg.Masc	11,846
3	VINF.Full	7,296
4	ART.Indef.Acc.Sg.Masc	4,918
5	N.Reg.Acc.Pl.Fem	3,914
6	N.Reg.Nom.Sg.Neut	3,722
7	N.Reg.Dat.Sg.Masc	2,782
8	APPR	2,530
9	ADJA.Pos.Acc.Sg.Masc	2,399
10	ADJD.Pos	1,961

Figure 12—Elements ending with "-en".

68x86mm (220 x 220 DPI)