

Online-Supplement

Audiodeskription für Unterrichtsvideos

Eine anwendungsorientierte Erläuterung
anhand von Pilotvideos aus den *Degree*-Projekten

Carsten Bender, Leevke Wilkens,
Finnja Lüttmann & Christian Bühler

Online-Supplement 4:
Pilotvideo für das Fach Musik.
Darstellung der Besonderheiten des Videos und der Audiodeskription

Leevke Wilkens^{1,*}, Finnja Lüttmann¹,
Carsten Bender¹ & Christian Bühler¹

¹ Technische Universität Dortmund

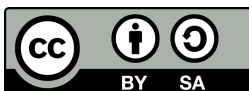
* Kontakt: Technische Universität Dortmund,
Fachgebiet Rehabilitationstechnologie,
Emil-Figge-Str. 50, 44227 Dortmund
leevke.wilkens@tu-dortmund.de

Zitationshinweis:

Bender, C., Wilkens, L., Lüttmann, F. & Bühler, C. (2025). Audiodeskription für Unterrichtsvideos. Eine anwendungsorientierte Erläuterung anhand von Pilotvideos aus den *Degree*-Projekten [Online-Supplement 4: L. Wilkens, F. Lüttmann, C. Bender & C. Bühler: Pilotvideo für das Fach Musik. Darstellung der Besonderheiten des Videos und der Audiodeskription]. *DiMawe – Die Materialwerkstatt*, 7 (2), 133–147. <https://doi.org/10.11576/dimawe-7817>

Online verfügbar: 18.03.2025

ISSN: 2629–5598



Dieses Werk ist freigegeben unter der Creative-Commons-Lizenz CC BY-SA 4.0 (Weitergabe unter gleichen Bedingungen). Diese Lizenz gilt nur für das Originalmaterial. Alle gekennzeichneten Fremdinhalte (z.B. Abbildungen, Fotos, Tabellen, Zitate etc.) sind von der CC-Lizenz ausgenommen. Für deren Wiederverwendung ist es ggf. erforderlich, weitere Nutzungsgenehmigungen beim jeweiligen Rechteinhaber einzuholen. <https://creativecommons.org/licenses/by-sa/4.0/legalcode.de>

Vorbemerkung

Die Videoausschnitte und Aufgabenstellungen, die in diesem Steckbrief dargestellt werden, wurden von Johanna Langner im Teilprojekt Musik im *Degree* Projekt entwickelt und eingesetzt. Für die Erstellung der Audiodeskription wurden diese an das Teilprojekt Rehabilitationswissenschaften kommuniziert, damit eine didaktisch passende Audiodeskription erstellt werden konnte.

Art des Videos

Aufnahme einer Videokonferenz mit fünf Schüler*innen einer sechsten Klasse. In der Videokonferenz sprechen die Schüler*innen über ein vorab gehörtes Musikstück. Vier der Schüler*innen sind in einer Videokachel sichtbar; eine Kamera ist ausgeschaltet. Durch die fünf Videokacheln nebeneinander werden fünf parallele Situationen gezeigt. Die Internetverbindung ist nicht stabil, wodurch die Tonqualität der Aufnahme teilweise schlecht ist. Die Namen der Schüler*innen werden gemutet; stattdessen werden die Bezeichnungen S1–S5 verwendet. Zusätzlich werden die Schüler*innen verpixelt.

Durch die Kameraperspektive – frontal am Computer sitzend – sind nicht alle Handlungen im Video zu sehen. So geht S1 beispielsweise an einer Stelle aus dem Bild und spielt auf einem Keyboard, das nicht zu sehen ist.

Die Untertitel wurden als Closed Captions am unteren Bildschirmrand zur Verfügung gestellt. Da es in der Aufgabenstellung um das „Wie“ in der Gesprächsführung geht, wurden parasprachliche Elemente (z.B. Stottern, tiefe Tonlage etc.) mit in die Untertitel aufgenommen.

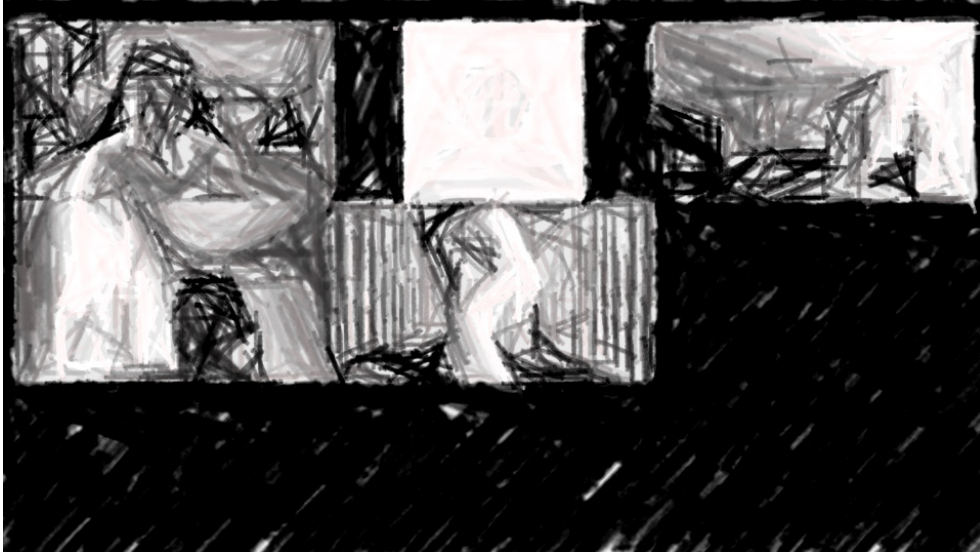


Abbildung 1: Screenshot aus dem Video

Aufgabenstellung

Studierende sollen anhand dieses Videos erfahren, wie Schüler*innen über Musik sprechen. Dafür soll das Schüler*innengespräch im Hinblick auf markante Stellen gesichtet werden, die dann tiefenanalytisch näher betrachtet werden sollen. Es sollen Stellen ausgewählt werden, die einen möglichst breiten und vielfältigen Eindruck bieten, in welchen „Modi“ die Schüler*innen über die gehörte Musik sprechen. Für die Umsetzung dieser Aufgabe soll die Funktion „Annotation“ genutzt werden, um diese Stellen zu markieren und erste Eindrücke festzuhalten (vgl. Langner, 2020).

Audiodeskription

Sowohl der Fokus der Aufgabe als auch das Geschehen im Video selbst sind primär auditiv. Dementsprechend werden nur kurze Beschreibungen eingefügt. Für diese gab es jedoch nicht immer ausreichend Tonpausen, sodass sehr kurze Tonpausen eingefügt wurden, um Platz für die Beschreibungen zu schaffen. Für eine Szenenbeschreibung am Anfang des Videos wird ein Standbild eingefügt.

In der folgenden Tabelle ist ein Ausschnitt des Audiodeskriptionsskripts mit Erläuterungen zu den Beschreibungen dargestellt.

<i>Audiodeskriptionsskript</i>	<i>Erläuterungen</i>
00:00–00:00 (Vorspann) 3 Schülerinnen und 2 Schüler in Videokonferenz. S1, S2, S3, S4 und S5. S5 mit ausgeschalteter Kamera. S1:	Durch die Fokussierung auf die Gesprächssituation anstatt auf die visuellen Handlungen wurde sich gegen eine parallele Situationsbeschreibung der fünf Kamerakacheln entschieden, sondern eine Eingangsbeschreibung inklusive Nennung der Akteur*innen vorgenommen, um einen Eindruck der Gesamtsituation zu vermitteln. Danach wurde der Fokus auf einzelne auffällige Handlungen gelegt.
00:15 S2 nickt.	Wenn möglich, werden die Reaktionen der Schüler*innen beschrieben.
00:18–00:19 S3:	Einordnung der Schüler*innenstimme.
00:30–00:31 S4 hat irritierten Gesichtsausdruck und gestikuliert wirr.	Diese Beschreibungen sind sehr interpretativ, da eine ausführlichere Beschreibung zu viel Zeit in Anspruch genommen hätte. Die Beschreibung „wirr“ soll dabei unterstützen, die später folgenden Gestiken („rhythmisch“) von diesen Beschreibungen zu unterscheiden.
00:34 S1 richtet sich an S4 und gestikuliert schnelle, wirre Bewegungen mit den Händen neben seinem Kopf.	Kommunikationswege (S1 und S4) werden verbalisiert. Handbewegungen werden möglichst kurz und prägnant beschrieben, wobei eine Interpretation in Kauf genommen wird.
01:08–01:09 S1 verlässt Videobild und spielt Keyboard.	Dadurch, dass es fünf parallele Situationen gibt, müssen auch „hörbare“, aber nicht sichtbare Handlungen mit beschrieben werden. So ist das Keyboard nicht erkennbar, aber durch den Weggang von S1 aus der Videokachel, während alle anderen weiterhin in der Kachel sichtbar sind, ist eindeutig, wer das Keyboard spielt, das direkt darauf zu hören ist.
01:37 S1 zeigt auf sich.	Handlungen der Schüler*innen werden, wenn möglich, beschrieben.
01:45 S1 haut rhythmisch auf Tischplatte.	Die Bewegung ist die erste Bewegung, in der S1 eine Bewegung vollführt, die die

<i>Audiodeskriptionsskript</i>	<i>Erläuterungen</i>
	Musik widerspiegelt bzw. diese beschreiben soll. Dementsprechend wurde die Handlung „haut auf die Tischplatte“ noch um die Beschreibung „rhythmisch“ erweitert.
01:57 S2:	Einordnung der Schüler*innenstimme.
02:18 S5:	Einordnung der Schüler*innenstimme.
02:45 S2 nickt.	Handlungen der Schüler*innen werden, wenn möglich, beschrieben, gerade wenn diese anscheinend für das weitere Gespräch relevant sind.
03:24 S1 deutet auf Schreibtisch an, Klavier zu spielen, und schlägt einmal auf Tischplatte.	Kaum Verwendung von musikspezifischem Vokabular, da die Studierenden Begriffe finden sollten, wie sie den „Modus“ benennen, in dem die Schüler*innen über Musik sprechen. Es wurde dementsprechend versucht, auf eine möglichst neutrale Wiedergabe zu achten.
03:38 S1 gestikuliert die gesungenen Töne rhythmisch passend mit den gehobenen Händen.	Hier liegt die Vermutung nahe, dass z.B. das „gestikulierend mit den gehobenen Händen“ auch als dirigierend erkannt wird. Dies als dirigierend zu beschreiben, wäre aber auf der einen Seite interpretativ, auf der anderen Seite würde dies auch wieder gegenläufig zur gestellten Aufgabe sein. Um aber trotzdem zu verdeutlichen, dass die Bewegungen von S1 im Zusammenhang mit den gesungenen Tönen stehen, wurde sich dafür entschieden, diese Bewegung als „rhythmisch passend“ zu beschreiben.

Literatur

Langner, J. (2020). *Beantwortung des Fragerasters zur Videovignette*. Unveröffentlicht.