Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

From creating to translating: Chinese students' audio description processes for film

Weiqing Xiao
Shanghai International Studies University
Ying Li
0234101599@shisu.edu.cn

Shanghai International Studies University

Article

Keywords: direct AD, AD translation, audiovisual translation, Chinese context, media accessibility

Posted Date: April 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9020666/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Audio description (AD), an assistive form of audiovisual translation, is essential for enabling people with visual impairments to participate equally in audiovisual and cultural experiences, yet scripting AD from scratch is costly and time-consuming. AD translation has been proposed as a more economical alternative, but comparative evidence on the production process between direct AD and AD translation remains scarce, especially in the Chinese context. This study therefore conducts a process-level comparison of the two production modes for films. Drawing on keystroke logging, screen recordings, and post-task questionnaires, it investigates participants' phase-level cognitive management (planning, drafting, and revising) and multidimensional perceived task difficulty across the two AD production methods. Findings revealed that overall duration was shorter for AD translation, though not significantly. Planning time was significantly longer in AD translation than in direct AD, whereas drafting and revision showed no significant differences. Questionnaire data indicated significantly higher overall perceived difficulty for direct AD, with only non-significant dimension-level trends. These patterns suggest that AD translation scaffolds script production by drawing on students' familiar interlingual routines, while the additional source-language AD script increases early multimodal integration demands and prolongs planning duration. Cross-linguistic adaptation, tight timing, and multimodal coordination appear to blur dimension-specific effects across the two modes. This study provides empirical evidence on AD production in the Chinese context, moving beyond existing product-oriented and largely conceptual analyses. It offers practical implications for cost-effective AD production and training and demonstrate the benefits of an interdisciplinary perspective that connects audiovisual translation scholarship with media accessibility research.

1. Introduction

Over the past decades, digital technologies have profoundly reshaped the audiovisual media landscape, particularly through the proliferation of video-on-demand platforms and smart devices. This shift has enabled more flexible and ubiquitous access to audiovisual content. However, this technological development has also exacerbated issues of discrimination and marginalization for persons with disabilities, especially those with visual and hearing impairments, who face barriers to fully access the content of audiovisual products (Greco and Jankowska 2020). In response to these challenges, media accessibility has been increasingly recognized as a crucial solution. It concerns the provision of media-based solutions that assist those who cannot, either partially or completely, in accessing media or non-media content, services, or environments in their original form—to perceive and engage with them effectively (Greco and Jankowska 2020).

Media accessibility is closely connected to audiovisual translation (AVT), particularly its assistive modalities such as audio description (AD). AD transforms essential visual content into verbal information, enabling people with visual impairments to participate more equally in audiovisual media and cultural events. Since its inception in the 1940s in Spain (Orero 2008; Reviers 2016), AD has made

significant progress in countries such as the United Kingdom and the United States, whereas in later adopters, such as China, its development has been slower (Author *et al*. 2025).

Regardless of the stage of development, ensuring both the quantity and quality of AD for its end-users remains a challenge due to its high production costs, time-consuming nature and shortage of trained AD professionals (Vera 2006; Remael and Vercauteren 2010; Jankowska 2015). Thus, exploring more efficient ways to produce high-quality AD at scale has become a pressing issue in both practice and research. One proposed solution is AD translation, which involves translating existing AD scripts rather than creating them from scratch. As the wording and content selection have already been established in the source AD (Remael and Vercauteren 2010), this approach may potentially reduce the time and effort required for scriptwriting. However, opinions on its effectiveness remain inconclusive (Hyks 2005; Benecke 2004). Moreover, existing studies have examined AD primarily from a product-oriented perspective, with empirical evidence drawn largely from European contexts. Research on AD translation in the Chinese setting remains scarce. Therefore, this study aims to examine the viability and efficiency of AD translation for films in the Chinese context by drawing on process-level evidence, moving beyond product-based analyses.

2. Direct AD and AD Translation Process

Synthesizing prior work, Jankowska (2015) groups AD production practices into three broad modes, namely individual work (Orero 2012), teamwork (Benecke 2004), and translation and adaptation of existing AD scripts (Vera 2006; Jankowska 2015). Individual production typically requires a single describer to pay sustained attention and single-handed management of multiple constraints and details. Individual work typically requires the describer to devote sustained attention to the task and to manage a wide range of details single-handedly. In team-based production, AD scripts are developed by more than one contributor, and end users, people with visual impairments, are often involved to provide feedback, which can make the final script more responsive to users' needs (Hirvonen 2025). In AD translation, describers work from an existing AD script in another language and then revise and localize it in their native language as needed.

These modes are not mutually exclusive because they reflect different dimensions of AD production. AD scripts can be produced individually or collaboratively, and they can be created from scratch (direct AD) or translated and adapted from an existing script (AD translation). In China, direct AD has been more common in organizational and volunteer settings, whereas in Anglophone and some European contexts AD translation is also practiced, including in AD training contexts.

Although AD can be produced through different methods, these approaches share a broadly similar workflow (Jankowska 2015; Fryer 2016), which can be delineated into three principal stages. Stage one is familiarization and preparation, in which AD script writers are encouraged to listen to the audio soundtracks several times without viewing the film, in order to anticipate audience needs and gauge the material's difficulty (Jankowska 2015) and then they would watch the entire materials again with visual

information to develop a holistic understanding of the material's overall content structure and salient visual cues (e.g., key events, character relations, and settings), and to consult reference sources as needed to verify any domain-specific terminology. In the second stage, AD script writers create the AD script in accordance with established AD guidelines, to decide which visual details to verbalize. This generally entails describing characters, major story developments and spatial-temporal settings. The third stage involves revision and script synchronization, in which AD script writers read aloud to ensure that the description is in accordance with the visual and can be delivered within the available gaps between lines of dialogue. After these three major stages, AD script then is proofread by other AD professionals and persons with visual impairments and then was recorded by voice talents. Figure 1 provides an overview of the AD script workflow described above.

Figure 1 Direct AD script workflow

The difference between direct AD and AD translation is that the source materials are different, in AD translation (See Fig. 2), except for the audiovisual materials, audio describers can also resort to the existing AD scripts.

Figure 2 **AD translation workflow**

3. Literature Review

By transforming visual information into verbal language, AD has been initially and mainly viewed as a form of intersemiotic translation (Jakobson 1959). As the source text of AD translation is itself the outcome of such an intersemiotic transformation, AD translation has been discussed as involving an additional layer of mediation compared to conventional interlingual translation. When the notion of AD translation was first proposed, early research primarily focused on its feasibility and the specific challenges involved. In the early stages of scholarly discussion, divergent views emerged regarding whether AD could be translated. Building on the question raised by Hyks (2005) and Benecke (2004) as to whether AD can be translated into other languages, Vera (2006) conducted an early empirical study involving professional audiovisual translators from different backgrounds, comparing direct AD scriptwriting and AD translation in terms of production time. The findings suggested that AD translation may reduce the time required for AD script production. However, the study also acknowledged that participants' audiovisual translation backgrounds may have influenced the results.

Following Vera's (2006) study, several subsequent studies have empirically compared direct AD and AD translation with respect to production time and associated costs. One notable study by Jankowska (2015) compared direct AD in Polish with the translation of English AD scripts into Polish, involving students majoring in English Studies and Audiovisual Translation as participants. Her findings showed that AD translation not only reduced production time but also led to improvements in overall output quality. More recent research has extended this line of inquiry to the context of video games. Zhang and Mangiron (2025) examined AD production for video games by involving professional practitioners in both

direct AD scriptwriting in Spanish and AD translation from English into Spanish. Their findings corroborated earlier studies, showing reductions in production time and economic costs.

In addition to empirically examining the feasibility of AD translation, another strand of research has addressed this issue mainly from a product-oriented, cross-linguistic perspective by comparing AD scripts across language pairs. This body of research has mainly examined differences in descriptive content and lexical choices, through language-pair comparisons such as Spanish–English (Bourne and Hurtado 2007; Feist and Cifuentes-Férez 2025; Alonso Alonso and Cifuentes-Férez 2025), English–Dutch (Remael and Vercauteren 2010), and Spanish–Chinese (Liu and Tor-Carroggio 2022; Liu et al. 2022). Findings from these comparative studies suggest that such differences foreground a set of recurring challenges in AD translation, notably cultural differences, mismatches between national AD guidelines, and the temporal and synchrony constraints inherent to AD. Among these challenges, cultural differences have received considerable scholarly attention, particularly regarding how culturally specific descriptive norms can be negotiated in AD translation. Several studies have analyzed strategies for translating culture-specific references in foreign films and proposed corresponding strategy frameworks (Szarkowska and Jankowska 2015; Jankowska 2022). In addition, Jankowska, Milc, and Fryer (2017) suggested the use of AD translation from minor languages into English as a means of preserving cultural references. Continuing along the line of contrastive research, Liu and her collaborators conducted corpus-based comparisons of Chinese and Spanish AD, arguing that appropriately translating and adapting Spanish AD scripts could serve as a viable approach to producing Chinese AD scripts (Liu and Tor-Carroggio 2022; Liu, et al. 2022). From a theoretical perspective, drawing on Reiss' text typology (Reiss 1989), Mazur (2024) proposed a textual typology for AD scriptwriting and translation.

With the advancement of technology, there has been growing scholarly interest in how various tools can facilitate AD translation. Matamala and Boix (2016) investigated the use of machine translation (MT) to enhance audiovisual accessibility, especially in multilingual countries, such as Spain where languages like Spanish and Catalan coexist. Their study highlighted the potential of MT but concluded that further development is needed for it to be fully applicable to AD translation. In the same year, Fernández-Torné and Matamala (2016) compared the time and effort in AD direct creation and AD translation. While their quantitative data indicated that post-editing required the least cognitive effort, qualitative feedback from participants revealed a conflicting perception, highlighting the need for further investigation. More recently, with the rise of generative artificial intelligence, Lüthi (2024) explored the use of ChatGPT in AD translation and found that it can be particularly helpful in improving lexical choice and phrasing. Beyond such tools, another study examined the use of templates in AD translation. Templates with embedded time codes were argued to provide an effective starting point, helping to save time and ensure greater consistency (Jankowska 2024).

Another important line of studies has focused on audience reception of AD translation (Liu 2023; Mangiron and Zhang 2025). Liu (2023) compared three versions of AD—a Chinese AD created from scratch, a localized version adapted from Spanish, and a direct translation from Spanish—among end-users with visual impairments. The findings indicated that while participants showed a preference for the

original Chinese AD, both the localized and directly translated versions were nevertheless considered acceptable, indicating that translated AD can meet basic audience expectations. Similar findings have been reported in video game AD, Mangiron and Zhang (2025) found that although AD created from scratch was favored by players with visual impairments, translated AD was likewise regarded as acceptable.

All in all, previous studies point to the growing recognition of AD translation as a viable means to improve media accessibility, with evidence highlighting its potential in areas such as content selection, lexical choices, cost efficiency, and audience acceptance. Yet, despite these encouraging findings, process-oriented research on both direct AD scriptwriting and AD translation remains scarce, and fine-grained comparative evidence between the two tasks is even more limited. Existing studies on the AD process have largely relied on overall time measures and general self-reported difficulty (Jankowska 2015; Fernández-Torné and Matamala 2016; Zhang and Mangiron 2025), an approach that, while informative, does not capture how time is distributed across different stages of direct and translated AD production, nor does it provide a more fine-grained comparison with respect to perceived difficulty across the two tasks. Moreover, although a few studies have examined Spanish-to-Chinese AD translation (Liu 2023), the current body of research remains largely confined to European language pairs, most notably English and Spanish, leaving AD translation between more typologically distant languages —such as English and Chinese—relatively underexplored. Beyond language-related constraints, relatively little attention has been paid to the role of translation students as participants in AD research. Despite calls to integrate both AD translation and direct AD into classrooms (Remael and Vercauteren 2010), translation students who possess the linguistic competences required for cross-cultural AD translation (Jankowska et al. 2017) remain underrepresented in AD process research. In China, AD practice is largely organized through grassroots initiatives, with student volunteers as key contributors (Gao and Cai 2020). Furthermore, most experienced describers and scriptwriters are without an educational background in languages or translation (Tor-Carroggio and Casas-Tost 2020). In light of this, translation students constitute a particularly relevant yet understudied population in the Chinese AD context.

These discussions call for empirical insights into the AD production process, which can be generated by process-oriented methods that trace production as it unfolds. Therefore, the present study focuses on translation students and examines their processes in direct Chinese AD scriptwriting and English to Chinese AD translation. Drawing on multiple sources of process evidence, the study investigates differences in process management and perceived difficulty across the two tasks, with the broader aim of offering empirical insights for both professional practice and early training in AD scriptwriting. Accordingly, the study seeks to address the following two research questions:

(1) How do the process stages differ between direct AD in Chinese and English-to-Chinese AD translation among translation students?

(2) How do translation students perceive and compare the task difficulty of direct Chinese AD and English to Chinese AD translation?

4. Methodology

4.1 Research design

Hayes and Flower's model conceptualized the writing process as three phases, namely planning, text generation (termed "translating" in their model), and revising (Hayes and Flower 1980; Flower and Hayes 1981). This staged view of writing resonates with Jakobsen's (2002) model of translation process, which operationalizes translation into three phases—an initial orientation phase, a drafting phase, and a final revision phase—based on patterns observed in logged keystroke data. Building on insights from writing and translation process research, Jankowska (2021) examined AD scriptwriting process through think-aloud protocols (TAPs). Her findings suggest that the AD scriptwriting process shares features with both writing and translation and can also be described in three main stages: pre-drafting, drafting, and post-drafting. Building on established staged approaches in writing, translation, and AD process research, this study also adopts a three-phase framework—planning, drafting, and revision—as an analytical lens to examine the AD production process in direct AD scriptwriting and AD translation (see Fig. 3 and Fig. 4).

Figure 3A **three-phase framework of direct AD**

Figure 4A **three-phase framework of AD translation**

In this study, a within-subjects repeated measures design was adopted: each participant completed two tasks, direct AD scriptwriting and AD translation. Task order was counterbalanced to control for order effects. The independent variable was therefore task type. Direct AD scriptwriting was defined as the production of a Chinese AD script from scratch using the video clip as the sole source of input. AD translation referred to the production of a Chinese AD script based on both the video clip and a pre-existing English AD script. The dependent variables comprised both process-related measures and self-reported ratings of perceived task difficulty. The process-related measures included overall task duration as well as time allocated to the three phases of AD scriptwriting, namely planning, drafting, and revision. These measures were derived from keystroke logging data and were operationalized as the total time (in seconds) spent in each phase based on Jakobsen's (2002) tripartite model. Perceived task difficulty was measured through a post-task questionnaire. Two levels of perceived difficulty were examined: overall task difficulty and difficulty across seven specific dimensions. These dimensions included content selection, diction, scene and object description, action description, style consistency, condensation, and temporal synchronization. More specifically, content selection referred to the perceived difficulty of identifying and selecting key visual information for AD; diction captured the difficulty of choosing appropriate words and expressions; scene and object description, as well as action description, reflected the difficulty of verbally rendering scenes, objects and actions in the video clip; style consistency denoted the difficulty of maintaining stylistic consistency between the AD script and the video; condensation represented the difficulty of compressing AD wording to fit the limited time available in the video; and temporal synchronization referred to the difficulty of aligning the AD script with the unfolding visual elements on screen.

To examine students' process management and perceived task difficulty, a quantitative process-based approach was adopted, drawing on keystroke logging, screen recording, and post-task questionnaires. Keystroke logging data were collected through Inputlog 7 (Leijten and Van Waes 2013) to capture participants' text-production behaviors across the three AD scriptwriting phases. The software unobtrusively records pauses, deletions, insertions and mouse movements. It also enabled the presentation of multimodal source texts, such as video clip in direct AD or combined video clip and AD script in AD translation task, making it possible to trace how students navigated and responded to different input types during the AD tasks. To complement the keystroke data, the EyeVeri (EV) screen recording function was enabled to synchronously record on-screen behaviors, including video playback, online searches, and task switching, providing valuable temporal and behavioral context for a more comprehensive and nuanced interpretation of students' process management. In addition, after each task, participants completed a self-report questionnaire translated from and drawn on the task-difficulty items in Bausells-Espín (2022). This questionnaire was originally developed for students performing didactic AD tasks. It assessed perceived task difficulty across multiple dimensions in both direct AD and AD translation and was thus suitable for the present study. The abridged questionnaire comprised ten items. Item 1 recorded participants' IDs. Item 2 measured their perceived general difficulty for each task. Items 3–9 assessed their perceived difficulty across seven dimensions: content selection, diction, scene and object description, action description, style consistency, condensation, and temporal synchronization (Bausells-Espín 2022). Item 10 assessed perceived usefulness of the English AD script. Items 1–9 were administrated after each task, whereas Item 10 was completed only after both tasks. Ratings were collected on a 5-point Likert-type scale (1 = not difficult at all, 5 = extremely difficult), with higher scores indicating greater perceived difficulty.

4.2 Participants

The participants in this study consisted of undergraduate students from two translation classes at a Chinese university in central China, representing common training backgrounds of translation students in China. Third-year translation students were selected because they had completed the core bilingual and translation courses needed for the AD-translation task, ensuring a basic curricular threshold rather than a presumed proficiency advantage. As participants had no previous experience with AD, they received an introductory instruction on AD before undertaking the tasks. The instruction was embedded in a university interpreting course and delivered as a three-week module (two sessions per week, one hour per session; approximately six hours in total). It aimed to familiarize participants with AD practice and task requirements, rather than to train them as professional audio describers. The study thus focused on novice learners' AD processes in both direct AD and AD translation. Participants ranged in age from 19 to 24 years. All participants were fully informed about the study and signed informed consent forms. They were also informed that their participation was voluntary and that they could withdraw from the study at any time without any consequences. A total of 83 students initially participated in the study. In total, 52 participants successfully completed both the AD tasks and the accompanying questionnaires. Usable Inputlog data were obtained from 22 participants, as technical issues prevented keystroke data capture for the remaining participants.

4.3 Materials

Research materials were taken from the first two films in the *Ice Age* series, *Ice Age* (2002) and *Ice Age: The Meltdown* (2006). This movie series was chosen for several reasons. First, it consists of multiple installments, providing consistent animation style and character design across films, which facilitates the extraction of comparable stimulus materials. Second, professionally produced English AD is available for these films. The AD audio was obtained from AudioVault (https://audiovault.net/), a non-commercial repository that provides access to audio-described soundtracks for research and accessibility purposes and was transcribed into written scripts for use as source texts in the AD-translation task. Third, the selected scenes are dialogue-free, which helped ensure comparable insertion opportunities for audio descriptions across tasks (Jankowska 2024) and reduced the need for participants to manage timing decisions around dialogue gaps, allowing them to focus on selecting and describing visual information. The film *Ice Age: The Meltdown (2006)* has also been used as research material involving student scriptwriters in one previous study (Jankowska 2015).

Two video clips, each approximately 95 seconds in length duration, were extracted from the two films. The two clips were matched for duration, background setting, character composition, descriptive density descriptive density, operationalized as AD word count (Tor-Carrogio 2020) and both depict action-oriented narrative sequences (see Table 1). Two experienced AD script writers independently rated the two video clips for overall descriptive difficulty on a 5-point scale (1 = very easy, 5 = very difficult). Both clips received identical difficulty ratings (2/5) from both raters, suggesting that they were comparable in overall descriptive difficulty.

Table 1
Video clips for AD tasks

Clip1	**Duration**	**Background setting**	**Character composition**	**Descriptive density**	**Content**
	100s	**On an ice shelf with an overhang**	**Single character (Scrat) only**	**246 words**	**Scrat struggles to extract an acorn as water jets burst through openings in the ice.**
Clip2	100s	On an ice sheet over water	Single character (Scrat) only	259 words	Scrat pursues an acorn underwater as thin ice fractures into floating blocks.

In addition, the two corresponding English AD scripts were comparable in lexical and syntactic complexity, as indexed by text length, lexical density (Read 2000), core vocabulary coverage (Lv and Liang 2019; Jiang *et al*. 2024), mean sentence length, moving average type-token ratio (50-word window; Jiang et al. 2025), and mean dependency distance (Liu 2008; Jiang et al. 2025). Details are provided in Table 2.

Table 2
Lexical and syntactic complexity measures for the two English AD scripts

Index	Clip1	Clip2
TL: text length (words)	246	259
LD: lexical density	0.520	0.552
CVC: core vocabulary coverage	0.467	0.481
MSL: mean sentence length	13.7	13.5
MATTR: moving average type-token ratio	0.948	0.912
MDD: mean dependency distance	2.235	2.428

Two experienced translators also rated the English AD scripts for overall text difficulty on a 5-point scale (1 = very easy, 5 = very difficult). Both raters assigned a score of 3 to each script, suggesting comparable difficulty across the two scripts. Both video clips were used in the direct AD scriptwriting and the AD translation tasks. In the direct-AD condition, the video clip served as the sole source input, and participants received a time-coded blank template to guide segmentation without providing any linguistic content. In the AD-translation condition, participants were provided with the video clip together with a time-coded English AD script. Providing access to both sources is recommended in AD-translation practice, as AD script writer may consult the original video when translating and adapting the AD script (Remael and Vercauteren 2010). The use of time-coded templates in both conditions served to facilitate temporal alignment and ensured that participants' outputs followed comparable segmentation for later analysis (Jankowska 2024).

4.4 Procedure

The study spanned eight weeks and comprised three sequential stages: pre-experimental preparation, software familiarization, and task administration. During Weeks 1–3, participants attended two weekly in-person sessions (1 hour each; about 6 hours in total) introducing key aspects of AD, including its definition, major categories of practice, and basic production methods. The sessions covered both direct AD scriptwriting and AD translation so that students understood how each task should be carried out. Instructional materials were partly drawn from the open-access course materials developed within ADLAB PRO, a European Union–funded Erasmus+ project (2016–2019). Because these materials are primarily designed for English AD training, supplementary Chinese AD guidelines were drawn from unpublished training materials developed by Sound of Light Barrier-free Film and TV Culture Development Centre, a private non-enterprise unit financially supported by the Shanghai Disabled Persons' Federation, which produces audio described films and disseminates them to residential communities via digital distribution channels. These supplementary AD guidelines were used to familiarize participants with conventions for Chinese AD.

In Week 4, participants received hands-on training in the use of Inputlog 7 and EV Capture and completed a short pilot task to gain familiarity with the software and to verify technical functionality. The pilot involved a short, unrelated video clip used solely for operational familiarization and not for data analysis.

During Weeks 5–8, each participant completed two AD tasks: one AD translation task and one direct AD scriptwriting task. Task order was counterbalanced to reduce potential order effects. Participants worked individually on their own laptops in a quiet and undisturbed environment of their choices. They were permitted to consult online resources and artificial intelligence (AI) tools during AD scriptwriting, and to complete the task without a time limit but as efficiently as possible, to approximate a naturalistic context.

Participants were instructed to activate the EyeVeri (EV) screen-recording function and Inputlog before starting each writing task to capture the script-production process. The desktop interface was organized into three areas: the left side was used for video playback and external resource searches, while the right side served as the writing workspace. In the direct AD condition, the writing workspace displayed only the timeline (see Fig. 5). In the AD translation condition, the writing workspace additionally presented an English AD script as a reference for translation (see Fig. 6).

Figure 5 **Screen setting in the direct AD condition**

Figure 6 **Screen setting in the AD translation condition**

After each task, participants completed a post-task questionnaire reporting their perceived level of task difficulty. All data were collected and analyzed in anonymized form.

4.5 Data Analysis

This study analyzed data from keystroke logs, screen recordings, and post-task questionnaires to examine Chinese students' AD scriptwriting processes and their perceived task difficulty.

Keystroke logs were analyzed through Inputlog's General Analysis module (Leijten and Van Waes 2013) to generate time-stamped traces of AD scriptwriting behavior, which were exported to Excel for further inspection and comparison.

Keystroke data were analyzed together with screen recordings to segment the AD scriptwriting process into three phases following Jakobsen's (2002) tripartite model. The planning phase was defined as the interval between the presentation of the visual source and the first productive keystroke, during which participants inspected the screen and prepared for AD script production. The drafting phase extended from the first productive keystroke to the entry of the final sentence-ending punctuation, reflecting continuous AD script generation. The revision phase covered the period from the completion of the initial draft to task submission.

Screen recording captured participants' on-screen activities, such as drafting, revising, and consulting external resources. It offers researchers a unique window into participants' thought processes and decision-making during AD tasks (Raído et al. 2025). Screen-recording data were analyzed alongside the keystroke logs by temporally aligning the recordings with typing traces to validate participants' keyboard-related actions and to support phase segmentation.

Two trained coders independently annotated phase boundaries using keystroke logs and screen recordings, and disagreements were resolved through discussion. Descriptive statistics (means, standard deviations, and medians) were computed for phase-level time allocation in each task. Given the within-subject design and the non-normal distribution of the data, paired comparisons between the two AD tasks were conducted using Wilcoxon Signed-Rank Tests in R. Data from post-task questionnaires containing five-point Likert-type items were summarized descriptively, and paired differences between the two AD tasks were likewise compared using Wilcoxon Signed-Rank Tests in R.

5. Results

5.1 Process Management

Descriptive statistics are reported for total task time and the temporal distribution across planning, drafting, and revision in the two AD tasks. Participants spent more total time on the direct AD task (M = 1899.97 s, SD = 1177.94, Median = 1555.54) than on the AD translation task (M = 1727.98 s, SD = 1000.63, Median = 1568.75). Across both tasks, the largest proportion of time was allocated to drafting, followed by revision and then planning. Specifically, Planning time was longer in the AD translation task (M = 51.74 s, SD = 55.98, Median = 41.28) than in the direct AD task (M = 24.68 s, SD = 12.88, Median = 24.06). During the drafting stage, participants spent more time in the direct AD task (M = 1805.48 s, SD = 1104.31, Median = 1516.69) than in the AD translation task (M = 1583.96 s, SD = 845.50, Median = 1477.15). In the revision stage, participants also spent more time in the AD translation task (M = 92.29 s, SD = 228.85, Median = 17.42) than in the direct AD task (M = 69.81 s, SD = 127.45, Median = 16.02). Time allocation across the three phases is summarized in Table 3.

As Shapiro–Wilk tests indicated non-normal distributions in several conditions (e.g., $W = 0.706$, $p < .001$ for planning in AD translation), Wilcoxon signed-rank tests were conducted to compare participants' total task time and phase-level time allocation between the two tasks. For total task time, no significant difference was observed between the two tasks ($V = 94$, $p = .305$, $r = .22$). Concerning the three specific stages, a significant difference was found in the planning stage ($V = 202$, $p = .013$), with higher planning time in the AD translation task than in the direct AD task. The corresponding effect size was $r = .52$. No significant differences were found for drafting or revision (drafting: $V = 87$, $p = .210$, $r = .27$; revision: $V = 92$, $p = .424$, $r = .17$). These results are summarized in Table 4.

Table 3
Descriptive statistics for phase-level time allocation

Stage	Task Type	Mean (s)	SD (s)	Min (s)	Max (s)	Median (s)
Planning	AD Translation	51.74	55.98	1.67	265.72	41.28
	Direct AD	24.68	12.88	4.80	48.89	24.06
Drafting	AD Translation	1583.96	845.50	0	3994.30	1477.15
	Direct AD	1805.48	1104.31	814.80	5543.46	1516.69
Revision	AD Translation	92.29	228.85	0	1073.58	17.42
	Direct AD	69.81	127.45	0	529.56	16.02
Total time	AD Translation	1727.98	1000.63	273.11	4268.56	1568.75
	Direct AD	1899.97	1177.94	882.83	5778.91	1555.54

Note. All durations are measured in seconds. n = 22 (keystroke data available).

Table 4
Wilcoxon signed-rank test results for phase-level time allocation

Stage	Wilcoxon V	*p* value	Effect size *r*	Significance
Planning	202	.013	**.52**	*
Drafting	87	.210	.27	
Revision	92	.424	.17	
Toal time	94	.305	.22	

Note. All durations are measured in seconds. Wilcoxon signed-rank tests were used for pairwise comparisons between AD Translation and Direct AD. Effect sizes (r) were calculated using the formula r = Z/√n, where n = 22. Asterisks indicate significance: *p* < .05 (*), *p* < .01 (**), p < .001 (***).

5.2 Perceived Difficulty for the Two AD Tasks

This section reports descriptive and inferential statistics for perceived-difficulty ratings across the two AD tasks. For the general-difficulty item, ratings were higher for the direct AD task (M = 3.44, SD = 0.70, Median = 3.5) than for the AD translation task (M = 3.10, SD = 0.85, Median = 3.0). More fine-grained descriptive statistics across individual difficulty dimensions showed a consistent numeric pattern, with consistently higher means for the direct AD task than for the AD translation task. For instance, content-selection ratings were 3.40 (SD = 0.80, Median = 3.0) for direct AD and 3.23 (SD = 0.83, Median = 3.0) for AD translation; lexical-choice ratings were 3.63 (SD = 0.71, Median = 4.0) and 3.54 (SD = 0.73, Median = 4.0), respectively. All descriptive statistics for perceived-difficulty items are summarized in Table 5.

Table 5
Descriptive statistics for participants' perceived task difficulty

Dimension	Task Type	Mean (s)	SD (s)	Median (s)
General difficulty	AD Translation	3.10	0.85	3.0
	Direct AD	3.44	0.70	3.5
Content selection	AD Translation	3.23	0.83	3.0
	Direct AD	3.40	0.80	3.0
Diction	AD Translation	3.54	0.73	4.0
	Direct AD	3.63	0.71	4.0
Scene and object description	AD Translation	3.44	0.75	3.0
	Direct AD	3.44	0.80	3.5
Action description	AD Translation	3.54	0.83	4.0
	Direct AD	3.63	0.77	4.0
Style consistency	AD Translation	3.60	0.75	4.0
	Direct AD	3.62	0.77	4.0
Condensation	AD Translation	3.42	0.75	3.0
	Direct AD	3.52	0.64	4.0
Temporal synchronization	AD Translation	3.25	0.86	3.0
	Direct AD	3.37	0.79	3.0

Note. M = mean; SD = standard deviation. Ratings were collected on a 5-point Likert scale (1 = very easy, 5 = very difficult). Median refers to the 50th percentile. All values are rounded to two decimal places. Results are based on self-reported ratings from N = 52 participants.

To test for differences in perceived difficulty between the two tasks, Wilcoxon signed-rank tests were conducted for the general-difficulty item and each individual dimension. A significant difference was found for the general-difficulty item ($V = 288.5$, $p = .0123$, $r = .35$), with higher ratings for the direct AD task. However, no statistically significant differences were observed for any of the remaining dimensions (all p-values > .28). Full inferential results are presented in Table 6.

Table 6
Wilcoxon signed-rank tests for participants' perceived task difficulty

Dimension	Wilcoxon V	*p* value	Effect size (r)	Significance
General difficulty	288.5	0.0123	0.347	*
Content selection	319.5	0.2807	0.15	
Diction	270.0	0.3905	0.119	
Scene and object description	129.0	0.9452	0.01	
Action description	282.5	0.4795	0.098	
Style consistency	215.0	0.7806	0.039	
Condensation	210.0	0.3509	0.129	
Temporal synchronization	222.0	0.4187	0.112	

Note. *V* = Wilcoxon signed-rank statistic; *p* = two-tailed significance value. *r* = effect size, calculated as $r = Z/\sqrt{N}$, where $N = 52$ (number of paired observations). Ratings were based on a 5-point Likert scale (1 = very easy, 5 = very difficult). Asterisks indicate significance: $p < .05$ (*), $p < .01$ (**), $p < .001$ (***).

After completing both AD tasks, participants rated the perceived usefulness of the English AD script. As indicated by participants' responses, 50 out of 58 participants (86.2%) agreed that having access to the English AD script made the AD scriptwriting process easier, whereas 8 (13.8%) disagreed.

6. Discussion

6.1 Process Management across Two AD Tasks

Addressing the first research question, the study compared participants' scriptwriting process in direct AD and AD translation by examining their overall temporal demands and time allocation across planning, drafting and revision.

Total task duration tended to be shorter for AD translation than for direct AD, but the difference was not statistically significant, indicating broadly comparable temporal demands between the two approaches. This finding echoes previous studies involving professional AD script writers as participants, which also did not identify significant differences in total duration between the two AD scriptwriting methods (Vera 2006; Liu 2023; Zhang and Mangiron 2025). One tentative explanation concerns participants' training background. After three years of translation learning, the AD translation task may have felt more familiar to them, which could have reduced the temporal demands to some extent. In addition, in the AD translation condition, two time-consuming steps—selecting relevant visual content and formulating wording—had already been completed in the source AD (Remael and Vercauteren 2010), which may

partly account for its shorter total duration. This interpretation, however, should be treated with caution given the null statistical result.

When the AD scriptwriting process is examined across the three phases, the results indicate largely parallel temporal distributions for the two task types. The drafting phase tends to occupy the largest proportion of time, whereas the planning phase accounts for the least. This pattern has also been documented in writing and translation process research (e.g., Jakobsen 2002; Zheng and Fan 2021). Although the overall three-phase pattern was similar, slight differences emerged in how time was allocated across phases in direct AD and AD translation. More fine-grained inspection of phase-level allocation showed that the drafting phase tended to take longer in direct AD than in AD translation, whereas revision showed slightly higher temporal durations in the translation condition. The planning phase also required more time in AD translation than in direct AD, and this difference reached statistical significance. Overall, the planning stage was the only phase in which time allocation differed significantly between direct AD and AD translation. The longer planning time in AD translation may be interpreted through the redundancy principle in multimedia learning (Sweller 2005). Given individuals' limited processing resources (Sweller 1988), presenting essentially the same information in more than one representational form requires participants to mentally coordinate these sources. Such coordination increases cognitive demand and may slow early conceptual decision-making and planning. In the AD translation condition, the written source AD script constitutes an additional stream of input that must be coordinated with the visual content and relevant audio cues in the video clip, potentially intensifying resource competition during planning and delaying the onset of fluent drafting. Although drafting tended to take longer in direct AD and revision was slightly longer in AD translation, these differences were small and did not reach statistical significance, suggesting broadly comparable allocation patterns for these phases under the present materials and participant profile. The directional trends are nonetheless interpretable but should be taken with caution. The tendency in drafting time may be because the source AD script in AD translation provides a starting point for content and wording. Revision was slightly longer in AD translation, in line with evidence that professional translators often revise more in translation than in monolingual writing (Immonen 2006). This may stem from cross-linguistic and cultural adaptation demands (Vercauteren et al. 2021) and revision-oriented habits shaped by translation training (Immonen 2006).

6.2 Perceived Difficulty for the Two AD Tasks

Turning to the second research question, post-task questionnaire responses showed that AD translation was rated as easier than direct AD, and this difference was statistically significant. The corresponding effect size ($r = 0.347$) was moderate (Cohen 1988), pointing to a substantive difference in perceived difficulty between the two tasks. A large proportion of participants also reported that the English AD script facilitated their scriptwriting process. This echoes findings from video game AD translation, where participants articulated that translating from an existing AD script as less cognitively demanding, because it provides guidance on what to describe and, to some extent, how to formulate it (Zhang and Mangiron 2025). This pattern can also be interpreted through the lens of scaffolding (Wood et al. 1976),

in that external support can make demanding tasks more manageable by offering guidance and reducing uncertainty. In AD translation, the English AD script provides participants with an external reference, which may make the task more akin to the translation work they typically undertake and thus easier overall. The more fine-grained comparison suggested that AD translation tended to be viewed as easier than direct AD in diction, action description, stylistic consistency, content selection, information condensation, and temporal synchronization. However, none of these differences reached statistical significance, possibly due to the inherently demanding nature of AD scriptwriting and partly because the advantages of AD translation were offset by the materials employed in this study.

As a multifaceted cognitive-linguistic and intermodal mediation activity (Braun 2007), AD scriptwriting involves two interrelated processes: the reception of the original audiovisual text, which requires AD script writers to comprehend the language, images, and sounds, and to interpret their interplay; and the production of the AD script through selecting essential information and integrating it into a coherent whole (Holsanova 2020). In both AD translation and direct AD, AD script writers are faced with the challenge of transforming multimodal information that is originally presented simultaneously on the screen into a linear format in spoken language, within a limited timeframe (Yan and Luo 2023; Vercauteren 2012; Remael and Vercauteren 2010).

Moreover, the characteristics of the research materials are an important factor in interpreting the results. The film clips selected in this study are narrative in nature with relatively straightforward plot development, providing a clear temporal–causal progression that facilitates AD content selection. In other words, the clips themselves may have been sufficiently clear to indicate what to describe, and the presence of the English AD scripts in AD translation may therefore not have made a noticeable difference in students' perceived content-selection difficulty. As narrative processing is widely regarded as a fundamental mode of human cognition, supported by both theoretical and neurocognitive accounts (Bruner 1991; Young and Saver 2001; Barthes 2004). Such processing enables individuals to track how events unfold in time and assemble them into a coherent whole. The materials also incorporate rapid pacing, exaggerated physical movements, vivid facial expressions, and synchronized sound cues that together contribute to an intense and humorous style. What is presented on the screen is no less important than how it is presented (Bardini 2017). As such, preserving this style in both AD translation and direct AD requires careful consideration and deliberate decision-making.

Additionally, the linguistic and cultural differences between Chinese and English constitute another possible cause. While the translation of English AD into Chinese itself is not considered particularly difficult, thanks to the fact that the language of AD is primarily easy to be understood (Arias-Badia and Matamala 2023), adapting it to align with Chinese AD conventions nevertheless presents new challenges. This is partly because, in English AD, actions are typically expressed with concise and specific manner verbs (Molés-Cases and Alonso 2025), which not only convey the action itself but also the way it is performed, whereas in Chinese, equivalent verbs may not exist. Moreover, due to the time constraints, translating from English AD often requires either content adjustments or condensing to fit within the available temporal duration, further complicating the task. In this study, participants worked

with a time-coded template, which, as noted by Jankowska (2024), can ease the burden of temporal synchronization by pre-allocating descriptive slots. This may account for why temporal synchronization was rated among the less difficult dimensions in both tasks.

The findings in this study are tentative and closely tied to the materials and task design adopted. As such, the present findings should be interpreted with caution when considering generalization beyond the specific context of this study.

7. Conclusion

This study presents empirical evidence on the AD processes and perceived task difficulty of translation students in direct AD and AD translation. Drawing on data from keyboard logging, screen recordings, and questionnaires, it constitutes, to the best of our knowledge, among the first empirical comparisons of the two AD script production modes in the Chinese context. By incorporating translation students as novice describers, it also provides insights into the early-stage development of AD scriptwriting training.

Overall, the results indicate that, consistent with patterns observed in other text-production tasks, both direct AD and AD translation allocate comparatively less time to planning and revision. Most of the task time is concentrated in the drafting phase. Overall task duration tended to be shorter for AD translation than for direct AD scriptwriting, suggesting that AD translation may be less time-consuming. This pattern is consistent with the questionnaire data, which indicates that students perceived translating from an English AD script as significantly easier. Specifically, at the phase level, planning time was significantly longer in AD translation than in direct AD, whereas revision showed only a slight, non-significant increase in the translation condition. This pattern may be interpreted through the lens of the redundancy principle in multimedia learning (Sweller 2005) and participants' educational background in translation. In the drafting phase, direct AD scriptwriting is slightly longer than Ad translation, though not significantly. Subtask-level perceived difficulty was likewise slightly higher for direct AD scriptwriting.

The absence of significant between-condition differences at the subtask level may be attributable to the characteristics of the selected materials. The video clips included in this study were primarily narrative in nature, featuring relatively clear temporal structures and event progressions. Under such conditions, content selection decisions may be sufficiently supported by visual input alone, such that the availability of the English AD scripts in AD translation does not make a significant difference in participants' perceived difficulty of content selection. Furthermore, both AD translation and direct AD appear to place substantial and largely overlapping cognitive demands on participants. Regardless of task type, participants are required to engage in continuous decision-making, maintain focused attention, and process multimodal information. At the subtask level, these shared demands may have resulted in comparable perceived difficulty across conditions. From this perspective, the absence of significant differences may not indicate that the two tasks are cognitively equivalent, but rather that their similarly high and overlapping processing demands may have masked more fine-grained distinctions in the

current experimental setting. Future work could therefore adopt more sensitive, process-oriented measures, such as eye-tracking, to capture subtle differences in AD script production process.

The findings suggest that AD translation may function as a pragmatic alternative to direct AD in the Chinese context when an English AD script is available, and time and staffing resources are constrained. However, the present process evidence indicates that AD translation should not be treated as a guaranteed shortcut, as its overall duration was only slightly shorter, while its planning and revision time tended to be longer. Moreover, the effects of AD script translation may be material dependent. For predominantly narrative clips with a clear temporal organization, the availability of an English AD script does not appear to significantly affect participants' perceived difficulty in content selection. Pedagogically, students' lower perceived difficulty suggest that AD script translation can serve as an entry point to AD script production, providing additional scaffolding for beginners. A staged progression from AD translation to independent AD production may be beneficial, as AD translation offers students concrete instances of AD practice and helps establish a foundational understanding of the task before they engage in fully independent production. However, AD translation training should place particular emphasis on the challenges posed by cross-linguistic adaptation, AD conventions, timing constraints, and clip characteristics. In addition, given the longer planning phase observed in AD translation, instruction should underscore early-stage decision-making, guiding students to integrate multimodal cues, prioritize relevant information, and engage with the source AD script effectively.

Nevertheless, this study is not without limitations. First, the participant selection was shaped by the specific context of AD practice in China. As professional AD scriptwriters are relatively limited and are not typically trained in English, student participants with relevant language backgrounds were recruited for the present study. While this choice is methodologically justified, it may limit the direct generalizability of the findings to professional AD practice. Second, technical constraints affected the scope of the process data collected. Due to compatibility issues with the Inputlog software on many participants' computers, complete keyboard logging data could only be obtained from a subset of participants. This limitation resulted in a reduced dataset for process measures. Future research with larger and more diverse participant pools would help to further validate the findings. Finally, the materials used in this study were primarily narrative in nature and dialogue-free, while the audiovisual materials might be more complicated in rea life. Caution is therefore warranted when generalizing the results to other types of audiovisual materials, such as informational or dialogue-heavy content.

Declarations

Acknowledgement

The research of this academic paper is supported by the China National Social Science Major Research Project "General History of Film Translation in China" (Grant number: 20&ZD313)

Data availability statements

The Inputlog and questionnaire data used in the analyses are publicly available and can be found in the Supplementary Materials section.

Competing interests

The authors declare no competing interests.

Ethical statements

Ethical approval was obtained from the Ethics Committee of the Key Laboratory of Brain–Machine Intelligence for Information Behavior–Ministry of Education, Shanghai International Studies University (Approval number: 202509100007) on 15 September 2025. All procedures in the study were conducted in accordance with the institutional ethical guidelines, safeguarding participants' autonomy, well-being, privacy, and rights.

Informed Consent: All participants provided written informed consent before taking part in the study. The consent process was administered by the researcher (the corresponding author) prior to the study sessions and before any data were collected. Written informed consent was obtained on 1 October 2025 from Chinese university student translators who met the study's eligibility criteria.

Author contributions.

Author A supervised the project, recruited the students, reviewed the study materials and experimental procedures, and contributed to writing through critical review and editing. Author B prepared the study materials, conducted the experiment, analyzed the data, drafted the manuscript and prepared figures and tables. Both authors reviewed the manuscript. Both authors approved the final version for submission.

References

1. Alonso Alonso R, Cifuentes-Férez P (2025) Boundary-crossing events in audio descriptions across English and Spanish. Across Languages and Cultures 26(2):189–204.
2. Arias-Badia B, Matamala A (2023) Audio description from an easy-to-understand language perspective: A corpus-based study in Catalan. The Journal of Specialised Translation 40:268–292.
3. Bardini F (2017) Audio description style and the film experience of blind spectators: Design of a reception study. Rivista Internazionale di Tecnica della Traduzione/International Journal of Translation 19(1):49–73.
4. Barthes R (2004) Introduction to the structural analysis of narratives. In: Bal M (ed) Narrative theory: Critical concepts in literary and cultural studies. Routledge, London and New York, pp 65–116.
5. Bausells-Espín A (2022) Audio description as a pedagogical tool in the foreign language classroom: An analysis of student perceptions of difficulty, usefulness and learning progress. Journal of Audiovisual Translation 5(2):152–175.
6. Benecke B (2004) Audio-description. Meta: Journal des traducteurs 49(1):78–80.

7. Bourne J, Hurtado CJ (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: Díaz-Cintas J, Orero P, Remael A (eds) Media for All: Subtitling for the Deaf, Audio Description and Sign Language. Brill, Leiden, pp 175–188.
8. Braun S (2007) Audio description from a discourse perspective: A socially relevant framework for research and training. Linguistica Antverpiensia, New Series–Themes in Translation Studies 6:357–369.
9. Bruner J (1991) The narrative construction of reality. Critical Inquiry 18(1):1–21.
10. Cohen J (1988) Statistical power analysis for the behavioral sciences. Lawrence Erlbaum, Hillsdale, NJ.
11. Feist MI, Cifuentes-Férez P (2025) Thinking-for-audio-describing: An English-Spanish corpus-based study and some didactic applications. Perspectives 1–20.
12. Fernández-Torné A, Matamala A (2016) Machine translation in audio description? Comparing creation, translation and post-editing efforts. SKASE Journal of Translation and Interpretation 9(1):64–87.
13. Flower L, Hayes JR (1981) A cognitive process theory of writing. College Composition & Communication 32(4):365–387.
14. Greco GM, Jankowska A (2020) Media accessibility within and beyond audiovisual translation. In: Bogucki Ł, Deckert M (eds) The Palgrave handbook of audiovisual translation and media accessibility. Palgrave Macmillan, Cham, Switzerland, pp 57–81.
15. Gao XH, Cai Y (2020) [illegible]: [illegible] [Free from barriers: An analysis of the current situation of China's barrier-free film industry]. [illegible] (1):19–30.
16. Hayes JR, Flower LS (1980) Writing as problem solving. Visible language, 14(4): 388-399.
17. Holsanova J (2020) Uncovering scientific and multimodal literacy through audio description. Journal of Visual Literacy 39(3–4):132–148.
18. Hyks V (2005) Audio description and translation: Two related but different skills. Translating Today 4:6.
19. Hirvonen M (2025) Shared cognition in the translation process: Information processing and meaning production as interactive accomplishments. Translation Studies 18(1):61–82.
20. Immonen S (2006) Translation as a writing process: Pauses in translation versus monolingual text production. Target: International Journal of Translation Studies 18(2):313–336.
21. Jakobson R (1959) On linguistic aspects of translation. In: Brower RA (ed) On translation. Harvard University Press, Cambridge, MA, pp 232–239.
22. Jakobsen AL (2002) Translation drafting by professional translators and by translation students. Copenhagen Studies in Language 27:191–204.
23. Jankowska A (2015) Translating audio description scripts: Translation as a new strategy of creating audio description. Peter Lang, Bern.

24. Jankowska A (2021) Audio describing films: A first look into the description process. The Journal of Specialised Translation 36:26–52.
25. Jankowska A (2022) Audio description and culture-specific elements. In: Taylor C, Perego E (eds) The Routledge handbook of audio description. Routledge, London and New York, pp 107–126.
26. Jankowska A (2024) Guidance but not the ultimate commandment: Pros and cons of using pivot templates in the audio description production workflow. The Journal of Specialised Translation 42:155–173.
27. Jankowska A, Milc M, Fryer L (2017) Translating audio description scripts into English. SKASE Journal of Translation and Interpretation 10(2):2–16.
28. Jiang X, Jiang Y, Zhang X (2025) Assessing effects of source text complexity on L2 learners' interpreting performance: A dependency-based approach. International Review of Applied Linguistics in Language Teaching 63(4):2717–2746.
29. Leijten M, Van Waes L (2013) Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. Written Communication 30(3):358–392.
30. Liu H (2008) Dependency distance as a metric of language comprehension difficulty. Journal of Cognitive Science 9(2):159–191.
31. Liu Y (2023) Audio description translation from Spanish into Chinese as an alternative to creating audio description: A reception study. Parallèles 2023:119–138.
32. Liu Y, Tor-Carroggio I (2022) Audio description translation: A pilot study in Chinese/Spanish. In: Taylor C, Perego E (eds) The Routledge handbook of audio description. Routledge, London and New York, pp 460–475.
33. Liu Y, Casas-Tost H, Rovira-Esteva S (2022) Made in China versus Made in Spain: A corpus-based study comparing AD in Chinese and Spanish. The Journal of Specialised Translation 38:180–207.
34. Lüthi N (2024) KI-gestützte Übersetzung von Audiodeskription: Eine quantitative und qualitative Untersuchung der Sendung *Passe-moi les jumelles*. MA thesis, ZHAW Zürcher Hochschule für Angewandte Wissenschaften.
35. Lv Q, Liang J (2019) Is consecutive interpreting easier than simultaneous interpreting? A corpus-based study of lexical simplification in interpretation. Perspectives 27(1):91–106.
36. Mangiron C, Zhang X (2025) Translated from English or created in Spanish? The reception of audio description for video games: An experimental pilot study. Translation Spaces 14(1):146–169.
37. Matamala A, Boix CO (2016) Accessibility and multilingualism: An exploratory study on the machine translation of audio descriptions. TRANS: Revista de Traductología 20:11–24.
38. Mazur I (2024) Same film, different audio descriptions: Describing foreign films from a functional perspective. Universal Access in the Information Society 23(2):511–521.
39. Molés-Cases T, Alonso R (2025) Audio description across languages: Motion and streaming services. Perspectives 1–18.

40. Orero P (2008) Introduction: Audiovisual translation: A new dynamic umbrella. In: Díaz-Cintas J (ed) Topics in Audiovisual Translation. John Benjamins, Amsterdam and Philadelphia, pp vii–xiii.
41. Read JA (2000) Assessing vocabulary. Cambridge University Press, Cambridge.
42. Reiss K (1989) Text types, translation types and translation assessment (trans. Chesterman A). In: Chesterman A (ed) Readings in Translation Theory. Oy Finn Lectura Ab, Helsinki, pp 105–115. (Originally published 1977)
43. Remael A, Vercauteren G (2010) The translation of recorded audio description from English into Dutch. Perspectives 18(3):155–171.
44. Raído VE, Angelone E, Olalla-Soler C (2025) Screen recording. In: Angelone E, Raído VE, Olalla-Soler C (eds) Research Methods in Cognitive Translation and Interpreting Studies. John Benjamins, Amsterdam and Philadelphia, pp 133–156.
45. Reviers N (2016) Audio description services in Europe: An update. The Journal of Specialised Translation 26:232–247.
46. Saldanha C (dir) (2006) *Ice Age: The Meltdown* [Film]. Blue Sky Studios; 20th Century Fox Animation.
47. Sweller J (1988) Cognitive load during problem solving: Effects on learning. Cognitive Science 12:257–285.
48. Sweller J (2005) The redundancy principle in multimedia learning. In: Mayer RE (ed) The Cambridge handbook of multimedia learning. Cambridge University Press, Cambridge, pp 159–167.
49. Szarkowska A, Jankowska A (2015) Audio describing foreign films. The Journal of Specialised Translation 23:243–269.
50. Tor-Carrogio I (2020) Time to start synthesising audio description in China? Results of a reception study. The Journal of Specialised Translation 34:171–191.
51. Vera JF (2006) Translating audio description scripts: The way forward? Tentative first stage project results. Paper presented at MuTra 2006 – Audiovisual Translation Scenarios, Copenhagen.
52. Vercauteren G (2012) A narratological approach to content selection in audio description: Towards a strategy for the description of narratological time. MonTI: Monografías de Traducción e Interpretación 4:207–231.
53. Vercauteren G, Reviers N, Steyaert K (2021) Evaluating the effectiveness of machine translation of audio description: The results of two pilot studies in the English-Dutch language pair. Tradumàtica: Tecnologies de la Traducció 19:226–252.
54. Wedge C, Saldanha C (dirs) (2002) *Ice Age* [Film]. Blue Sky Studios; 20th Century Fox Animation.
55. Wood D, Bruner JS, Ross G (1976) The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry 17(2):89–100.
56. Yan JX, Luo K (2023) Audio description and interpreting training: A comparison of assessment criteria from the perspective of learners. Perspectives 31(6):1049–1064.
57. Young K, Saver JL (2001) The neurology of narrative. SubStance 30(1):72–84.
58. Zhang X, Mangiron C (2025) Audio description in video games. Babel 71(4):527–547.

59. Zheng J, Fan W (2021) Different processes for translating expressive versus informative texts? A computer-assisted study of professionals' English–Chinese translation. Digital Scholarship in the Humanities 36(3):782–793.

Figures

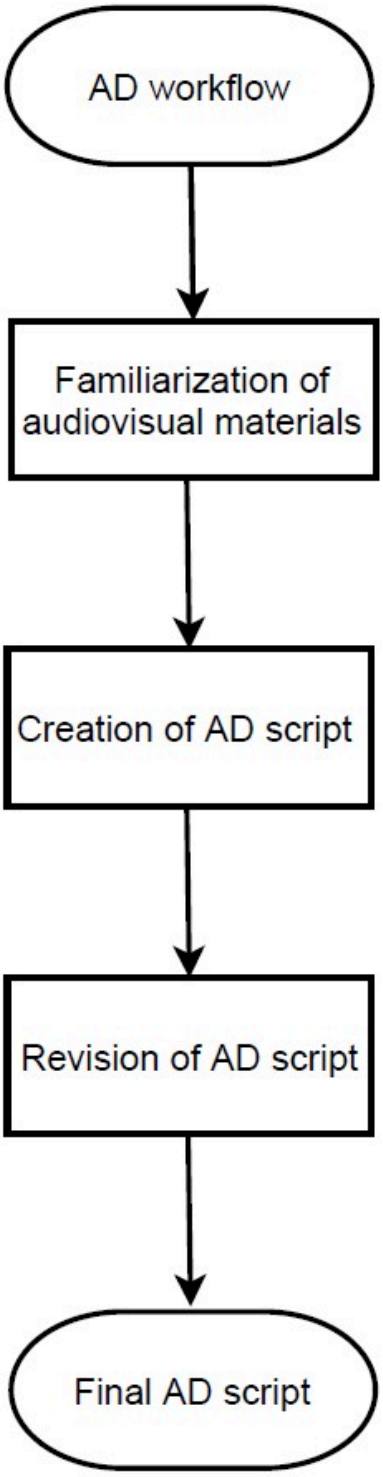

Figure 1

Direct AD script workflow.
Flowchart of the direct audio description (AD) script production procedure. The workflow comprises familiarization with the audiovisual materials, drafting of the AD script and revision of the AD script.

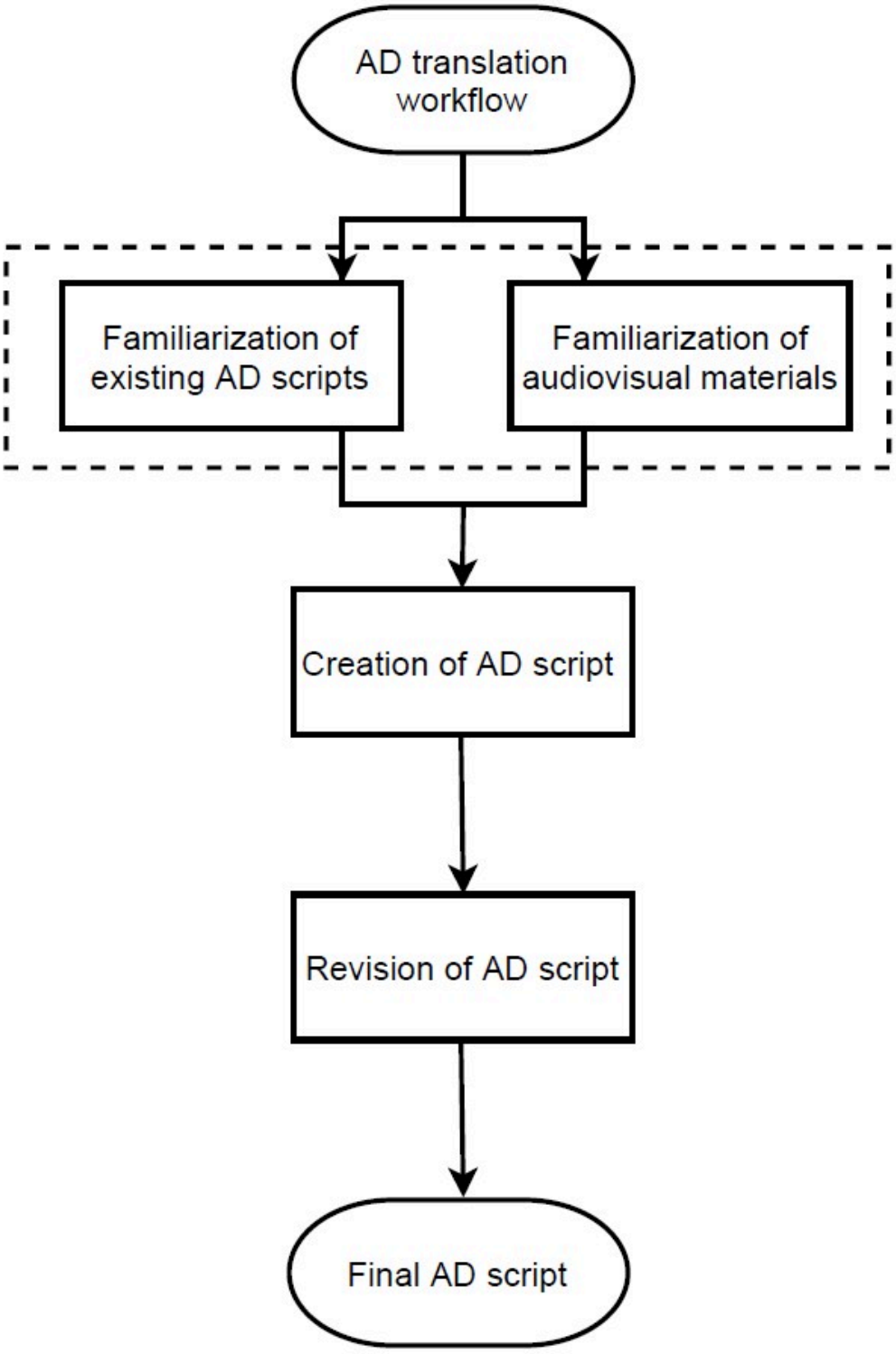

Figure 2

AD translation workflow.
Flowchart of the AD translation procedure. The workflow includes familiarization with the audiovisual

materials and the source AD, drafting of the AD script and revision of the draft.

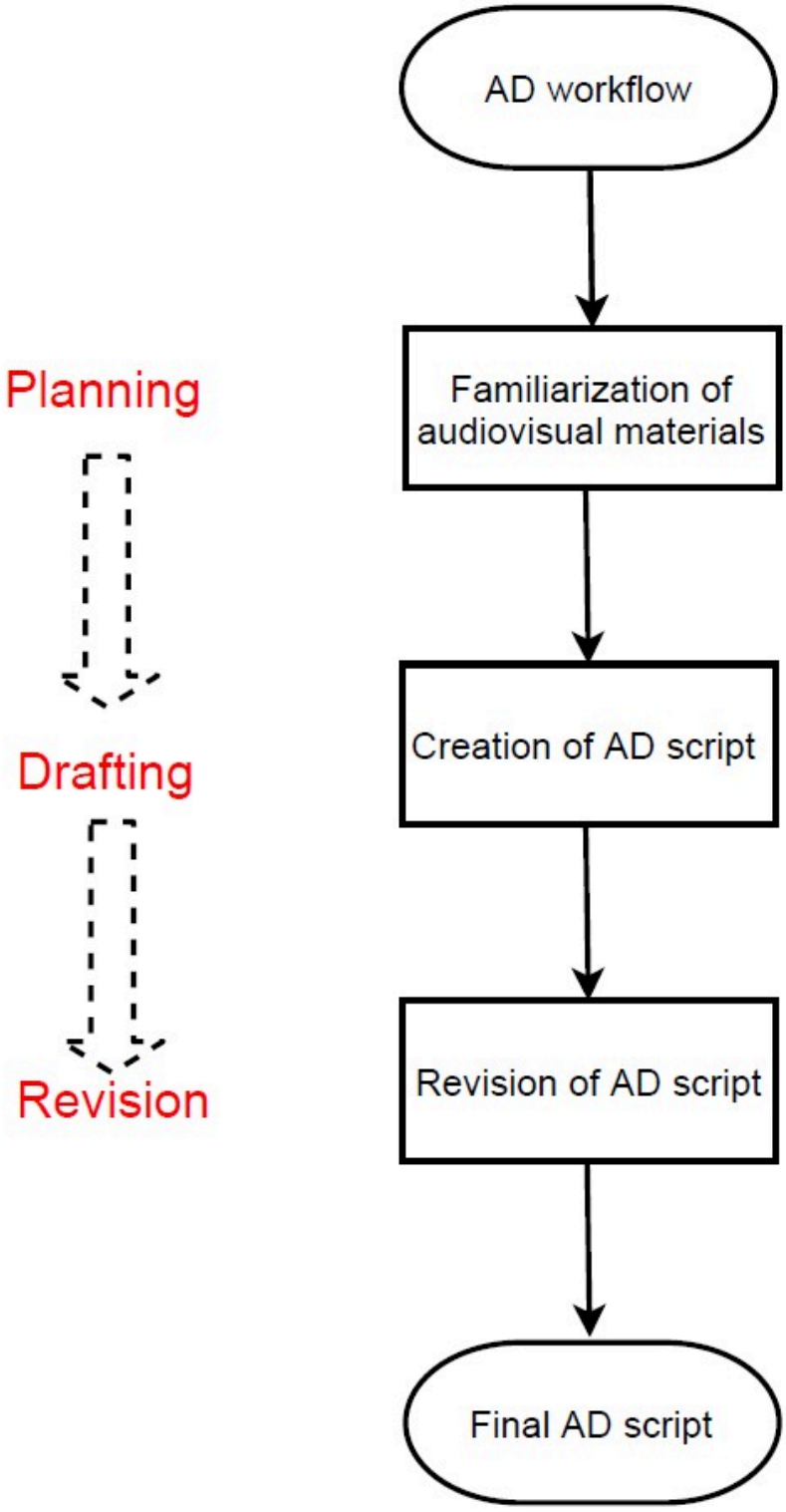

Figure 3

A three-phase framework of direct AD.
Three-phase framework used to conceptualize and segment the direct AD process. Familiarization, drafting, and revision activities were aligned with the planning, drafting, and revision phases.

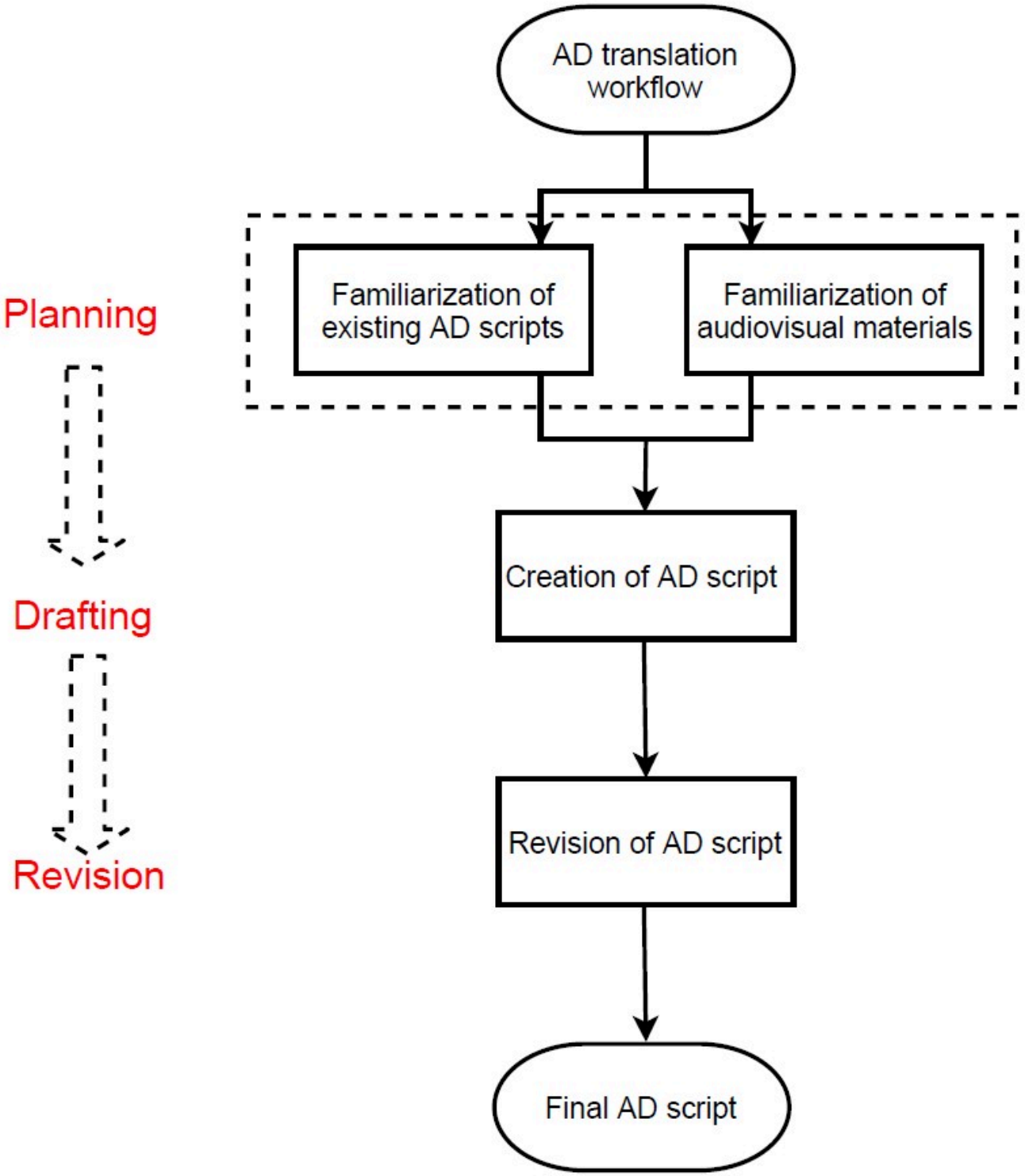

Figure 4

A three-phase framework of AD translation.

Three-phase framework used to conceptualize and segment the AD translation process. The planning phase involves processing both the audiovisual materials and the source AD, followed by drafting and revision of the AD script.

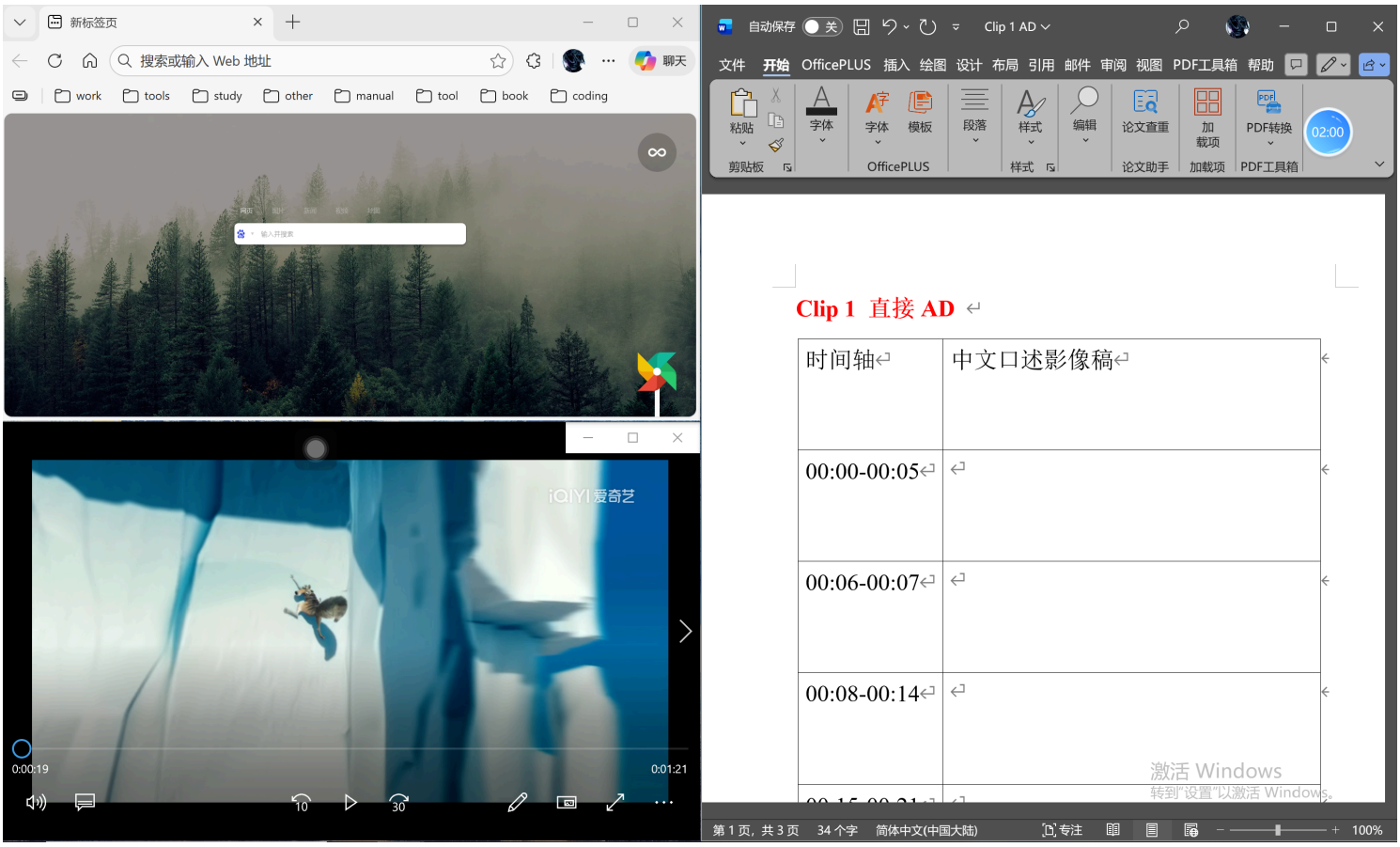

Figure 5

Screen setting in the direct AD condition.

Screen layout of the task interface in the direct AD condition, showing the arrangement of the video stimulus, external consultation resources, and the AD script-writing area.

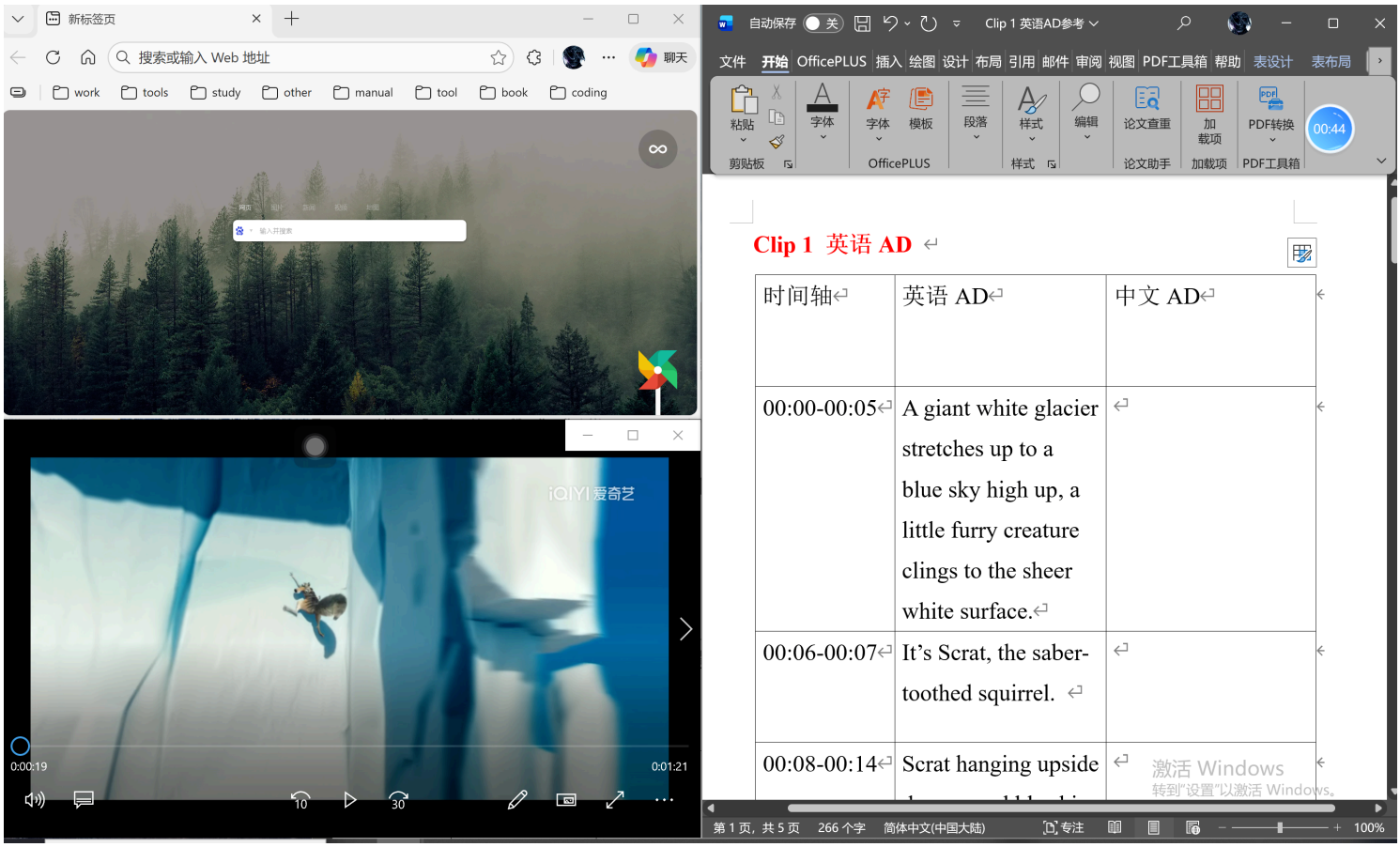

Figure 6

Screen setting in the AD translation condition.
Screen layout of the task interface in the AD translation condition, showing the arrangement of the video stimulus, external consultation resources, the source AD display, and the AD script-writing area.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Rscriptforphaseanalysis.docx
- Rscriptforperceiveddifficultyanalysis.docx
- AD.docx
- inputlogdataraw.xlsx
- inputlogdataclean.xlsx
- ADdifficultydata.xlsx
- ADTaskQuestionnaireinEnglish.docx