

Couto-Cantero, P., Sabaté-Carrové, M. & Tinedo-Rodríguez, A. (2022). *Effectiveness and assessment of English Production Skills through Audiovisual Translation. Current Trends in Translation Teaching and Learning E*, 9, 149 – 182.  
<https://doi.org/10.51287/cttl20225>

## **EFFECTIVENESS AND ASSESSMENT OF ENGLISH PRODUCTION SKILLS THROUGH AUDIOVISUAL TRANSLATION**

Pilar Couto-Cantero<sup>1</sup>, Mariona Sabaté-Carrové<sup>2</sup> and Antonio Tinedo-Rodríguez<sup>3</sup>

<sup>1</sup>Universidade da Coruña, <sup>2</sup>Universitat de Lleida, <sup>3</sup>Universidad Nacional de Educación a Distancia (UNED)

### **Abstract**

This article is framed within a research project based in Spain aimed at studying the use of Audiovisual Translation tools to develop communicative competences for teaching and learning English as a Foreign Language. The aims of the study are to get to know to what extent the oral and written production skills of the participants improve thanks to the implementation of an Audiovisual translation-based course, and to get to know if the overall performance of the post Integrated Skills Test is

better than the pre-Integrated Skills Test after the learning process. It is also aimed to prove the validity of those language assessment tests. A mixed methodology (qualitative and quantitative analysis) was used to obtain information from the participants. Research data was collected from different universities in Spain with an initial sample of 40 applicants, 8 of which finally completed the course. Data analysis shows that oral and written production skills improve thanks to the implementation of six audio description-based lesson plans completed during a pilot course offered during the summer of 2021 to volunteer adult participants with a B2 level in English language. Furthermore, the study illustrates that the results of the final tests are better than the results of the previous ones after the intervention. We discuss the results obtained in this study and conclude that they are harmonious with former studies on the topic as they validate the use of tests to improve the development of communicative skills. Although, there is a very limited number of participants in the sample, we consider this may be used as an example of a trend that will be explored in the future, as this is a pilot study included in a wider research project which is still underway.

**KEYWORDS:** audiovisual translation; audio description; language tests; student evaluation; second language learning; English as a foreign language; integrated skills approach.

## 1. INTRODUCTION

This article is framed within the TRADILEX project, aimed at studying the use of Audiovisual Translation (AVT) tools to develop communicative competences for teaching and learning English as a Foreign Language (EFL). From the different modalities offered by AVT: Dubbing, Subtitling, Subtitling for the Deaf and the Hard of Hearing

(SDH), Voice-over and Audio description, we have narrowed our research on a piloting study based on an Audio description course (AD) for B2 CEFR level adult participants.

The importance of this research for the field lies on two main concerns never studied and analysed to date. The first one is the final design of a series of original tailored-made Integrated Skills Tests (ISTs) i.e.: an Initial Test of Integrated Skills (ITIS) and a Final Test of Integrated Skills (FITIS) to be completed before and after the course (Couto-Cantero et al., 2021). The second one consists in the analysis of a piloting course focused on the implementation of 6 AD-based lesson plans to confirm the performance of the oral and written production skills of the participants.

The potential of using AVT for learning languages and developing communicative competences has been studied so far (Danan, 2015; Fernández-Costales, 2017; Frumuselu et al., 2015; Lertola, 2019; Talaván, 2013). Similarly, other AVT combinations with learning EFL have been studied, e.g.: dubbing, audio description or overlapping voices, as well as other subtitling modalities began to be used for teaching and learning purposes (Danan, 2010; Gómez Pérez & Roales Ruiz, 2015; Navarrete, 2013; Navarrete, 2021; Sánchez-Requena, 2016; Talaván & Lertola, 2016; Talaván et al., 2016; Talaván & Rodríguez-Arancón, 2014a,

2019). Despite these valuable contributions, there is a research gap in the lack of studies and specific research showing the actual results of implementing such tools for teaching and learning foreign languages on students.

Therefore, the main purpose of this article is precisely intended to demonstrate the effectiveness of these Integrated Skills Tests (IST) based on implementing audio description (AD) lesson plans for learning English as a Foreign Language (EFL). Additionally, it aims to confirm the validity of the ITIS and FITIS for a B2 level (CEFRL) by analysing the results obtained in pilot tests according to variables and other data analysis obtained from this research. The specific objectives planned to achieve are the following: (a) to get to know if the participants' Oral Production improves after the learning process acquired by means of AD activities from the lesson plans, (b) to get to know if the participants' Written Production improves after the learning process acquired by means of AD activities from the lesson plans, and (c) to get to know if the overall performance of the B2 FITIS is better than the B2 ITIS after a learning process acquired by means of AD-based lesson plans.

The virtual campus of the TRADILEX project was used as a platform to collect all tests, tasks, lesson plans and information gathered to elaborate this research. This analysis uses a mixed methodology

with quantitative and qualitative data obtained from the piloting practice delivered asynchronously during the Summer Term of 2021. Initially, forty adult participants applied for the course as volunteer learners resulting in a total of only eight who finally completed the course together with the initial and final tests. Instruments used for this research were an initial questionnaire, initial test (B2 ITIS), a set of 6 AD-based lesson plans, a final test (B2 FITIS), and a final questionnaire.

## **2. THEORETICAL FRAMEWORK**

The current global pandemic situation has challenged education techniques across the planet (United Nations Educational, Scientific and Cultural Organization, 2020). Educators have been compelled to shift from face-to-face classes to emergency remote teaching (Hodges et al., 2020), re-plan, re-schedule and redesign their classes to an online format of language teaching (Ferdig et al., 2020; Gacs et al., 2020). Foreign language teachers have provided countless innovative opportunities to continue engaging and motivating themselves in each aspect of the teaching process. This unprecedented global pandemic has brought to the fore the need for urgent changes in all aspects of teaching. One of these aspects is language assessment.

Although this pandemic has accelerated exponentially the urgency for changes in foreign

language teaching, in the last decade, language assessment has veered from testing the examinees' skills separately to doing so by integrating and combining the examinees' abilities (Weir, 1983). Thus, it is not uncommon to find tests that integrate source reading or listening material into their writing and speaking performance in ways that simulate the cognitive, communication, and literacy demands of real-life academic or vocational tasks (Cumming 2013a, 2013b; Gebril & Plakans, 2013; Plakans & Gebril, 2012; Shaw & Weir, 2007; Tunkay, 2014; Weigle, 2004). This assessment based on testing integrated skills incorporates several skills within one test to determine whether a student can tackle the complexity of real-world tasks in academia that require multiple skill sets (Horowitz, 1986; Sawaki et al., 2013; Strickland, 2015). This type of assessment is known as Integrated Skills Test (IST).

Previous ISTs can be found on and off the internet. To name just a few, TEAP (Test of English for Academic Purposes) by the Associated Examinations Board (Weir, 1983), IELTS (Writing Task 1: Cambridge English), TOEFL IBT, PTE Academic by Pearson, ISE (Integrated Skills in English) Test by Trinity College, TEAP (2016) by the Eiken Foundation and Sophia University. There are several advantages of an IST over more conventional tests. First and foremost, an IST is more authentic, as the test tasks replicate a real-life language-in-use task in a specific context and situation (Davidson, 2009; Davidson & Dalton,

2003; Davidson & Hobbs, 2003) and are grounded on interaction and authenticity upon which meaning is built (Davidson, 2009). In this respect, an IST constitutes a more faithful representation of what communication is and is in line with the Common European Framework of Reference for Languages (CEFR) 's general measures "to promote, encourage and support the efforts of teachers and learners at all levels to apply in their own situation the principles of the construction of language-learning systems [...] by developing suitable forms and instruments for the evaluation of learning programmes." (Council of Europe 2001:3). With its communicative language activities and strategies, the CEFR replaces the traditional model of the four skills (i.e., listening, speaking, reading and writing), which has increasingly proved inadequate to capture the complex reality of communication, with an IST that will better test the students' skills prior to and after a course.

Admittedly, an IST shares certain features with traditional assessment tests, e.g., both require from students to produce written or oral work that incorporates meaningful uses of source evidence, both conceptually -to comprehend, synthesize, and present ideas from sources- and through writing -to conform with stylistic convention for presenting ideas from sources, and acknowledging those sources, as suggested by Cumming (2013a, 2013b). However, most traditional assessment tests fail to incorporate elements from other areas that would

give a more eclectic nature to tests. One of them is the audiovisual component, specifically audiovisual translation (AVT). It is paradoxical that after years of rejection on the entertaining nature of AVT, and its ensuing acceptance as a tool for language learning (Torralba, 2020), AVT (or the AV component for the sake of it) is still absent or poorly represented in ISTs. We believe that there is room for improvement in this area and thus we have incorporated AVT in the design of our Integrated Test of Initial Skills (ITIS) and Final Test of Integrated Skills (FITIS) (Couto-Cantero et al., 2021).

Finally, in line with CEFRL tenets, our ISTs (ITIS and FITIS) align with the importance of mediation. In mediation, the user/learner acts as a social agent who creates bridges and helps to construct or convey meaning, sometimes within the same language, sometimes from one language to another (cross-linguistic mediation). The focus is on the role of language in processes like creating the space and conditions for communicating and/or learning, collaborating to construct new meaning, encouraging others to construct or understand new meaning, and passing on new information in an appropriate form. The context can be social, pedagogic, cultural, linguistic or professional (CEFRL Companion Volume with new descriptors, p. 103). ITIS and FITIS need to include mediation in a variety of ways before the student does the activity, as suggested by North & Piccardo (2016): by

explaining activities, by supplying information that may anticipate obstacles, by giving examples that may illustrate ways to carry out the activities, by relaying specific information, by giving contextual information (linguistic, cultural, semantic or technical) of the (oral or written) text.

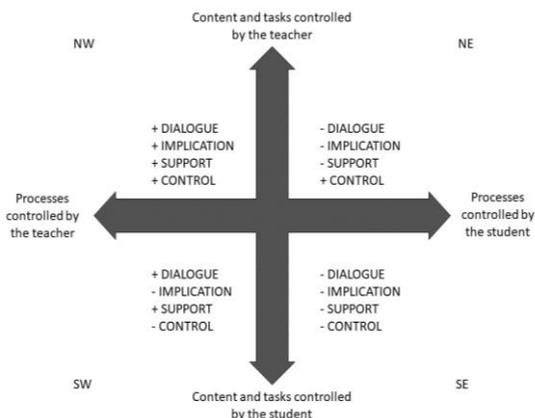
### **3. METHOD**

#### **Context**

The context of the study was more challenging than traditionally because it was implemented through a virtual course designed in a virtual platform of our research team in the summer of 2021. Following the model of Coomey & Stephenson (2001) on e-learning courses, the course corresponded to the upper-right sector of Figure 1. The course was implemented by fully autonomous students. It is also important to highlight that the circumstances were not auspicious because during that period of time the 5<sup>th</sup> wave of the COVID-19 pandemic hit the whole country.

## Figure 1

*Theoretical paradigms of e-learning by Coomey & Stephenson (2001)*



Besides, even though the TRADILEX project follows a holistic approach to language learning focusing on fluency and accuracy, this study emphasises the development of fluency, explores the possible relationship of this didactic resource with communicative methods and delves into the role of mediation in teaching EFL.

## Mixed Research

The nature of this research is mixed since apart from the quantitative data gathered from the ITIS and FITIS tests, it is crucial to triangulate that information with the qualitative evidence retrieved

from the questionnaires to gain a deeper insight into the research topic.

When designing our assessment tests, we borrowed many of the rating criteria suggested by Chan et al. (2015), e.g., standardisation sessions, double marking led by two researchers, blind marking, monitoring of evaluator's performance, and had a small pool of evaluators. Likewise, checking for plagiarism (SafeAssign, Turnitin) was an essential feature in assessing results. Therefore, the assessment was as authentic, reasonable and well-grounded as possible by: a) allowing plenty of time for students to read the source written texts, b) allowing plenty of time for students to view the source oral texts, c) allowing plenty of time for students to complete the task, d) using multiple topics and motivating source texts, e) allowing students to ask questions about the source texts (feedback), f) allowing students to take notes on their laptop, g) allowing students access to internet, spell checker, thesaurus. More details on the design of the tests (ITIS and FITIS) in Couto-Cantero et al. (2021).

### **Test administration**

The whole course was internally available online and asynchronous and was hosted in the TRADILEX platform. Students were asked to meticulously adhere to the guidelines set out in the virtual course and do the activities in the exact order in which they were displayed in the platform. They completed the test online asynchronously through a

link to Google Forms. Once completed, they uploaded their test as PDF files onto the virtual platform (Figure 2).

## Figure 2

### *Virtualisation of the ITIS test*

INITIAL TEST OF INTEGRATED SKILLS (ITIS)

 ITIS (RECEPTION SKILLS) [60 MINUTES]

Please make sure you can spend 1 hour before doing the test.

 ITIS (RECEPTION SKILLS) [60 MINUTES]

 ITIS (PRODUCTION SKILLS) [60 MINUTES]

Please make sure you can spend 1 hour before doing the test.

 ITIS (PRODUCTION SKILLS) [60 MINUTES]

*Note.* (Own elaboration)

## **Sample selection, data collection and participants**

The sampling method was random, and the sample was stratified as enrolment was limited to B2 students. The course was offered to students of the Centro Universitario de Idiomas a Distancia (CUID) and to students at the University of A Coruña (UDC). Those interested (n=40) enrolled the course, yet only (n=8) students completed it (n=8) with varying age groups, as seen in Table 1.

Data collection involved Moodle and Google Forms. The information was collected in the programme SPSS v.27.0.1.0.

**Table 1**

*Information about the age of the participants*

<b>Age group</b>	<b>Percentage</b>
18-30	30% (n=12)
31-40	15% (n=6)
41-50	42.5% (n=17)
51 or more	12.5% (n=5)

Regarding gender, it is important to highlight that 80% (n=32) were women, whilst 17.5% (n=7) were men, and one student (n=1) preferred not to say.

### **Instruments and materials**

The instruments used for this research are the following ones:

- a) Initial questionnaire. This questionnaire was implemented asynchronously online and aimed at gathering data on sociodemographic variables and on former experiences with AVT.
- b) The ITIS test. The purpose of this test is to quantitatively measure the 4 skills before the intervention. According to Couto-Cantero et al. (2021) the innovation of this test lies on mediation.
- c) The set of 6 LP on AD for B2 level. This set of LP conforms the intervention. Students are expected to create their own voice-over version of the videos in the lesson plans and to complete a set of activities aiming at both

facilitating the understanding of the video and developing their language skills.

- d) The FITIS test. This test is parallel to the ITIS test and is implemented at the end of the intervention to measure the possible improvement.
- e) The final questionnaire. This test is an online questionnaire in which students provide feedback to the research team with information on their experience. It contains different sections: the process, the result, the development of interculturality, and the degree of satisfaction.

The materials used were the following:

- a) Virtual platform. The virtual platform (tradilex.es) was supported by the Moodle technology and an *ad hoc* course was created for the experimental design.
- b) Invitation letter. Participants were invited through an invitation letter which contained the registration procedure and a welcoming message.
- c) User's guide. A user's guide was created for participants to know how the virtual course works and to provide them with links to tutorials on how to use the recording software (Lightworks and Screencastify).

### **Validation of the instruments**

To guarantee the validity and objectivity of the data, the reception variables (written reception and oral

reception) were automatically assessed by the virtual platform. To assess the production tasks linked to the variables like written production and oral production, the research team took the decision of making use of the TRADILEX rubric and bearing the ALTE “Can do statements” in mind (Council of Europe, 2001). Each production test was assessed with a double-blind review process, and the final mark consisted of the average mark of the evaluators. Moreover, to avoid plagiarism, the texts submitted by students were uploaded and run onto the Turnitin plagiarism checker.

## **Variables**

Due to the nature of a mixed research, qualitative and quantitative variables were used to test the research questions. Regarding the former variables, the final questionnaire had the objective of gathering qualitative information on the process. The item 6 specifically focused on the self-perception of the improvement of the four skills. Each of the aforementioned variables (oral reception, oral production, written reception and written production) consisted of a non-binary Likert-based scale with six degrees of agreement: Strongly agree, Agree, Slightly agree, Slightly disagree, Disagree, Strongly disagree.

As far as the quantitative variables linked to the four skills (reading, writing, speaking and listening) are concerned, these were assessed through the ITIS and FITIS tests (Couto-Cantero et al., 2021). Each skill was measured through a continuous quantitative

variable whose value ranges from 0 to 10 and they were assessed in two different ways:

- Reception skills. These skills were assessed automatically because they consisted of multiple-choice test whose right answers were pre-recorded when the test was designed.
- Production skills: These skills were assessed by the research group by making use of rubrics.

To sum up there were 8 variables which are displayed in table 2:

**Table 2**  
*Variables of the study*

Variable	Categorisation	Instrument
Oral reception	Continuous quantitative	ITIS and FITIS
Oral production		
Written reception		
Writing production		
Oral reception (self-perception)	Politomic qualitative	Final questionnaire
Oral production (self-perception)		
Written reception (self-perception)		
Writing production (self-perception)		

Table 2 summarises the eight variables which have been considered for the main conclusions of the study. It consists of the four skills (oral reception, oral production, written reception, and written production) and the self-perception of improvement that students experience after the intervention of the same variables. The objective is to contrast through triangulation whether the quantitative data match the

qualitative data. The four skills are measured through the ITIS (pre-test) and the FITIS (post-test) and they are continuous. Each skill consists of 10 points and the highest possible mark for each test is 40 points. The self-perception variables are qualitative and have been measured through a six-point Likert scale in the final questionnaire.

#### 4. RESULTS

Considering the size of the final sample ( $n=8$ ), the inferential methods applied were non-parametric. Particularly, to explore the differences between the different skills in the ITIS and FITIS, the Wilcoxon signed-rank test was applied, and obtained the results displayed in Table 3:

**Table 3**

*Quantitative analysis of the results of the pre-piloting experience*

	Oral production	Written production	Oral reception	Written reception
n	8	8	7	7
p-value	0.004	0.004	0.187	0.343
Positive differences	8	8	4	4
Negative differences	0	0	1	2
Match	0	0	2	1

Table 3 shows how the difference is statistically significant for oral production ( $p\text{-value} = .004$ ) and for written production ( $p\text{-value} = .004$ ). The Wilcoxon signed-rank test confirms the tendency towards the improvement after the intervention

because the number of positive differences is 8 for both variables. These results are harmonious with the ones obtained by Lertola & Talaván (2015) in their study on active audio description to promote speaking skills online.

Nonetheless, with reference to reception, the difference is not statistically significant for oral reception (p-value = .187) nor for written reception (p-value = .343). It is worth commenting that a participant did not complete the reception final test, hence  $n = 7$  in the reception section. Likewise, if we take the Wilcoxon signed-rank test (Positive differences, Negative differences, and Match) into account, it is important to highlight that as much as 57% of participants in the sample improved their reception skills (oral and written reception) after the intervention.

## **Results of the analysis of RQ 1**

Bearing the limitations of the sample in mind because of its size, the intervention based on the AD sequence seems to have a positive impact on the oral production skills of the participants because the p-value = 0.004 implies that the differences between the pre-test (ITIS) and the post-test (FITIS) of that skill are statistically significant. Moreover, the Wilcoxon signed-rank test confirms that the tendency is that of improvement since 100% of them obtained a better mark in the post-test. Results are displayed in Figure 4.

## Figure 4

*Comparison of the pre-test and post-test (oral production)*

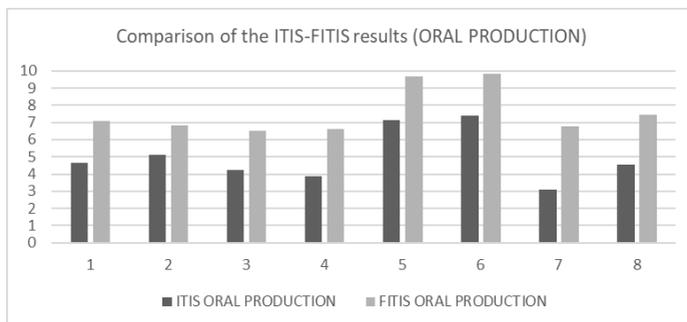


Figure 4 shows the mark out of 10 of the 8 participants (x-axis) in the pre-test (ITIS) and in the post-test (FITIS). The y-axis scale reaches up to 10 since it was the maximum possible mark for this variable. As the statistical analysis confirms, the tendency in the 8 participants is that of improvement because the ITIS marks are lower than the FITIS marks.

## Results of the analysis of RQ2

The intervention also seems to have a positive impact on the written production skills since the p-value is 0.004 and is thus statistically significant. Furthermore, the Wilcoxon signed-rank test confirms the improvement of students since 100% of them obtained a better mark in the post-test. Results are displayed in Figure 5.

## Figure 5

*Comparison of the pre-test and post-test (written production)*



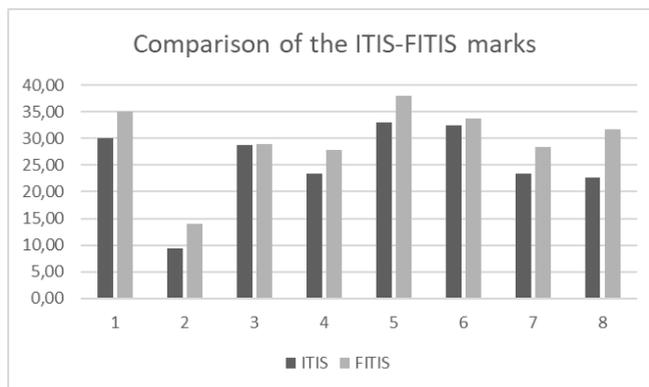
In Figure 5 we can observe the performance of the eight participants with regards to written production by comparing their marks out of 10 in the pre-test (ITIS) and in the post-test (FITIS). It is important to note that the y-axis scale reaches up to 10 because it was the maximum possible mark for this variable. The results in the figure are harmonious with the statistical analysis being the tendency that of improvement.

### Results of the analysis of RQ 3

To answer this RQ 3, triangulation is required because it is important to observe the results obtained quantitatively, but it is also important to contrast them against the ones gathered qualitatively on student's perceptions. Figure 6 shows the ITIS-FITIS B2 marks of the participants:

**Figure 6**

*Comparison of the marks of the pre-test and post-test*



*Note.* Own elaboration

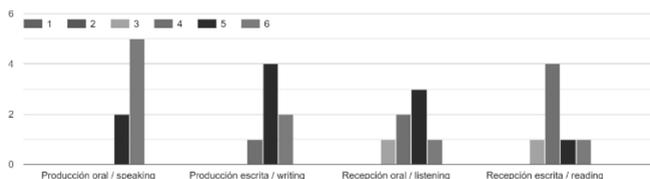
As the reader may observe, all of them improved after the intervention because the FITIS marks are higher. By looking at Figure 6 we can observe the general performance of the eight participants with regards to the four skills. Marks are out of 40 because each skill consists of 10 points. The results in this figure are harmonious with the statistical analysis being the general tendency that of improvement. It is important to note that participant number 2 only handed two parts of each test and that is why the mark is much lower than the one of other participants.

## Final results of the analysis

Finally, it is crucial to compare quantitative data obtained from the ITIS and FITIS tests together with qualitative data gathered from the questionnaires. Particularly, the participants' self-perceptions of improvement of the four skills were compared with the real marks they obtained. Results shown in Figure 7 are harmonious with the ones retrieved from the final questionnaire in which participants were asked which communicative skills were improved thanks to learning obtained from the set of 6 AD-based LPs.

### Figure 7

*Self-perceptions of improvement of the four skills after the intervention*



*Note.* Likert scale: 1=Totally disagree, 2=Disagree, 3=Slightly disagree, 4=slightly agree, 5= Agree and 6=Totally agree

All the participants, 100% (n=7) strongly agreed or agreed with the fact that the AD sequence had improved their production skills (written and oral), whilst when it comes to reception skills there were divergences. Regarding oral reception, 57.1% (n=4)

strongly agreed or agreed but 42.8% (n=3) slightly agreed or disagreed. As regards written reception 57.1% (n=4) slightly agreed with the fact that this skill was improved through AD. The results seem to match, and they are harmonious because written production and oral production are the skills whose improvement was noticeable in the quantitative analysis, and they are the skills which students perceived they have developed the most.

## **5. DISCUSSION AND CONCLUSION**

Despite the size of the sample, the results of AD as a resource to learn EFL are promising because it seems to enhance production skills in online environments. Lertola & Talavan's (2015) study served as a departure point, and the present study confirms that the use of AD fosters both speaking and writing performance. Recent research available by means of an unpublished doctoral thesis also validates the improvement of oral skills through AD (Navarrete, 2021). Therefore, the results obtained in this study are harmonious with former studies on the topic as they validate the use of tests as a means to improve the development of communicative skills (Danan, 2015; Fernández-Costales, 2017; Frumuselu et al., 2015; Lertola, 2019; Talaván, 2013), even though the literature on didactic AD is still scarce.

The potential of AD to foster production skills lies on the fact that participants have to produce their own script (written production) and they have to

record a new audio track with their voices over the original track explaining every single detail that a blind person may need to understand the whole clip (oral production). It is thus an interesting mediation activity which subsidiarily may foster awareness on the importance of accessibility.

The ITIS and FITIS tests have been designed from a plurilingual, mediation-based approach. As we considered that traditional designs for tests, in the current context, are not complete and do not assess all skills properly and in an engaging way, we agreed that our ITIS and FITIS B2 should be more authentic, they should capture what students really do more accurately, have a better context, cognitive, consequential and predicative validity. They should also have a major positive washback effect upon teaching and learning and better prepare students for demands of higher study.

On the other hand, evaluators need to reconsider key assessment principles such as validity, reliability, practicality, test design and control of environmental conditions. Moreover, the adoption of integrated-skills assessment has the potential to transform the way we think about assessment.

Our results confirm the validity of our 3 RQ as it has been proved that the participants' Oral Production improves after the AD intervention (RQ 1), the participants' Written Production improves after the learning process acquired by means of AD activities (RQ 2) and the overall performance of the B2 FITIS is better than the B2 ITIS after a learning process

acquired by means of AD-based lesson plans (RQ 3). Although the results obtained confirm the RQs posed at the outset of the study and show promising results, we are aware of several limitations.

One of the main obstacles faced at the outset was tradition and resistance to break it. If traditional tests worked, why bother designing new ones? The answer to this is our impression that traditional tests do not actually work, as the communicative function of language activities (i.e., involving reception, production, interaction, or mediation) requires significant improvement. Likewise, our target students were used to more traditional approaches to assessment. Another challenge we faced was in the marking and evaluation process, namely diagnosing why a student fails is rarely a straightforward clear-cut endeavour. Teachers often ignore why a student has failed a question especially in multiple choice tests. This is then a likely limitation of our study. Other practical decisions were the actual duration of the ISTs. In order to properly test a student using an IST, we needed to be thorough and cover as many areas of their EFL knowledge as possible. We finally reached a compromise of 2 hours altogether and divided the section into time-lapses of 15 minutes for each section. However, alterations of such parameters could result in different outcomes, i.e., shorter time or division of sections. Foreign language learners usually have more difficulties when it comes to carry out production tasks than to reception tasks. The fact of being assessed and the

limitation of time is also a handicap for them to feel nervous and worried or even insecure.

An added limitation of this research was the varying environmental test conditions, whereby technical feedback was not offered immediately. Some examples of such test conditions resulted in questioning the reliability of data (danger of plagiarism/cheating), suspicion of subjective rating since evaluators were human (especially in production activities), problems with standardization (which we overcame by establishing a pattern of 5 questions and 3 possible answers), and need for training, as computer literacy of participants cannot always be taken for granted.

One of our recommendations for further research is to promote this kind of ISTs and AVT-based proposals in other languages different to English. Moreover, they could also be implemented, not only in a B2 level, as the example we have provided in this article, but also in all levels of the CEFRL. It would also be very interesting to translate this model, which was born within a funded National Project in Spain (TRADILEX), to other international countries trying to find partners for the European Union competitive calls.

## Funding

PID2019-107362GA-I00  
AEI/10.13039/501100011033. Spanish  
Government, Science, and Innovation Ministry /  
Gobierno de España, Ministerio de Ciencia e  
Innovación.

Antonio J. Tinedo-Rodríguez is funded with a  
Santander-UNED predoctoral grant.

## References

- Burns, C., & Foo, M. (2009). Integrating content and academic skills: working in tandem to meet learning outcomes and assist the transition of international students. *EMERGE, 1*, 32-40.
- Chan, S. H. C., Inoue, C., & Taylor, L. (2015). Developing rubrics to assess the reading-into-writing skills: A case study. *Assessing Writing, 26*(1), 20-37.
- Coomey, M., & Stephenson, J. (2001). Online learning: it is all about dialogue, involvement, support and control-according to research. In J. Stephenson, *Teaching and Learning Online: Pedagogies for New Technologies*. Routledge.  
<https://doi.org/10.4324/9781315042527>
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Press Syndicate of the University of Cambridge.

- Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing. [www.coe.int/lang-cefr](http://www.coe.int/lang-cefr)
- Couto-Cantero, P., Sabaté-Carrové, M. & Gómez Pérez, M. C. (2021). Preliminary design of an Initial Test of Integrated Skills within TRADILEX: an ongoing project on the validity of audiovisual translation tools in teaching English. *REALIA. Research in Education and Learning Innovation Archives*, 27(2), 73-88. <https://doi.org/10.7203/realia.27.20634>
- Cumming, A. (2013a). Assessing integrated skills. In A. Kunnan (Ed.). *The Companion to Language Assessment*, Volume 1 (pp. 216-229).
- Cumming, A. (2013b). Assessing Integrated Writing Tasks for Academic Purposes: Promises and Perils. *Language Assessment Quarterly*, 10,1, 1-8. <https://doi.org/10.1080/15434303.2011.622016>
- Danan, M. (2010). Dubbing projects for the language learner: A framework for integrating audiovisual translation into task-based instruction. *Computer Assisted Language Learning*, 23(5), 441–456.
- Danan, M. (2015). Subtitling as a language learning tool: past findings, current applications, and future paths. In Y. Gambier, C. Mariotti y A.

- M. Caimi (Eds.), *Subtitles and Language Learning. Principles, strategies and practical experiences* (pp. 41–61). Peter Lang.
- Davidson, P. (2009). Authentic assessment in EFL classrooms. In C. Coombe, P. Davidson & D Lloyd (Eds.), *The Fundamentals of Language Assessment: A Practical Guide for Teachers*, 2<sup>nd</sup> Edition (pp. 213-224). TESOL.
- Davidson, P. & Dalton, D. (2003). Multiple-measures assessment: Using 'visas' to assess students' achievement of learning outcomes. In C. A. Coombe & N. Hubley (Eds.). *Assessment Practices* (pp. 121-134). TESOL.
- Davidson, P. & Hobbs, A. (2003). Using academic discussions to assess higher order speaking skills. In S. Phipps (Ed.), *Proceedings of the 8<sup>th</sup> Bilkent University School of English Language ELT Conference: Speaking in the Monolingual Classroom* (pp. 200-216). Bilkent University.
- Ferdig, R. E., Baumgartner, E., Hartshorne, R., Kaplan-Rakowski, R., & Mouza, C. (Eds.) (2020). Teaching, technology, and teacher education during the COVID-19 pandemic: Stories from the field. *Waynesville: Association for the Advancement of Computing in Education (AACE)*. <https://bit.ly/3rUOvNh>
- Fernández-Costales, A. (2017). Subtitling in CLIL: Promoting Bilingual Methodologies through

- Audiovisual Translation. In María Elena Gómez Parra & Richard Johnstone (eds), *Educación bilingüe: tendencias educativas y conceptos claves/ Bilingual Education: Trends and Key Concepts* (pp. 185–196). Ministerio de Educación.
- Frumuselu, A. D., De Maeyer, S., Donche, V. & Gutiérrez-Colon-Plana, M. M. (2015). Television Series Inside the EFL Classroom: Bridging the Gap Between Teaching and Learning Informal Language through Subtitles. *Linguistics and Education*, 32(B), 107–117.
- Gebril, A & Plakans, L. (2013). Toward a Transparent Construct of Reading-to-Write Tasks: The Interface Between Discourse Features and Proficiency. *Language Assessment Quarterly*, 10, 1, 9-27. <https://doi.org/10.1080/15434303.2011.642040>
- Gómez Pérez, C. & Roales Ruiz, A. (2015). Propuestas didácticas de la aplicación LRN-2Sub a la clase de TAV alemán-español. En M. A. Recio Ariza, M. B. Santana López, M. de la Cruz Recio, P. Zimmermann González (coord.) *Interacciones: reflexiones en torno a la Traducción e Interpretación del/ al alemán*. (pp. 47-64). Peter Lang.
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020, March 27). The difference between emergency remote

- teaching and online teaching. *Educause Review*. <https://bit.ly/3r3xWQ0>
- Horowitz, D. M. (1986). What professors actually require: Academic tasks for the ESL classroom. *TESOL Quarterly*, 20, 445-460.
- Lertola, J. (2019). Second Language Vocabulary Learning through Subtitling. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, 32(2), 486-514.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22, 140-155.
- Moore, T., & Morton, J. (2005). Dimensions of difference: A comparison of university writing and IELTS writing. *Journal of English for Academic Purposes*, 4, 43-66.
- Navarrete, M. (2013). El doblaje como herramienta de aprendizaje en el aula de español y desde el entorno Clipflair. *MarcoELE*, 16, 75–87.
- Navarrete Ramírez, M. (2021). *Active audio description as a didactic resource to improve oral skills in foreign language learning*. (Doctoral Thesis). Universidad Nacional de Educación a Distancia. Madrid. <https://bit.ly/3s1v45I>
- North, B., & Piccardo, E. (2016). Developing illustrative descriptors of aspects of mediation for the Common European Framework of Reference (CEFR): A Council of Europe project. *Language Teaching*, 49(3), 455-459.

- Plakans, L. M., & Gebril, A. (2012). A close investigation into source use in L2 integrated writing tasks. *Assessing Writing*, 17(1), 18-34.
- Sánchez-Requena, A. (2016). Audiovisual Translation in Teaching Foreign Languages: Contributions of Dubbing to Develop Fluency and Pronunciation in Spontaneous Conversations. *Porta Linguarum*, 26, 9–21.
- Sawaki, Y., Thomas Quinlan & Yong-Won Lee (2013). Understanding Learner Strengths and Weaknesses: Assessing Performance on an Integrated Writing Task. *Language Assessment Quarterly*, 10, 1, 73-95.  
<https://doi.org/10.1080/15434303.2011.633305>
- Shaw, S., & Weir, C.J. (2007). Examining writing: Research and practice in assessing second language writing. *Studies in Language Testing*, 26. Cambridge University Press and Cambridge ESOL.
- Strickland, D.L. (2015). *Student Success Program Assessment Focusing on the Impact of the Student Success Program on Retention*. <https://bit.ly/3ucenXA>
- Talaván, N. (2013). *La subtitulación en el aprendizaje de lenguas extranjeras*. Octaedro.
- Talaván, N. & Lertola, J. (2016). Active Audiodescription to Promote Speaking

Skills in Online Environments. *Sintagma*, 28, 59–74.

- Talaván, N.; Lertola, J. & Costal, T. (2016). iCap: Intralingual Captioning for Writing and Vocabulary Enhancement. *Alicante Journal of English Studies*, 29, 229–247.
- Talaván, N. & Rodríguez-Arancón, P. (2014). The Use of Interlingual Subtitling to Improve Listening Comprehension Skills in Advanced EFL Students. In Beatrice Garzelli & Michela Baldo (eds). *Subtitling and Intercultural Communication. European Languages and Beyond*. Pisa: InterLinguistica ETS, pp. 273–288.
- Talaván, N. & Rodríguez-Arancón, P. (2019). Voice-over to improve oral production skills. En J. D. Sanderson y C. Botella Tejera (Eds.), *Focusing on Audiovisual Translation Research* (pp. 211–236). UPV. Publicacions Universitat de Valencia.
- Torralba, G. (2020). El uso de subtítulos pasivos y activos en la enseñanza de lenguas extranjeras: una revisión de las posibilidades didácticas de ambos recursos. *Íkala Revista de Lenguaje y Cultura*, 25(1), 231-250. <https://doi.org/10.17533/udea.ikala.v25n01a11>
- Tunkay, H. (2014). An integrated skills approach using feature movies in EFL at tertiary level. *Turkish Online Journal of Educational Technology* 13(1), 56-63.

- United Nations Educational, Scientific and Cultural Organization (2020, December 27). #LearningNeverStops. <https://bit.ly/32BOF3v>
- Weigle, S. C. (2004). Integrating reading and writing in a competency test for non-native speakers of English. *Assessing Writing*, 9(1), 27–55.  
<https://doi.org/10.1016/j.asw.2004.01.002>
- Weir, C.J. (1983). The Associated Examining Board's test of English for academic Purposes: An exercise in content validation. In A. Hughes & D. Porter (Eds.). *Current Developments in Language Testing*. Academic Press, (pp. 147-153).