Evaluating the effectiveness of ORCIT's note-taking resource

Dalia Mankauskienė

Vilnius University
Lithuania

Svetlana Carsten

University of Leeds
United Kingdom

Dalia Mankauskienė: Institute for Literary, Cultural and Translation Studies, Faculty of Philology, Vilnius University, Lithuania. Orcid: 0000-0003-2806-0892. | E-mail: dalia.mankauskiene@flf.vu.lt
Svetlana Carsten: Centre for Translation Studies, School of Languages, Cultures and Societies, University of Leeds, United Kigdom. Orcid: 0000-0001-9808-4254. | E-mail: s.carsten@leeds.ac.uk

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

40

Abstract

The paper presents a study conducted over several weeks with 22 undergraduate fourth year Translation BA students taking the compulsory Consecutive Interpreting module. The aim of the study was to ascertain whether the use of the ORCIT note-taking resource had an impact on learner knowledge and note-taking skills. The study followed the relevant parameters of the five-level holistic TEL evaluation framework by Pickering and Joynes (2016). The results did not show a statistically significantly better performance of the ORCIT group versus the control group during the trial but, comparing the pre-trial interpreting performance of students with their interpretation during the trial, the ORCIT group showed a statistically significant improvement, while the control group did not. This suggests a possible beneficial effect of the note-taking resource in terms of transfer of learning (learner gain) rather than knowledge retention.

Keywords: interpreter training; TEL; resource evaluation; ORCIT.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

41

1. Introduction

Technology-enhanced learning has been well-established in world education thanks to breathtaking developments in information technology in the late 1980s and all through the 1990s. Distance learning became one of the flagships of higher education early this century, attracting thousands of people to open online courses known as MOOCS (Massive Open Online Courses). E-learning, online or virtual learning, went hand in hand with 'face-to-face' traditional learning and there was no turning back. Translation and interpreting studies were in the forefront of these developments. These two disciplines embraced technology with amazing zeal, deploying AI to aid human translators and interpreters in their demanding and often stressful work. At the threshold of this century, universities across the world adopted information and digital technology in their training of the translators and interpreters of the future. In the field of interpreting, technology-enhanced training has become an indispensable aid to the well-established and universally recognised interpreter training pedagogy (Sandrelli and De Manuel Jerez, 2007; Class and Moser-Mercer, 2013; Sandrelli, 2015; Carsten, 2015). In the last two decades, there has been a proliferation of digital platforms and online resources offering a multitude of digital affordances for both trainers and trainees.

However, there has been a noted scarcity of researched and reported evaluation of the effectiveness of such resources on learners (Bradin, 1999; Lamy and Goodfellow, 1999; Fiori, 2005; Chapelle, 2008; Pickering and Joynes, 2016). This is not unique to interpreting. Although some qualitative, survey-based, evaluation, or quantitative Google Analytics data became available as early as the 1980s, the difficulty of evaluating the effectiveness of CAL tools remains pertinent. Academics have argued that objective, empirical, evaluation of online resources was well overdue (Pickering and Joynes, 2016). Trials started emerging in the 2000s with some small-scale pre- and post-testing evaluations conducted in the US (Issa and others, 2011; Chen and Aimee, 2013; Cook and Ellaway, 2015). One such evaluation (Issa and others, 2011) was taken on board by UK researchers working in the medical field (Pickering and Joynes, 2016). The UK team went beyond a small-scale trial and proposed an ambitious five-level holistic TEL evaluation framework, which will be addressed in chapter 2.1.

The focus in this paper is on a) an online interactive pedagogical tool ORCIT (Online Resources for Conference Interpreter Training) developed for interpreter-trainers and students of interpreting, and b) an empirical evaluation of the effectiveness of a key ORCIT interpreting resource on note-taking, using a pre- and post-testing method. In the first instance, we will give a brief overview of ORCIT, its aim and target audience; we will, then, comment on the results of ORCIT's earlier qualitative and quantitative evaluation, including the relevance of the five-level holistic TEL evaluation framework; we will discuss the relevance of the framework's *level 1B* evaluation; and, finally, we will move on to the methodology and the results in the trial of the note-taking resource, in order to ascertain whether there has been an impact of the resource on the learner.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

42

2. ORCIT: Online resources for conference interpreter training

The ORCIT project ran from October 2010 to September 2018, resulting in a set of 130 open-access, multimedia, multilingual and interactive learning resources for interpreter training. It was funded by DG SCIC (European Commission) within its broad programme *Grants for projects aimed at supporting conference interpreter training.* The project brought together partners from nine EU universities offering courses in interpreter training[1]. Its aims included:

- Consolidation of conference interpreting skills through a technology-enhanced environment (the focus was on five main interpreting competences: listening and analysis, mastery of the mother tongue and public speaking, early and advanced consecutive interpreting, early and advanced simultaneous interpreting, research skills).
- Demonstration of pedagogical best practice for trainers.
- Enhancement of students' flexible learning experience, including guided and independent approaches.

The ORCIT website (www.orcit.eu) has eight webpages in eight European languages. The original language of the resources, English, was subsequently localised into Czech, French, German, Greek, Lithuanian, Slovenian and Spanish, and these appear in a bookshelf format, offering guidance to users as well as interactive demo exercises. ORCIT is not a stand-alone training tool. It complements conference interpreting courses and follows the standards set by the European schools with interpreting training provision. ORCIT has a specific target audience, i.e., trainers and trainee interpreters, and offers a clear idea of its pedagogic content and how it should be presented[2].

2.1. ORCIT: previous evaluation

A necessary pre-requisite for any project funding is to determine its 'value for money.' Therefore, evaluation of the output is a natural conclusion to a project. In the case of ORCIT, evaluation of the created resources was systematic and incremental and was carried out at the end of each project year, when a set of completed resources was released on an open-access basis. Initial collection of data took place within the institutions participating in the project and the results were reported to the funder annually, which was a require-

---

1   The partners included the University of Leeds (UK), Vilnius University (Lithuania), Charles University, Prague (Czech Republic), Aristotle University, Thessaloniki (Greece), Comillas Pontifical University (Spain), La Laguna University, Tenerife (Spain), ISIT, Paris (France), Heidelberg University (Germany), and the University of Ljubljana (Slovenia).

2   See the link ORCIT EXPLAINED in www.orcit.eu.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

43

ment stipulated by the grant application. The initial success indicator was in the form of anecdotal evidence from the tutors using ORCIT resources in the classroom or from their students. The ORCIT dissemination campaign also helped to collect evidence from non-participating universities where the resources were used[3]. In addition, quantitative evaluation through Google Analytics was carried out on an annual basis and served as a further indicator of the popularity of ORCIT resources.

Upon completion of the full set of English resources in 2016, the evaluation questionnaire was devised and uploaded onto the English page of the website. It was, subsequently, localised into seven other languages and became available in 2018 on each relevant page of the ORCIT website. The evaluation questionnaire allowed users to comment on relevance of the resources to their curricula, their value as an aid to acquiring and developing the skills, quality of instruction and navigation, and overall usefulness. Furthermore, the questions about the user status (student, trainer, other, e.g., practicing professional) and the context in which resources were used (classroom, self-study, etc.) offer important pedagogical evidence of "how these resources are used, and by whom" (McGill, 2010/2014).

A large-scale qualitative and quantitative evaluation of ORCIT resources took place after the completion of the project at the end of 2018 with the results published in *The Interpreter and Translator Trainer* 15 (4) in December 2021 (490-505). The conclusions were very encouraging as these demonstrated significant world popularity of ORCIT resources in the open questionnaire responses and registered impressive 16,228 hosted sessions on its website.

In the 2018 study the authors considered a number of approaches to evaluation. For quantitative evaluation these included content analysis (CA) methods and learning analytics data mining. For qualitative evaluation, JISC[4]/HE Academy 2009-2012 holistic approach to evaluate the impact of OERs (Open Educational Resources) was adopted. Finally, three out of five relevant parameters of the five-level holistic TEL evaluation framework (Pickering and Joynes, 2016) were applied to ORCIT resources and discussed in detail in Carsten and others (2021). Briefly, the five levels of the evaluation framework include:

---

3   To give an example of comments about ORCIT resources that comprised part of the report to SCIC in 2012-2013: 'ORCIT materials are an excellent supplement to in-class exercises and discussions' (Monterey Institute of International Studies). 'The ORCIT resource explains simply and clearly some of the key areas of interpreting in a user friendly, bookshelf format that is as welcoming as it is easy to use' (English A trainer, ISIT). 'I recommend the ORCIT resources to all my conference interpreting students and incorporate the recordings and exercises into the "homework" I ask them to do' (English A trainer, University of Manchester, UK).

4   JISC is the UK higher education, further education and skills sector's not-for-profit organisation for digital services and solutions: https://www.jisc.ac.uk/.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

44

- *Level 0*: Preliminary evaluation of need (to ensure TEL is not introduced for its own sake)
- *Level 1A*: Learner satisfaction (cost-utility)
- *Level 1B*: Learner gain (cost-effectiveness)
- *Level 2*: Learner impact (cost-benefit)
- *Level 3*: Institutional impact (cost-feasibility).

Evaluation of need (*level 0*) and learner satisfaction (*Level 1A*) were included in the above study and a detailed methodology for evaluation at *level 1B* was proposed for the next stage of empirical trial. *Level 1B* addresses *learner gain* indicators—i.e., how a particular technology-enhanced resource compares with existing resources or traditional methods of teaching.

The trial to assess learner gain effectiveness of ORCIT was planned to be launched in the 2020-2021 academic year, but the Covid pandemic prevented the project team from going ahead with it. One of the prerequisites for *level 1B* evaluation was to compare online learning with that of face-to-face learning. Therefore, it was not possible to carry out any experiment during the restrictions imposed by universities on classroom learning. Furthermore, the original ambitious plans for a multi-institutional trial had to be postponed for an indefinite period of time as the pandemic had a detrimental impact on student numbers on Master's programmes. So as not to lose the sight of the popularity and effectiveness of ORCIT resources, the project team has decided to run a trial with a cohort of final-year undergraduate students of Translation at Vilnius University. This is a small-scale pilot project which focuses on the acquisition of the note-taking skill—one out of five skills necessary for developing sound interpreting competence. The authors hope that this pilot trial will pave the way to a bigger study of the effectiveness of the range of ORCIT resources.

2.2. Note-taking resource: pre-test baseline for participants

The methodology used in this study follows the proposed pre- and post-testing approach for *level 1B* evaluation as outlined in Carsten and others (2021: 502-503) and discussed in the following chapter. *Level 1B* addresses *learner gain* indicators—i.e., how this particular resource compares with existing resources or traditional methods of teaching. If the evaluation were to be carried out with postgraduate students on Master's programmes, the pre-test baseline would be given because the students would have the same level of ability right from the beginning. In the 2021 publication it was explained that:

> the students will have undergone rigorous selection for entry to this type of programme to determine their aptitude; they will hold a university degree or equivalent; have an excellent command of their mother tongue and an in-depth knowledge of their working languages; have a good knowledge base in a number of disciplines; demonstrate an aptitude in concentration, analysis and synthesis skills; be good communicators, etc. (502).

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

45

With an undergraduate cohort of students as in Vilnius, a different approach to the trial had to be adopted to arrive at a uniform baseline. As explained in the methodology chapter, the tested cohort of students of the presented research were in their final year of study. They had reached the same level of attainment in their disciplines by virtue of following the same programme and the necessary assessment for progression to the final stage of their degree. However, knowledge of the set of disciplines that students have studied, as opposed to acquisition of interpreting skills in their final year, should be given special consideration.

Moser-Mercer in her 2008 publication *Skill Acquisition in Interpreting: A Human Performance Perspective* ascertains that performance is central to interpreting. She compares it to performance in sports and acting as an example (2008: 1). Moser-Mercer maintains that although "performance psychology draws heavily on cognitive psychology for the theoretical description of human performance…, performance psychology lends itself much more to direct practical application" (2008: 2). Therefore, the focus in this study is very much on practical application and performance in one interpreting skill, note-taking.

In the absence of aptitude testing for performance ability of undergraduate students, the authors of this study decided to allow training in a set of early interpreting skills to run its course first. This normally happens in the first half of the first semester. Interpreting training progresses according to the scaffolding principle, i.e., "breaking down of the interpreting process into sub-tasks, which would be tackled one by one and ultimately be brought together again when performing the whole task" (see Moser-Mercer, 2008: 14). The early skills for Vilnius students are those of listening and analysis, public speaking and consecutive without notes. Only after studying and practicing these early skills and having been assessed for performance, would the participants reach a uniform preparedness to take a test in a more advanced skill, that of note-taking, in the second half of the first semester. It is at that stage that they were invited to take part in the study outlined below.

3. Level 1B: methodology

3.1. Participants

In total, 22 Vilnius University students participated in the study: 3 male and 19 female, all but one were 22-23 years old. At the time they were all fourth year Translation BA students, taking the compulsory Consecutive Interpreting module led by one tutor. Before moving on to note-taking, all students had studied a compulsory Rhetoric module, similar to the Public Speaking module in the ORCIT set of resources. They had almost two months of training on the early interpreting skills, those of listening and analysis and early consecutive. Thus, the participants' aptitude for the whole group was comparable, with marginal deviation, to that of the beginner interpreting trainee. It has to be borne in mind that, in general, the Consecutive Interpreting course follows the same pedagogical approach as delineated in ORCIT,

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

46

in terms of the sequence when introducing certain skills at a certain time, i.e., applying the principle of "scaffolding". Although it was not possible to completely block access to the resources due to its open-access nature, students were not encouraged to use ORCIT by themselves before or during the time of the trial.

According to the course requirement, for the Consecutive Interpreting module the students are divided into two groups and have separate seminars, one after another, to facilitate small-group interpreting practice. One of these groups was used as the trial group (further referred to as 'ORCIT group') and the other one as 'control group'. Each group consisted of 11 students who had similar levels of prior knowledge and preparedness. However, as the study took place in three different sessions (28/10, 04/11, 11/11), not all students participated in all sessions. Their participation and sample size per session are presented in table 1. Another session was conducted with one of the exercises on 18/11 (note-taking from an article), but the student attendance was very low, therefore it was excluded from the analysis.

TABLE 1

Number of participants per session (1 indicating that the person was in the class)

STUDENT	28/10	04/11	11/11	STUDENT	28/10	04/11	11/11
Stud_O1 (F)	1	1	1	Stud_C1 (F)	1	1	1
Stud_O2 (F)	1	1	1	Stud_C2 (F)	1	1	1
Stud_O3 (F)	1	1	1	Stud_C3 (F)	1	1	1
Stud_O4 (F)	0	1	1	Stud_C4 (F)	1	1	1
Stud_O5 (F)	1	1	1	Stud_C5 (F)	1	0	1
Stud_O6 (M)	1	1	1	Stud_C6 (F)	1	1	1
Stud_O7 (M)	1	0	1	Stud_C7 (F)	1	0	1
Stud_O8 (F)	1	0	1	Stud_C8 (F)	1	0	0
Stud_O9 (F)	0	1	0	Stud_C9 (F)	1	0	0
Stud_O10 (F)	0	1	0	Stud_C10 (F)	1	1	0
Stud_O11 (M)	0	1	1	Stud_C11 (F)	1	0	1
Total_ORCIT	**7**	**9**	**9**	**Total_Control**	**11**	**6**	**8**

Although the study was conducted during regular class time, all participants were given the option of not participating. Yet after they were explained the purpose of the trial and allowed to ask questions, all of them chose to participate and signed the informed consent form. The students did not receive any financial remuneration for their participation.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

47

3.2. Procedure

Before the trial for note-taking and during the last two weeks of October 2022, the students' performance was observed in the regular class to establish the pre-trial baseline of their interpreting skills acquisition. This was done as part of their regular assessment and would have happened irrespective of this study. The assessment consisted of interpreting tasks of short segments and at least one speech longer than 3 minutes (all done without taking notes).

As has been mentioned earlier, the data for this study was collected during three separate sessions coinciding with the regular Consecutive Interpreting classes during the period relevant to the introduction of note-taking skills. In these classes other activities pertaining to the regular module were also taking place in order to maximise the benefits of the pedagogical principles of reinforcement, repetition and redundancy. On 07/11/2022 (between the second and third data collection sessions) all students had a lecture together for the introduction of basic symbols and abbreviations as explained in the ORCIT Exercises on note-taking section 'Build your own glossary of symbols'. Due to national holidays in Lithuania, there was no lecture on 31/10/2022 (i.e., between the first and the second data collection sessions).

The content for both groups was identical, although one was presented through ORCIT; the other, by the tutor. It is important to note that, after the data had been collected, both the ORCIT and the control groups were encouraged to use the resources at their leisure.

The topics for the trial were chosen following the pattern of the ORCIT resources, namely Introduction to Note-Taking and Exercises on Note-taking. For the more theoretical parts of the resources an open-ended question format was chosen to allow students to express freely what they remembered instead of prompting them through multiple-choice questions. To this end, the participants were offered an 'Introduction to note-taking' questionnaire and 'Exercises on note-taking: step-by-step, part one' respectively (see table 2 and appendices 1 & 2). All questions focused on the retention of knowledge and the interpreting-specific metalanguage, while the practical interpreting tasks in 'Exercises on note-taking: step-by-step, part two' (see table 2, 11/11/2022, and appendix 2, part 2) served as the basis for evaluating the transfer of learning, or learner gain (Pickering and Joynes, 2016).

Almost two months after the initial trial the students were asked to fill in the final questionnaire, which was made up of two parts:

1. Testing their information and metalanguage retention by asking some of the same questions as in Intro and Exercises questionnaires. Certain questions were excluded from the post-trial questionnaire because they would have been difficult to understand outside the context (see appendix 3);

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

48

2. Overall feedback on ORCIT resources which was based on the evaluation form available on the ORCIT website[5] (see appendix 4).

One of the concerns of measuring the performance on repeated tasks is the carryover effect which can happen between the measurements, i.e., if participants are being asked the same questions, they will learn the answers, but not necessarily have a better understanding and, thus, show improvement merely due to completion of the same task. A few steps were taken in order to minimise the carryover effects for information-retention testing: there was a two-month gap between measuring performance on the same questions, no provisional feedback was given to the students after the trial questionnaire, the participants were not told during the trial they would be tested again with the same questions.

TABLE 2

Data collection for the study

DATE	DATA SET NAME	ORCIT RESOURCE UNDER STUDY	DATA OUTCOME
October 2022	Pre-trial		Assessment of pre-trial student performance
28/10/2022	Intro	Introduction to Note-Taking[6]	Answers to 8 open-ended questions were given in writing right after hearing the information either from ORCIT (the ORCIT group) or from tutor (the control group) (see appendix 1 for the evaluation form with the questions)
04/11/2022	Exercises	Exercises on Note-taking[7]: Step-by-Step, Part One	Answers to 7 open ended questions were given in writing right after hearing the information either from ORCIT (the ORCIT group) or from tutor (the control group) (see appendix 2 for the evaluation form with the questions)
11/11/2022	Speech	Exercises on Note-taking: Step-by-Step, Part Two	Interpretation of a speech
First week of 2023	Post-trial evaluation		1. Answers to some of the same questions as in Intro and Exercises (see appendix 3).
	Satisfaction assessment		2. Overall assessment of the satisfaction of using ORCIT (see appendix 4).

---

5   See https://docs.google.com/forms/d/e/1FAIpQLSd4HJIyaWGIdWFj7Jf4ppab64w0hQ5HJS5Wdw0kHvE_wSWqRg/viewform for reference.

6   See https://orcit.eu/resources/nti-en/story_html5.html for reference.

7   See https://orcit.eu/resources/nte-en/story_html5.html for reference.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

49

The data then were evaluated according to the percentage of information retained and were given a mark from 1 to 5 in the following manner:

1 Very poor (20-39 percentage remembered)	2 Poor (40-49 percentage remembered)	3 Pass (50-59 percentage remembered)	4 Good (60-69 percentage remembered)	5 Very good (70-79 percentage remembered)

Answers to each question in both questionnaires were assessed separately and then the average per student was calculated, which was later used in the other calculations. The results of this evaluation are presented in the next section.

3.3. Results

The analysis of the trial data was performed on all three sets of data separately (due to different sample sizes) and on the aggregated data set (referred to as 'Overall'). In the Overall dataset, the missing data entries were replaced with the average of the group (i.e., if a student had not filled in the first questionnaire, his/her answers to a specific question were replaced with the average of the ORCIT group's response to that particular question). The descriptive statistics of these datasets can be seen in table 3.

TABLE 3

Descriptive statistics of all datasets by group

	ORCIT GROUP			CONTROL GROUP			TOTAL					
	n	Mean	SD*	n	Mean	SD	n	Mean	SD	t	p	g
Intro	7	4.39	0.41	11	4.31	0.38	18	4.34	0.38	0.45	0.66	0.21
Exercises	9	4.17	0.16	7	4.14	0.26	16	4.16	0.20	0.28	0.78	0.14
Speech	9	4.33	0.71	8	4.25	0.71	17	4.29	0.69	0.24	0.81	0.11
Overall	11	4.29	0.20	11	4.23	0.27	22	4.26	0.23	0.60	0.55	0.24

*SD (standard deviation)

Welch's t-test was performed to determine if there was a statistically significant difference in the results of a questionnaire or interpretation between the two groups. The test revealed that there was no statistically significant difference in mean scores between the two groups in any of the datasets, which means that there is no statistical evidence that the difference in groups is not due to chance. However, as has been suggested earlier in this paper, the

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

50

number of students on interpreting courses has been steadily decreasing, so bigger samples are difficult to come by without cooperation in such a study of several institutions.

Another statistical measurement was calculated, which arguably is a more important one in this study due to its more practical application—the effect size. This statistic shows how meaningful is the relationship between variables or the difference between the groups. It indicates the practical significance of a research outcome. A large effect size means that a finding has practical significance, while a small effect size indicates limited practical applications. The sample size under study is quite small, therefore Hedges' g measure was chosen and calculated for all data sets. As a general guideline, Cohen (1988) suggested that an effect size of 0.2 should be considered a small effect; size 0.5, medium, and 0.8, large. As can be seen from table 3, the effect size of the Intro dataset and the Overall data set can be considered small (0.21 and 0.24, respectively), yet not negligible.

On the other hand, it is probably not realistic to expect one resource out of many to have a large impact, as acknowledged by Pickering and Joynes: "being able to realistically quantify the impact of one single resource within a multi-faceted setting is both reductionist and extremely difficult because it is unrealistic to assume that a student would be able to progress through their course with only one package of information influencing their learning" (2016: 1245). Moreover, the course itself follows the pattern of ORCIT quite closely, so the change in the way information is presented might not have had such a big impact as it might have had in a different context and perhaps with a different methodology.

Apart from the overall analysis of the datasets, we may also look at the individual answers to the questions in the Intro and the Exercises questionnaires (table 4 and table 5, respectively). As can be seen from the tables below, none of the differences are statistically significant. However, certain observations still can be made based on the kind of questions one group answered more precisely than the other.

The answers to the third Intro question—"what is the difference between active and passive listening"—were surprising, because active and passive listening is a topic that was discussed within the module a few times before the trial as it is an important part of consecutive without notes. Yet students in the ORCIT group were able to answer this question much better than those in the control group.

Question 7 from the Intro dataset and questions 1, 4 and 6 from the Exercises dataset are very similar—whether/why you should not write down as much as possible, whether/why your notes should not be dense, etc.— The answers that the same students gave to all of these questions are not consistent, i.e., the same student might have recalled the majority of information in answering three out of four of these questions but was not able to remember as much answering the fourth question. So, it is not the case that the student does not understand the concept of notes having to be brief, visual, etc., but rather that certain points were

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

51

not as clearly perceived as others. However, there is no one specific question that the majority of students failed to answer, so specific reasons for such a situation cannot be named.

TABLE 4

Descriptive statistics of Intro dataset questionnaire per question

QUESTION	ORCIT GROUP (7)		CONTROL GROUP (11)		TOTAL (18)		t	p
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**		
1	4.14	0.90	4.73	0.47	4.50	0.71	-1.59	0.15
2	4.57	0.79	4.73	0.65	4.67	0.69	-0.44	0.67
3	4.86	0.38	4.45	0.69	4.61	0.61	1.60	0.13
4	4.71	0.49	4.82	0.40	4.78	0.43	-0.47	0.65
5	4.00	0.82	3.55	1.04	3.72	0.96	1.04	0.32
6	4.00	1.00	4.00	0.77	4.00	0.84	-	1.00
7	4.43	0.79	4.09	0.30	4.22	0.55	1.09	0.31
8	4.43	1.13	4.09	1.30	4.22	1.22	0.58	0.57

TABLE 5

Descriptive statistics of exercise dataset questionnaire per question

QUESTION	ORCIT GROUP (9)		CONTROL GROUP (7)		TOTAL (16)		t	p
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**		
1	3.56	0.88	4.14	0.69	3.81	0.83	-1.49	0.16
2	3.67	0.71	4.00	0.58	3.81	0.66	-1.04	0.32
3	5.00	-	4.86	0.38	4.94	0.25	1.00	0.36
4	3.67	0.87	3.86	0.38	3.75	0.68	-0.59	0.57
5	4.22	1.20	3.86	0.38	4.06	0.93	0.86	0.41
6	5.00	-	4.43	0.79	4.75	0.58	1.92	0.10
7	4.11	0.78	3.86	0.38	4.00	0.63	0.85	0.41

Comparing pre-trial student performance with their interpreting with notes results during the trial, we can see that there is a statistically significant ($p < 0.05$) overall improvement (see table 6). However, the improvement in the control group is not statistically significant

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

52

(p = 0.0796), i.e., their performance during the trial is similar to their performance prior to it. Meanwhile, in the ORCIT group the improvement is statistically significant (p = 0.0353), indicating that these students were able to utilize the new knowledge on note-taking more effectively. The effect size for both groups exceeds 0.5, which means that the effect size of training that happened between the pre-trial assessment and the speech interpretation with note-taking (i.e., learner gain) during the trial is medium (cf. Cohen, 1988). Although the impact of cumulative benefit from other elements of training cannot be ruled out as most of the training in this period was on note-taking resulting in improved performance.

TABLE 6

Descriptive statistics comparing the pre-trial performance with speech interpretation with notes during the trial by group

	PRE-TRIAL		SPEECH		t-TEST PAIRED		g
	Mean	SD	Mean	SD	t	p	
ORCIT group (n=9)	3.89	0.78	4.33	0.71	2.30	0.0353	0.57
Control group (n=8)	3.88	0.64	4.25	0.71	2.36	0.0796	0.53
Total (n=17[8])	3.88	0.70	4.29	0.69	2.12	0.0041	0.58

This is not fully consistent with the study carried out by Issa and others on medical students (2011), where the control group followed the traditional lectures using the traditional slide design and the trial group followed the material with a multimedia slide design based on Mayer's multimedia design principles (Mayer, 2005). The researchers pointed out that "applying multimedia principles when designing slides for medical students results in greater improvement in student knowledge retention compared with traditional slide design" (Issa and others, 2011:824). However, they also did not find statistically significant improvement in transfer of learning and concluded that "empirical research in still needed to determine how these principles affect transfer of learning" (Issa and others, 2011: 818). The instructional design behind the ORCIT resources is based on Mayer's multimedia design principles which are discussed in detail in Carsten and others (2021: 5-6). It is, therefore, encouraging to see an improved transfer of learning in the ORCIT group compared to that of the control group, although the knowledge retention was not found to be statistically significant in this study.

The post-trial evaluation questionnaire showed that both groups were able to answer the same questions statistically significantly better two months after the trial than during the

---

8   As 17 students interpreted the speech during the study, only the performance of these 17 students is compared with their pre-trial assessment.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

53

trial (t-Test paired $p < 0.05$ for ORCIT group (0.016), control group (0.009) and total (0.0003)). It shows that their understanding of note-taking process and concepts have improved in the two months between the trial and the post-trial evaluation. However, repeated-measures ANOVA analysis did not show the effect of the group to be statistically significant. This might be because after the trial all students were allowed to use ORCIT or perhaps due to general practice during seminars which also took place after the trial.

Finally, based on the answers to the post-trial ORCIT evaluation form, the following results can be observed with regard to overall satisfaction:

- All students noted that ORCIT resources were relevant to their curriculum, stating that they gave "a very detailed information using a clear structure" and "it was great for practice; the examples and tips given were easy to understand and were useful" and everyone agreed with the statement "overall I found the resource to be useful".

- The following answers were not mutually exclusive. Most of the respondents, unsurprisingly, noted that they used ORCIT in class: 85.7 % to do the interactive exercises; 76.2 % when introducing a new skill; 66.7 % for revision. The main reason why the students used the resources outside the classroom was for revision (52.4 %).

- All students agreed that the demonstrations are helpful, while all but one said that the explanations are easily understood. 6 students stated that they "neither agreed nor disagreed" with the statement that they liked the way the material was presented and with the statement that "this resource has consolidated my skills in this area", while the others agreed with these statements. Yet the biggest difference was in assessing the statement "I found the material motivating"—10 students agreed with the statement, 11 neither agreed nor disagreed and 1 disagreed.

- All students noted that the exercises were a good way to practice note-taking, mentioning that speeches were short and the various topics discussed were good for practice, "they provide many examples and useful tips for beginners" and the resource was useful because "the exercises and training material were in one place".

- Evaluating the note-taking resource, 19 % of the respondents rated it as a 5 (very useful), while 76.2 % gave it 4 out of 5, in the comments calling for "more practice for interpretation" and stating that "it would have been more useful if it had more exercises, more speeches to practice".

4. Conclusions

The aim of the study was to ascertain whether the use of the ORCIT resource, note-taking, had had an enhanced effect on information retention and transfer of learning, if com-

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

54

pared to the traditional tutor-led form of instruction. As a result of this study, we can conclude that the statistical evaluation of student performance did not reveal statistically significant changes and that during the trial the ORCIT group did not outperform the control group in a statistically significant way. However, from the data collected we can also conclude that using ORCIT can be as effective as live tutoring, though we do not advocate the use of one over the other. Naturally, the overall training of the students should be borne in mind. At Vilnius University the approach to interpreting training is through blended learning where ORCIT resources are routinely and successfully used, especially outside the classroom setting.

The results of the interpreted speech with notes (learner gain) presented an interesting outcome though. Comparing the results of the pre-trial evaluation with the assessment of the interpreted speech during the trial, we found that the ORCIT group statistically significantly improved their performance, while the improvement of the performance of the control group was not statistically significant. This might suggest that the strength of ORCIT is perhaps not in improving information retention, but rather the transfer of learning, or impact on learner, in other words, which for interpreter-trainee is an important aspect. This perception is seen in the evaluation questionnaire where students called for more practice of interpretation and stated that the note-taking resource "would have been more useful if it had more exercises, more speeches to practice". However, we must reiterate that the impact of cumulative benefit from other elements of training cannot be ruled out as most of the training in this period was on note-taking resulting in improved performance.

With the benefits of ORCIT outlined above, and especially as an independent learning tool, we believe that ORCIT could be extremely useful for non-professional interpreters, working in the field of community interpreting, for example. That is a different avenue for ORCIT's use that should be explored further. Perhaps a similar study could be conducted with a group of non-professional interpreters, who are not enrolled in university programmes to assess the potential of ORCIT in different interpreting contexts.

5. References

Bradin, Claire, 1999: "CALL issues: Instructional aspects of software evaluation" in Joy Egbert and Elizabeth Hanson-Smith (eds.): *CALL environments: Research, practice, and critical issues*, Alexandria VA: TESOL Publications, 159-175.

Carsten, Svetlana, 2015: "Mapping a theoretical framework for ORCIT", *Vertimo Studijos* 8, 7-22.

Carsten, Svetlana, Dragoș Ciobanu and Dalia Mankauskienė, 2021: "The challenge of evaluating open interpreter training resources: case study of ORCIT", *The Interpreter and Translator Trainer* 15 (4), 490-505, doi: 10.1080/1750399X.2020.1867407.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service…: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

55

CHAPELLE, Carol, A., 2008: "Challenges in Evaluation of Innovation: Observations from Technology Research", *Innovation in Language Learning and Teaching* 1 (1), 30-45. doi: 10.2167/illt041.0.

CHEN, Baiyun, and Denoyelles AIMEE, 2013: "Exploring students' mobile learning practices in higher education" [http://er.educause.edu/articles/2013/10/exploring-students-mobile-learning-practices-in-higher-education, accessed 3 December 2022].

CLASS, Barbara, and Barbara MOSER-MERCER, 2013: "Training conference interpreter trainers with technology – a virtual reality" in Ollala GARCÍA BECERRA, Macarena PRADAS MACÍAS and Rafael BARRANCO-DROEGE (eds.): *Quality in Interpreting: Widening the Scope* 1, Granada: Editorial Comares S. L., 293-313.

COHEN, Jacob, 1988: *"Statistical power analysis for the behavioral sciences"* (2nd edition), Hillsdale, NJ: Lawrence Erlbaum Associates.

COOK, David A., and Rachel H. ELLAWAY, 2015: "Evaluating technology-enhanced learning: A comprehensive framework", *Medical Teacher* 37 (10), 961-970.

FIORI, Melissa L., 2005: "The development of grammatical competence through synchronous computer-mediated communication", *CALICO Journal* 22 (3), 567-602.

ISSA, Nabil, Mary SCHULLER, Susan SANTACATERINA, Michael SHAPIRO, Edward WANG, Richard E. MAYER and Debra A. DAROSA, 2011: "Applying Multimedia Design Principles Enhances Learning in Medical Education", *Medical Education* 45 (8), 818-826, doi: 10.1111/j.1365-2923.2011.03988.

LAMY, Marie-Noëlle, and Robin GOODFELLOW, 1999: "Reflective conversation' in the virtual language Classroom", *Language Learning & Technology* 2 (2), 43-61.

MAYER, Richard E. (ed.), 2005: *The Cambridge Handbook of Multimedia Learning*, Cambridge: Cambridge University Press.

MCGILL, Lou, 2010/2014: "Full Guide Open Educational Resources" [https://jisc.ac.uk/full-guide/open-educational-resources, accessed 3 December 2022].

MOSER-MERCER, Barbara, 2008: "Skill Acquisition in Interpreting", *The Interpreter and Translator Trainer* 2 (1), 1-28. doi: 10.1080/1750399X.2008.10798764.

PICKERING, James D., and Viktoria C. T. JOYNES, 2016: "A Holistic Model for Evaluating the Impact of Individual Technology-enhanced Learning Resources", *Medical Teacher* 38 (12), 1242-1247. doi: 10.1080/0142159X.2016.1210112.

SANDRELLI, Annalisa, 2015: "Becoming an interpreter: The role of computer technology", *MonTI* 2, 111-138.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

56

Sandrelli, Annalisa, and Jesús de Manuel Jerez, 2007: "The impact of information and communication technology on interpreter training", *The Interpreter and Translator Trainer* 1 (2), 269-303.

6. Appendices

6.1. Appendix 1: Introduction to note-taking questionnaire

1. Are core skills of interpreting related to note-taking only? Provide a brief comment.

2. Which are the core skills of interpreting? Name the ones you remember.

3. What is the difference between active and passive listening? Provide a brief comment.

4. How much do we need to focus on words? Provide a brief comment.

5. PITFALLS. Provide a brief description of these.

6. NUMBERS AND NAMES. Describe how you deal with these.

7. IDEAS. Is it a good idea to write as much as possible in your note pad while taking notes. Explain why yes or not.

8. MAIN PRINCIPLES OF NOTE TAKING. Name the ones you remember and comment on them.

6.2. Appendix 2: Exercises on note-taking questionnaire

1. Why don't we use short-hand or write down information like lecture notes? Provide a brief answer.

Listen to the speech and the discussion that follows and briefly answer the following questions:

2. What are the margins for?

3. What are the lines for?

4. Why should we avoid dense notes?

5. What purpose do arrows serve?

6. Why do we say less is more in note-taking?

7. What are the basic principles listed in this section? Name the ones you remember.

6.3. Appendix 3: Information retention questionnaire

1. What are the core skills of interpreting? Name the ones you remember.

2. What is the difference between passive and active listening? Provide brief comment.

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

57

3. How much do we need to focus on words when interpreting? Provide a brief comment.

4. What should your notes be like?

5. Why don't we use short-hand or write down information like lecture notes?

6. What are margins for?

7. What are lines for?

8. What purpose do arrows serve?

9. What are the basic principles of note-taking?

6.4. Appendix 4: ORCIT evaluation

10. Are ORCIT resources relevant to your curriculum?

11. Why/why not?

12. Why do you use the ORCIT Resources? Select all that apply.

☐ I have been advised to use ORCIT resources by my trainer

☐ Using ORCIT is compulsory part of my interpreting course

☐ I use ORCIT resources on my own initiative

13. How do you use the ORCIT Resources? Select all that apply.

☐ I use ORCIT resources outside of class when learning about a new skill

☐ I use ORCIT resources outside of class for skills consolidation

☐ I use ORCIT resources outside of class for revision

☐ I use ORCIT resources in class with my trainer when introducing a new skill

☐ I use ORCIT resources in class with my trainer to do the interactive exercises

☐ I use ORCIT resources in class with my trainer for revision

☐ I use ORCIT resources during group practice sessions with classmates

ONOMÁZEIN | Special Issue XIII - Present and future challenges for public service...: 39 - 58
Dalia Mankauskienė and Svetlana Carsten
Evaluating the effectiveness of ORCIT's note-taking resource

58

Note-taking resource evaluation

14. For each of the statements below, select which response matches your opinion on this resource:

	Agree	Neither agree nor disagree	Disagree
I found the resource easy to navigate	☐	☐	☐
I thought the explanations were easily understood	☐	☐	☐
I found the demonstrations helpful	☐	☐	☐
I liked the way the material was presented	☐	☐	☐
I found the material motivating	☐	☐	☐
I understand this area/skill better having completed this resource	☐	☐	☐
This resource has consolidated my skills in this area	☐	☐	☐
Overall I found this resource useful	☐	☐	☐

15. Are the exercises a good way to practice note-taking?

16. Why/why not?

17. How useful did you find this particular resource overall? (1 – not useful at all, 5 – very useful)

18. Please use this box to add any other comments you have about this resource.