Lingue e Linguaggi
Lingue Linguaggi 19 (2016), 23-40
ISSN 2239-0367, e-ISSN 2239-0359
DOI 10.1285/i22390359v19p23
http://siba-ese.unisalento.it, © 2016 Università del Salento

INVESTIGATING STUDENT CHOICES IN PERFORMING HIGHER-ORDER COMPREHENSION TASKS USING TED TALKS IN LEARNWEB

FRANCESCA BIANCHI[1], IVANA MARENZI[2]
[1]UNIVERSITÀ DEL SALENTO, [2]LEIBNIZ UNIVERSITÄT HANNOVER FORSCHUNGSZENTRUM L3S

Abstract - The aim of the current experiment was to test the teaching and research potential of interactive features of selection, deselection, tagging and logging in the analysis of reading-comprehension processes. To this aim, LearnWeb – an interactive platform integrating TED talks – was used to involve 25 Italian MA students of consecutive interpreting in analytical tasks gauging their reading-comprehension abilities in English. Their selections, deselections, and annotations were automatically collected by the system and manually analysed by the researchers. The analyses provided an answer to the following research questions: Was any of the tasks perceived as difficult by the students? How was each task faced by the students? How did the logs contribute to understanding the students' approaches to the tasks? The types of exercises used fit a large range of learning scenarios, and the resources, analytical methods and results described in this paper may be relevant to anyone interested in discourse comprehension.

Keywords: TED talks; LearnWeb; log analysis; text comprehension; processes

1. Introduction

The current paper describes a tentative experiment in the use of TED talks and the interactive learning platform LearnWeb in order to investigate student choices in performing analytical tasks that gauge discourse-comprehension abilities. The experiment took place within an MA module on consecutive interpreting – discourse comprehension being a fundamental skill in interpreting.

LearnWeb[1] (Marenzi, Zerr 2012) is a collaborative and multimedia educational platform. This platform is characterised by an iterative evaluation-driven design-based research approach (Marenzi 2014), i.e. an approach in which researchers and practitioners collaborate in real-world settings to cyclically assess and improve the functionalities of a system (Wang, Hannafin 2005, pp. 6), tailoring them to the practitioners' needs. LearnWeb has been used in several and varied learning scenarios (Bortoluzzi, Marenzi 2014; Marenzi, Kantz 2013; Holdack-Janssen, Marenzi 2012), but never before in interpreting courses.

LearnWeb has recently integrated a range of multimedia open resources (Taibi *et al*. 2015), including TED talks,[2] an open set of video conferences that is showing great potential as educational resource. Indeed, as several authors have observed (e.g.: Park, Cha 2013; Bortoluzzi, Marenzi 2013), TED talks expose students to authentic spoken material that is engaging in terms of form, as well as content; furthermore, they present a variety of topics (i.e., vocabulary), and accents; finally the videos are accompanied by transcripts and multilingual translations. These features make of them a flexible type of material.

[1] http://learnweb.l3s.uni-hannover.de
[2] http://ted.org/

TED talks are gaining momentum as learning as well as research material in the area of language studies, including translation and interpreting. Aston and Rodi (2012) resorted to TED talks for the creation of an aligned corpus of spoken discourse to be used with trainee conference interpreters. Shimizu *et al.* (2014) chose to include TED talks in their corpus of simultaneous translations because of their format and breadth, which made them suitable for analysing broad-coverage speech translation. Finally, Sung (2014) used TED videos to raise the students' listening-comprehension skills in undergraduate interpreting classes.

Accessing TED talks in LearnWeb offers teachers several advantages, as this platform provides interactive features designed to support learning tasks. These features are detailed in Section 3.4.

Within the general aim of finding possible uses of the platform in interpreting courses, the experiment was specifically designed to see if and to what extent one of its unique features, i.e. the automatic logging of the actions performed by students when working on TED transcripts, could provide insight into the students' discourse-comprehension abilities. This was operationalised into the following research questions: Given a set of reading tasks, is any of the tasks perceived as difficult by the students? How is each task faced by the students? How do the logs contribute to understanding the students' approaches to the tasks?

Section 2 introduces the rationale for focusing on higher-order comprehension skills in an MA course of interpreting, presents the theoretical framework informing the current exercises and analyses, and outlines previous uses of automatic data logging. Section 3 describes the learning scenario, the tasks, the platform, the subjects taking part in the experiment, and the analytical methods adopted in the study. Section 4 presents the results of our analyses, organised by research question. Finally, Section 5 draws conclusions and outlines the limitations of the current study and directions for future research.

2. Conceptual framework

This section outlines the conceptual models that informed the exercises and the analyses in the current experiment. In particular, Hatim and Mason's (1997) model on the interaction between interpreter and text suggested the use of these exercises with trainee interpreters; van Dijk and Kintsch's (1983) cognitive model of text comprehension provided the general theoretical framework, while Brown and Day's (1983) macro-rules for text summarization were used in the analysis of the data.

Each of these is briefly sketched below.

2.1 Interpreting

Interpreting is a form of translation in which the interpreter accesses the source text only once and produces a target text impromptu (Pöchhacker 2003, pp. 11). Comprehension is a fundamental skill in all forms of interpreting. Memorization and summarization abilities are also required, above all in consecutive and liaison interpreting, the two forms of interpreting trained in the current academic module.

As regards comprehension, several authors underline "the ability to process texts cognitively and analytically" (Kalina 2000, p. 18; see also: Moser-Mercer 1994, 2000,

among others). Hatim and Mason (1997) focus on the importance of textual analysis[3] in interpreting. In doing so they distinguish the following three composing elements of textuality: texture, structure, and context. Texture refers to "aspects of text micro-organization which contribute to the overall effects of texts hanging together internally" (Hatim 2013, p. 296). Structure is the "compositional plan" (Hatim 2013, p. 194) of a text. Context refers to the communicative, pragmatic and semiotic aspects of message construction (Hatim 2013, p. 284), interpreted by the listener with reference to some conceptual scheme that provides a general framework for text interpretation (Hatim, Mason 1997, p. 36). Each level of textuality predominates in a specific type of interpreting scenario. In consecutive interpreting, structure predominates, as the interpreter produces her output after the entire source text has been delivered and focuses on information that is relevant to text structure as a means to store context and texture effectively. Finally, context predominates in liaison interpreting, where – due to the dialogic nature of the events requiring this approach – texture and structure unfold piecemeal to the interpreter, and the interpreter largely resorts to contextual information in order to negotiate meaning.

Expert trainers, such as Moser-Mercer (2000), suggest that students should be gradually guided to the complex task of interpreting through a step-like approach, beginning from comprehension skills. Interestingly, as we shall see in the next section, some macro-rules that guide comprehension are also at work in memorization and summarization processes.

2.2 Comprehension skills

Van Dijk and Kintsch's cognitive model of text comprehension provides the theoretical framework for the current experiment. This model, which applies to both written and spoken texts, has largely informed empirical research on text comprehension, memorization and summarization.

According to van Dijk and Kintsch (1983), a text is composed of three layers: microstructure, i.e. the local organization of propositions into coherent pieces of discourse; macrostructure, i.e. the essential points of a text (i.e. the *gist*), organised as a coherent whole; and superstructure, i.e. conventional schemata that may or may not be present. Text comprehension involves mental representations at three different levels roughly corresponding to the three text structures: a *surface representation* focuses on microstructures; a *textbase representation* contains macrostructural information; finally, when information in the text is integrated with experience of the world, a *situation model* is created. Macrostructures are created by applying three macro-rules, respectively involving processes of deletion (of less important information), generalization (of propositions in the text) and construction (of new propositions/sentences). These mental processes underlay comprehension as well as summary activities. Brown and Day (1983) elaborated on van Dijk and Kintsch's model and listed the following six macro-rules for text summarization: 1. deleting unnecessary information; 2. deleting redundant information; 3. substituting a list of items with a superordinate word; 4. substituting a list of actions with a word expressing a superordinate event; 5. selecting a topic sentence; 6. inventing a topic sentence if none is available. Less experienced readers resort primarily or exclusively to deletion strategies, while more advanced readers make ample use of strategies involving substitution and construction (Winograd 1983).

[3] See Garzone (2000) for the preference of this term over others.

Empirical experiments have shown that surface representations decay rapidly from memory (see, for example, the literature review in Zwaan, Rapp 2006, p. 737). On the other hand, abilities connected to macrostructure, such as identifying important information and text structure, are significantly related not only to text comprehension, but also to memory and summarization skills (see, for example, the literature review in Armbruster *et al.* 1987). The above further explain the interpreter's focus on macrostructural features in consecutive.

Although comprehension is trained at all levels of schooling and from very early grades, studies on university students suggest that the latter do not necessarily master deep comprehension skills, not to mentions strategies. Frazier (1993), for example, observed that even a task which mirrors a basic reading strategy like distinguishing important from trivial information – one of the tasks used in the current experiment – may prove to be challenging at university level. Indeed, this task – like all strategies – involves conscious awareness of the reading process by the reader (Winograd, Bridge 1986; Williams 1988), i.e. it implies some level of metacognition.

2.3. Methods to investigate text-comprehension processes

Reading abilities can be analysed by assessing reading products (such as summaries), or by investigating the reading process. Several methods have been used to investigate text-comprehension processes, including think-alouds, interviews, error detection, questionnaires, eye tracking, and learning logs (manual or automatic).[4] These are all essentially (and inevitably) indirect methods of testing the comprehension process. The current experiment adopted automatic logs.

Automatic logging of the students' actions in a digital environment (e.g.: keystrokes, events, mouse movement, etc.) record explicit, automatically and instantaneously captured data, thus facilitating the collection, coding, aggregation and analysis of students' learning activities and performance, and offering researchers insight into the learning process (Long, Siemens 2011; Siemens, Baker 2010). Indeed, it has been shown that a person's interaction with an environment generates data from which cognitive and metacognitive interpretations can be made (Azevedo *et al.* 2010).

By tracing, recording and timestamping students' interactions (such as mouse click, mouse wheel, and which item was clicked or selected) with the content and questions of a reading comprehension task, Peckham and McCalla, (2012) were able to highlight different patterns of student behaviour. Their experiment demonstrated that an automatic recognition of the cognitive strategies used by students is possible. Zushi *et al.* (2015) discuss ways to analyse recorded mouse trajectories, response time, and drag and drop (D&D) logs in order to identify hesitations in word-reordering tasks, seen as potential indicators of the degree of learners' understanding. Researchers have used LearnWeb's automatic logging of the students' actions – alone or in combination with other methods – for the analysis of search processes (Holdack-Janssen, Marenzi 2012; Marenzi *et al.* 2016).

[4] For an overview of the differences between think-alouds, interviews, error detection, and questionnaires see Scott (2008). On think-alouds see also Charters (2003). On the use of eye tracking to identify and explain cognitive processes underlying reading, see Bax (2013), Rayner (1998), and Roberts (2012). On manual logs for evaluating reading strategies and text comprehension skills see Smith (1996), Soldner (1998), Kamijo (2012), and Ochoa Delarriva and Basabe (2015).

3. Materials and methods

This section describes the learning scenario, the participants, and the tasks. Furthermore, it illustrates the system's interface, with specific reference to the features used by the students and/or by the authors. Finally, it presents the corpus of data collected and the analytical methods.

3.1 Learning scenario

In the present experiment, LearnWeb and its TED-related features were tested in a module on interpreting in the first year of an MA curriculum for translators and interpreters. The module introduces students to consecutive interpreting without notes – also called short consecutive (Pöchhacker 2003, p. 19) –, seen as a preparatory activity to traditional consecutive as well as to liaison interpreting (which will be trained in a separate module on second year). The core skills trained in the current module are discourse comprehension, memorization, and summarization. At the very beginning of the module, attention is devoted to understanding discourse structure, distinguishing key elements of discourse from peripheral elements, and identifying discourse functions.

3.2 Participants

The experiment involved 25 Italian students studying English as their main foreign language. The students were in the 21-23 age range, with the exception of a senior student aged 47. This MA curriculum for translators and interpreters attracts graduate students in foreign languages from all over southern Italy. Consequently, the students involved in the experiment had similar, though not necessarily identical, study backgrounds, as every Italian university is free to decide on the contents and modules of their language curricula. As regards the students' knowledge of English, no specific assessment was made during the experiment, but it can be safely stated that they were B2+ level or above of the European Framework of reference, B2+ being the threshold for admittance to the MA curriculum.

3.3 Tasks

The experiment was based on three separate tasks revolving around a single TED video, titled "The simple power of hand-washing,"[5] by Myriam Sidibe, and its transcript. Like all TED talks, this video shows a monologic talk. The camera moves between long or close shots on the speaker, close shots on the projected slides, and long shots on the listening audience. The students were invited to watch the video before and while doing the exercises.

The tasks are reported in Figure 1. They mirror exercises previously done in class; however, their formats were adapted to the possibilities of the platform (see Section 3.4), and to the aim of exploring the potential of the logging system.

[5] http://www.ted.com/talks/myriam_sidibe_the_simple_power_of_hand_washing

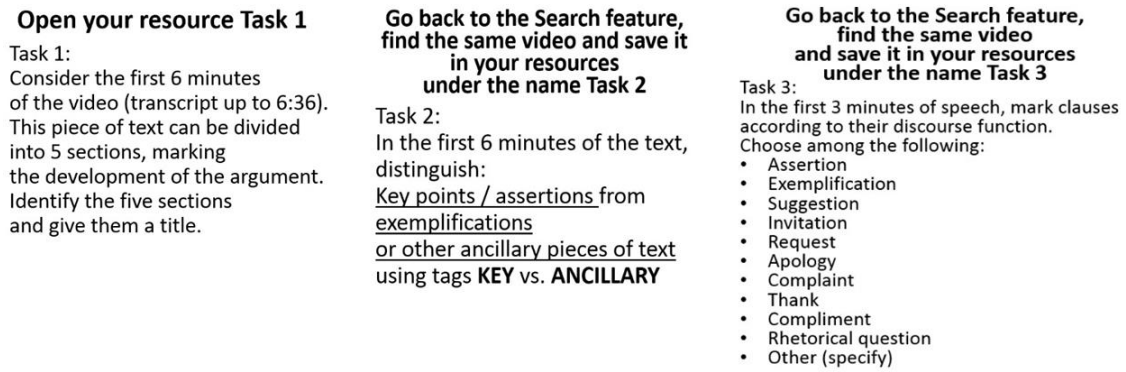

Figure 1
The three tasks.

Task 1 tests the students' understanding and summary skills. It involves the creation of a textbase representation of the text and also requires the students to verbalise it. Task 2 draws the students' attention to the application of the most basic reading and summary strategy, i.e. deletion. Finally, Task 3 draws the students' attention to the communicative function of discourse, and involves abilities at the level of both textbase representation and situation model, not to mention metalinguistic abilities. The list of functions provided includes all the discourse functions that appear in the given text, plus a few extra ones, added to provide a broad view of discourse functions; the last option (*Other*) was included to suggest that the list was incomplete.

To avoid crowding the same stretch of text with several annotations from different exercises, the students were instructed to save the same video three times with different names (Figure 1). This trick also proved useful when analysing the logs, as filtering the logs by resource gave us a focused view on each exercise.

LearnWeb is a very convenient platform for e-learning activities, but when the experiment took place the students were new to it; furthermore, the TED functionalities were still in an experimental phase and had not yet been tested with students. For these reasons, the exercises were planned and performed as class work, so that the teacher could intervene should technical problems arise.

The experiment stretched over two days. On day 1, the students were instructed on the use of the platform, were given a PowerPoint file containing a description of the tasks (Figure 1), and were invited to undertake Task 1. On day 2, they worked on Tasks 2 and 3.

3.4 LearnWeb

This section illustrates LearnWeb, with particular emphasis on the features used in the current experiment by the students and/or by the authors.

3.4.1 The system

LearnWeb is a learning and competence development environment that allows users to share and collaboratively work on resources collected from the web or user-generated (Marenzi, Zerr 2012). It provides users with a search interface for discovering and sharing resources across a range of Web 2.0 services such as YouTube, Flickr, and Slideshare, and LearnWeb itself, and offers a personal learning space. Resources in LearnWeb can be bookmarked, described, tagged, rated, commented and discussed by its users. By creating folders, users can organize resources belonging to the same learning context. In this way, users can rely on resources that have been collaboratively collected by the LearnWeb

Lingue e Linguaggi

community and categorized for specific learning scenarios. For a description of the affordances of LearnWeb as a collaborative platform see Marenzi and Nejdl (2012), Marenzi and Zerr (2012), and Marenzi (2014).

The following section illustrates LearnWeb's interface with respect to the steps used in the specific learning scenario.

3.4.2 TED-related features

LearnWeb allows users to display TED videos and their transcripts in the same window, as shown in Figure 2. Thus, students can watch a video and read its transcript simultaneously. The transcript is preceded by a brief explanation of the interactive features available. The current experiment took advantage of the selection and annotation features only, illustrated in Figure 3.

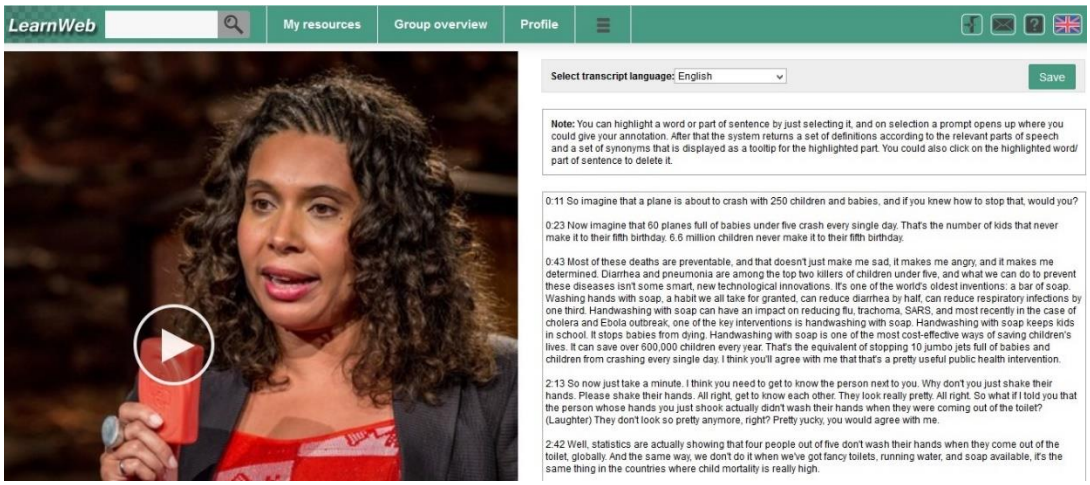

Figure 2
Transcript displayed next to the corresponding video.

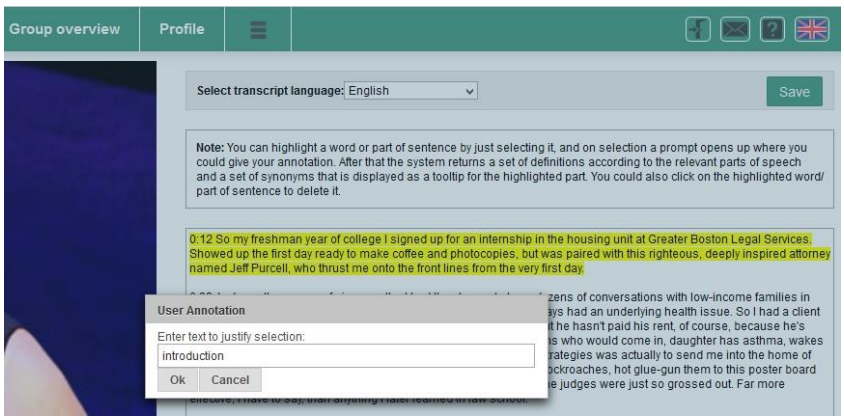

Figure 3
The annotation window.

When the mouse passes over a selected and annotated piece of text, the annotation is displayed in a black box at the end of the annotated text (Figure 4). Annotations and selections can be easily deleted by clicking on the annotated text (Figure 5), but, in order

to modify an annotation, it is necessary to delete the corresponding selection and make a new one.

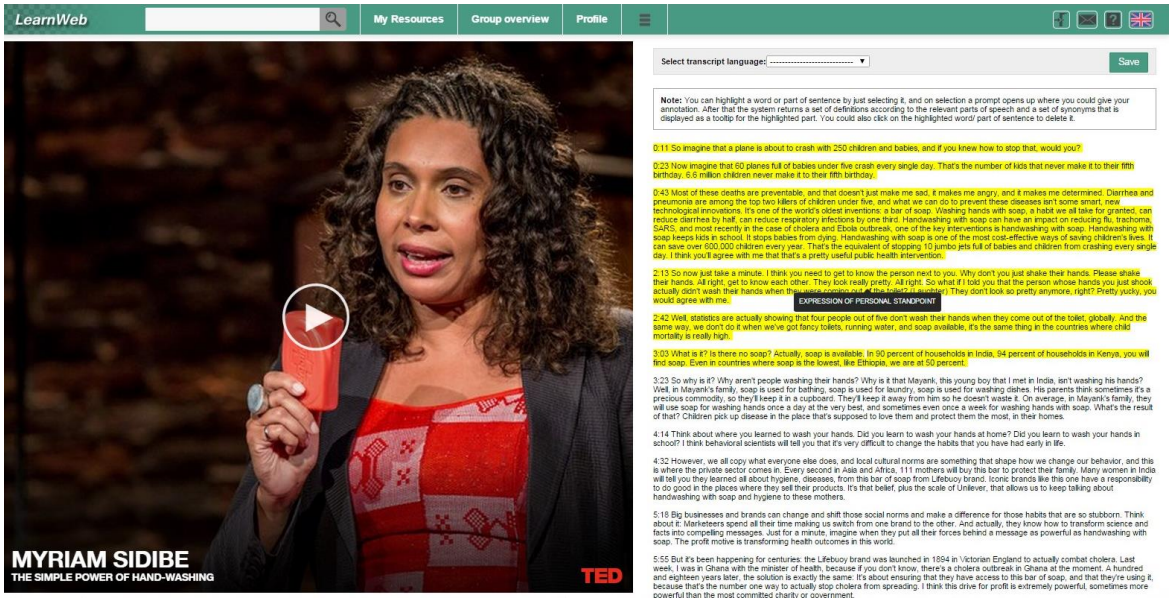

Figure 4
Student annotation.

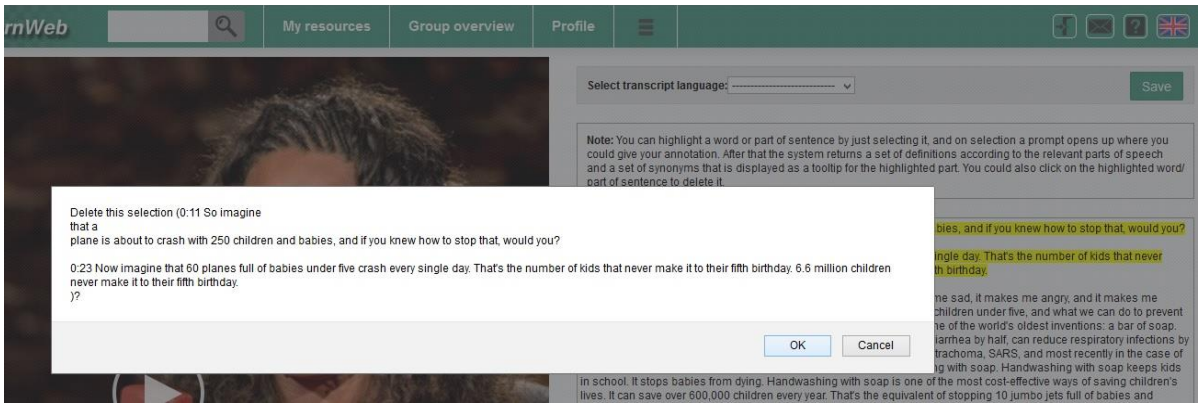

Figure 5
Deleting a selection and the corresponding annotation.

Finally, LearnWeb automatically logs the students' activities (their selections, tags, deletions, etc.) and provides the teacher and researcher with two different views of the logged data (Figures 6 and 7).

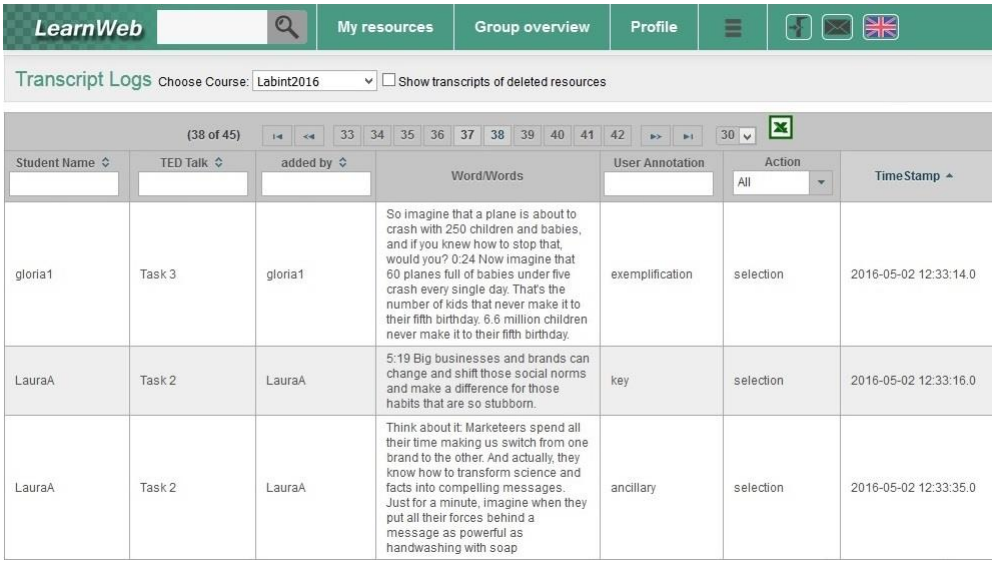

Figure 6
Detailed log, extract.

Figure 7
Simple log, extract.

The detailed log shows: the name of the student who added the tag (Student Name); the name given to the video when saving it into the local resources (TED Talk); the name of the student who added the video to the local resources (Added by); the selected text (Word/Words); the tag added by the student (User Annotation); the student's action (i.e. selecting or deselecting); and the time of the action (Time Stamp). The simple log offers a summary count of all selections, deselections, and annotations performed by a given student on a given resource. It also provides the teacher with the possibility to access the student's annotated resource, by clicking on the name of the resource in the TED Talk column. Filters allow the teacher to search the logs for a specific student, resource or other type of data. Furthermore, the detailed log data can be sorted by timestamp, while the simple log data can be sorted by count, in increasing or decreasing order.

3.5. The data and the research design

An initial look at the detailed log and at the annotated resources showed that some resources, as well as the corresponding logs, presented 'empty annotations', i.e. the system had recorded that an annotation was made, but – for a technical reason which was later discovered and fixed – had not memorised or logged its contents. The resources containing

empty annotations were excluded from the analyses, and the data analysed eventually included the annotated resources of 14 students in Task 1, 15 students in Task 2, and 13 students in Task 3, along with the corresponding logs.
Furthermore, the logs of some resources included multiple identical selections without annotation. As we subsequently discovered, these were the consequence of minor issues in the logging system and in the students forgetting to save their work frequently. However, it soon became clear that using the collected log data would prove less straightforward than expected, and in particular that the simple log could not be used for the analyses, as its counts reflected the technical problems encountered in the use of the platform, more than anything else. The primary source of information for the analyses was thus the detailed log.

The analytical phase was divided into three steps. Step 1 focused on the students' changes of mind: did they change their mind frequently while selecting and annotating? what types of changes did they make? Changes were considered as an indication of students not feeling more or less confident with the task; in other words, as a sort of measure of subjective difficulty. In step 2, we searched the student annotations for cues that could give us insight into the students tagging and annotation processes. Analytical Steps 1 and 2 took advantage of LearnWeb's log data, but not exclusively. Finally, Step 3, which is still in progress, sees the first author, who is also the teacher of the module, assessing the acceptability of the students' annotated resources. This step has two related aims: 1. to see whether correlations exist between the observed processes and the quality of the final products; and 2. to provide the developers of the platform with a rubric upon which to try and implement a learning analytics system.

The following sections describe the methods and results relating to analytical steps 1 and 2, and briefly hint at the issues involved in assessing the acceptability of the students' annotated resources for learning analytics.

4. Methods of analysis, results and discussion

The following paragraphs present the analyses and their results, organised by research question.

4.1 Was any of the tasks perceived as difficult by the students?

To answer this question, the detailed log was filtered by student name and task, and the student data were ordered in ascending chronological order. Next, the data were manually scanned to identify all the cases where a student had selected and annotated a piece of text and subsequently deselected it to change the selection span or the annotation. Finally, of all the cases found, only those where the changes were substantial rather than cosmetic were counted. Cosmetic changes included, for example: reselecting a sentence to include its first or last word or letter, which had been omitted before; or reselecting the same piece of text to substitute a grammatically awkward annotation with a neater one (e.g. passing from tag "Introduction: how can we reduce high children mortality rate?" to tag "Introduction: how can we reduce child mortality?"). On the other hand, substantial changes included cases such as: passing from the selection of a single sentence, to the selection of two or more sentences as a single piece; or replacing tag "Why they [people] do not do that [wash their hands]" with "How to teach people to wash their hands."

The percentage of students who substantially changed at least one selection/tag was 29%, 7%, and 38% in Tasks 1, 2, and 3, respectively; and the number of substantial

changes recorded was, respectively, 7, 1, and 8. These data suggest that the students felt rather confident of their analyses in Task 2. On the other hand, Task 1 and above all Task 3 were more problematic, with Task 3 recording a higher number of substantial changes across a higher percentage of students.

These results quantify and reflect the students' confidence in the given skill.

4.2 How was each task faced by the students?

At a general level, the data showed that a minority of students analysed more text than required (14%, 13%, 38% respectively in Task 1, Task 2, and Task 3). Furthermore, in all the tasks, the majority of students annotated the text in linear fashion, proceeding from the beginning down (86%, 67%, 69%, respectively for Tasks 1, 2, and 3). Interesting exceptions to this rule included the senior student, who followed a specific path when annotating Task 2. The time stamps of her annotations in this task suggest that she would read the text until she reached a key element and marked it as 'key'; after that she would go back and tag the preceding lines as ancillary; next, she would move to the text following the first key element and would go on reading until she reached a new key element and tagged it, after which she would go back to the untagged preceding lines and annotate them as ancillary; and so on, through the entire text. Another interesting case is that of a student who, in Task 2, annotated (in linear order) key elements only, thus showing that the annotation of ancillary elements could be considered redundant. Observation of the order in which selections and annotations were made was possible only thanks to the detailed log and its recording of the time of the students' actions in the TimeStamp column (Figure 6).

Let us now consider each task individually.

4.2.1 Task 1

Task 1 required the students to divide the text into five sections marking the development of the argument and to give each section a title. The students' annotations were compared to Brown and Day's (1983) macro-rules (see Section 2.2). The analysis showed that almost all the annotations matched macro-rule 6 (i.e. inventing a topic sentence), and involved strategies of substitution and creation. An exception to this was the annotation made by three students to the same piece of text: they all copied an existing topic sentence from the text. This pattern held true in the final versions, as well as in the 'draft' versions of the students' annotations, when present. Furthermore, it was observed that the students' annotations were structured as:

- Complete sentences (38%); e.g. *Every day 6.6 million children die for diseases*; *A bar of soap could save most of them*; *The problem is people don't wash hands frequently*; *Iconic brands' messages can make the difference in changing habits*. The use of sentences is a typical summary device.
- Long (noun)phrases (47%); e.g. *Introducing the problem*; *What can be done*; *Handwashing in countries with high child mortality rate*; *Handwashing statistics*; *The handwashing habit in America and in Africa and Asia*. The use of noun phrases instead of sentences can be considered an indication of good synthesis skills.
- Wh questions (16%); e.g. *How many children die and why?*; *How many people do not wash their hands?*; *Why they do not do that?* Questions are a valid and suggested method to support active reading (Rosenshine *et al.* 1996; Day, Park 2005; Taboada, Guthries 2006; Urlaub 2012).

Frequently, a single student adopted mixed types of annotations, as in the following example: *How can we reduce child mortality?*; *Contrast between soap availability and how many times do people wash their hands*; *Mayank's family: how do they use soap?*; *How can brands influence habits?*; *Profit versus charity in combating diseases such as cholera*.

The techniques adopted by the students in this exercise suggest good summary skills; furthermore, they are suitable to support memorization, a fundamental task in short consecutive interpreting.

4.2.2 Task 2

Task 2 required the students to distinguish key elements from ancillary ones. The tags to use were suggested in the instructions. No comment can thus be made about the form of the students' annotations. On the other hand, analysis of the students' selections showed that the students' unit of reference was the paragraph (47% of the students), the sentence (20%), or a mix of the previous two (33%).

Considering key elements entire paragraphs of a text, though theoretically possible in some contexts, is not the best approach in the current scenario. Unlike many written expository texts, the given text – being spoken – does not feature a summary paragraph. The paragraphs in the current transcript are frequently long and include several examples and repetitions of concepts that can hardly be considered key elements. The very notion of paragraph should be questioned in this particular text, since the paragraphs visible in the transcript – though rather logic – are constructions of the transcriber, not of the author.

So, despite the students' certainty in what they were doing, 80% of them adopted an unsuitable approach, which among other things posed serious problems to the teacher when it came to assessing and marking the students' work (see Section 3.4).

4.2.3 Task 3

Task 3 required the students to mark each clause in the first three minutes of the speech and specify its discourse function by choosing from a given list of functions. The list was open, in that the students could add extra, unlisted functions, if necessary.

Two students felt the need for unlisted functions, though this was not necessary. In fact, the functions they added were nothing but synonyms of given tags.

Furthermore, 77% of the students selected and tagged full paragraphs or sentences, instead of clauses. Such an approach, though in open contrast with the task directions, cannot be dismissed as totally wrong, given the fact that, more than once, the text included several clauses with the same function in a row. However, these students selected full sentences or paragraphs systematically, even when this did not make sense. This of course can be attributed to the students not reading the instructions carefully or not understanding the term 'clause'. However, it also suggests little sensitivity for discourse functions and limited familiarity with this type of pragmatic task.

Finally, it was observed that the students' deselections and changes of tag were primarily located towards the beginning of the text. This may be a further indication of their limited familiarity with this type of task, or it may be a consequence of their linear approach (i.e. selecting and annotating text sequentially, from the beginning of the text down).

4.3 How did the logs contribute to an understanding of the students' approaches to the tasks?

Most of the analyses were based on evidence that is visible also in the students' final products, i.e. in the annotated resources. However, the system's logs played more than one important role. In particular, the detailed log proved fundamental to:

- Highlight which students had experienced technical problems in which task, and help the researcher select a corpus of data devoid of biases.
- Access the draft versions of the students' work and understand whether a student had hesitations while performing the task.
- Observe the order in which selections and annotations were made, thanks to the presence of TimeStamps.
- Understand the beginning and end of selections, when multiple selections and tags appeared in contiguous pieces of text.
- Finally yet importantly, it allowed the authors to access the student's work even when the students forgot to save it.

The data in the simple log could not be used in the current analyses. In fact, it was soon noticed that the simple log counts reflected the technical problems encountered in the use of the platform, more than anything else. However, once the problems described in Section 3.2.3 are solved, this log may provide an automatic image of the students' hesitations. Finally, the simple log was fundamental to give the authors access to the students' final products. Without direct access to the students' personal resources, some logged data would have been difficult to interpret.

4.4 Assessing the students' work

The final step of this project – i.e. assessing the acceptability of the students' annotated resources in order to see whether correlations exist between the observed processes and the quality of the final products, and to provide the developers of the platform with a rubric upon which to try and implement a learning analytics system – is still in progress. The current section, therefore, will focus on issues, rather than results, and outline the problems currently being faced by the authors in assessing the quality (i.e. level of acceptability) of the students' final products.

Assessing students' work is not always easy; but it can be really challenging when it has to be based on an assessment scheme that is suitable to inform a learning analytics system. By their very nature, learning analytics systems are incompatible with holistic marking, and require an analytical marking system, anchored in automatically trackable and identifiable features. The exercises used in this experiment are highly problematic in this respect.

Task 1 asked students to divide the text into five sections and give them a title. In this type of task, it is rather easy to assess the students' final annotations manually: the teacher reads the students annotations and decides, one by one, whether their meanings hit the core topics of the corresponding selected paragraph(s) and whether, taken in sequence, the annotations manage to create a suitable structure of the given text. In the current experiment, the students did a generally good job, with the minor exception of two students (14% of the group) who included, respectively, one and two annotations that did not target the contents of the corresponding section. On the other hand, deciding upon their correctness can hardly be done by a computer. A major problem is the fact that the

students' selection spans were all different, or could potentially be all different; another major issue is that annotations were open, and the students' phrasings were not only all different, but made ample use of strategies of substitution and creation. In other words, Task 1 is an exercise where an infinite number of correct solutions exist.

Tasks 2 and 3 pose no fewer problems, despite their being basically exercises with a clear set of limited choices. Task 2 required the students to distinguish key items in the text from peripheral (ancillary) ones. The first issue in marking such an exercise is the fact that there is a margin of subjectivity in the notion of key item. This could be overcome by having the works marked by a panel of raters, and then by looking for some level of interrater agreement. The second issue is linked to the fact that, before classifying a piece of text as key or ancillary, the students had to select a text span, and they frequently adopted unsuitable selection spans. Similarly, Task 3 – requiring the students to select clauses and specify their discourse function by choosing from a given (open) list – saw the majority of students opting for unsuitable selection spans, although the selection span (clauses) was specified in the task directions. How should this unexpected behaviour be considered? From our perspective, this is an issue even in holistic marking: prioritising selection over annotation would mean failing the majority of these students in these tasks, without appeal; prioritising annotation over selection would mean adopting a rather flexible perspective, and this would leave us with a feeling of uncertainty and unfairness.

Despite these issues, the authors are working to see whether some analytical assessment scheme can be developed that may help establish whether correlations exist between the observed processes and the quality of the final products. Furthermore, they are investigating the possibility of tracking alternative, automatically manageable parameters to instruct a learning analytics system without drastically changing the structure of the exercises. After all, learning analytics are not meant to judge the students work, but to provide constructive on-line feedback to teachers and students.

5. Conclusions

The current paper has illustrated a possible use of TED talks and the interactive learning platform LearnWeb in a module on consecutive interpreting. However, the types of exercises used would fit a large range of learning scenarios, and the resources, analytical methods and results described in this paper may be relevant to anyone interested in discourse comprehension.

In particular, the paper has described how LearnWeb can help the teacher to draw the students' attention to some of the composing factors of active text comprehension, on the one hand, and to verify whether the students possess higher-order comprehension abilities and strategies, on the other, by means of simple selection and annotation exercises. LearnWeb's automatic logging system collected data on the students' selection and annotation processes. These data were used by the researchers to address the following questions: Was any of the tasks perceived as difficult by the students? How was each task faced by the students? How did the logs contribute to understanding the students' approaches to the tasks?

The analyses have shown that the students felt rather confident of their analyses in Task 2. On the other hand, Task 1 and above all Task 3 were more problematic, with Task 3 recording a higher number of substantial changes across a higher percentage of students. At the same time, however, many of the students made gross mistakes in Task 2, in so far as the majority of them adopted, either systematically or sporadically, totally unsuitable selection spans. This was also true, though to a lesser extent, for Task 3. These data

suggest that a large number of the students in this group had good summarising abilities – attested by their annotations in Task 1, based almost exclusively on strategies of substitution and creation –, but little awareness of the strategies and skills they automatically apply in the reading process – attested by the generally poor results in Tasks 2 and 3. Furthermore, the limited number of changes, especially in Task 2, suggest that most of the students were not aware of their limited metacognitive and analytical abilities.

The analyses were largely based on evidence that was visible in the students' final products, i.e. in the annotated resources. However, the system's logs played more than one important role, as summarised in Section 4.3. In particular, the detailed log gave the researchers access to the draft versions of the students' work, which was necessary to understand the students' confidence in performing individual tasks. Furthermore, it showed a generalised linear approach of the students in the selection and annotation tasks. Although no detailed and quantitative analysis of the acceptability of the students' annotated resources was made, the overall results of the students at Tasks 2 and 3 suggest that a linear approach to selecting and annotating text may be a poor strategy, compared to other, less linear ones. Hypothetically, this may even be a reflex of a linear reading approach. However, further research in this direction is necessary. Finally, the simple log, though not used in the current experiment because of technical issues, may provide an automatic image of the students' hesitations once those issues are solved.

Automatic logging of the students' selection and annotation processes have thus proven to be a method that may support researchers in the investigation of higher-order comprehension skills, if used in conjunction with other analytical methods. The kind of information this type of log provides largely depends on the types of actions being logged and on the types of exercises being used to assess reading abilities. At the time of the experiment, LearnWeb offered a limited set of interactive functionalities on TED talks and its logging features recorded a rather limited set of data. At the time of writing, LearnWeb has already been updated and some of the issues outlined in the current experiment have been fixed. Furthermore, in keeping with LearnWeb's characteristic approach of cyclical assessment and improvement of the system's functionalities based on the needs of teachers in real scenarios, L3S researchers and the teacher are working at the design of new functionalities – to support the use of the platform in translation and interpreting courses and in reading-comprehension tasks – and at repeating the experiment in the updated environment.

6. Limitations of the current experiment and future research

A major limitation of the current research is its relying on two sources of data only (logs and final products). The amount of evidence at our disposal highlighted the students' choices and changes in performing higher-order comprehension tasks and allowed us to make hypotheses about the students' cognitive and metacognitive reading skills. However, in order to reach more interesting and conclusive findings, other investigative methods should be adopted, in complementary relation with the current ones. In particular, think-alouds and/or retrospective interviews or questionnaires could be used to gather subjective information about the tasks and further information about the cognitive processes applied by the students while performing the exercises. Furthermore, eye-tracking techniques could be used to verify whether linear selections and annotations reflect a linear reading approach.

Another limitation is the fact that the final step in the original design of this project – i.e. assessing the acceptability of the students' annotated resources in order to see

whether correlations exist between the observed processes and the quality of the final products – has not been completed and is proving problematic. Therefore, besides making some last attempts at the development of an analytical assessing scheme that may prove useful to the above purpose, we are also investigating the possibility of tracking alternative, automatically manageable parameters to instruct a learning analytics system without drastically changing the structure of the exercises.

Finally, as the selection phase in the three exercises proved highly problematic for the students and created difficulties to the researchers in the analyses of the data, we are considering to (partially) alter the structure of the exercises, and display transcripts in which the sections to tag are already clearly indicated. This may eventually provide a wider and more easily analysable set of data.

Authorship: This study was planned and carried out by both co-authors. Ivana Marenzi is responsible for writing Section 3.4. Francesca Bianchi for writing the remaining sections.

Bionota: Francesca Bianchi, PhD., is lecturer of English Language and Translation at the University of Salento. She has been involved in the creation of pedagogical units using multimedia authoring systems, in the development of computerized interactive teaching environments, and in the use of subtitling as a CLIL task. She has participated in nationally financed projects and collaborated with a range of Italian and European universities, including Pavia, Trieste, Hannover, Lancaster, and CNR-Pisa. She has published several papers and has organized international conferences on the use of technologies in either language learning or linguistic research.
Ivana Marenzi, PhD, is senior researcher at the L3S Research Center in Hannover, Germany. Throughout her career she has specialised in the relationship between technology and communication. She obtained her PhD on the interplay between Content and Language Integrated Learning and Web 2.0. Her main area of research in Technology Enhanced Learning (TEL) includes the support of collaborative and lifelong learning. She is referee for international journals (e.g. Elsevier LeaInD, IJTEL) and conferences (e.g. EC-TEL, ICWL, CSEDU, ICALT, LILE), and has collaborated in organizing conferences (e.g. ICWL 2016) and workshops (EbTEL 2012; TEFA 2013, 2016).

Email addresses: francesca.bianchi@unisalento.it; marenzi@l3s.de

Acknowledgements: The authors would like to thank Paola Leone and the two anonymous reviewers for their comments on earlier versions of this paper.

References

Armbruster B.B., Anderson T.H. and Ostertag J. 1987, *Does text structure/summarization instruction facilitate learning from expository text*, in "Reading Research Quarterly" 22 [3], pp. 331-346.

Aston G. and Rodi D. 2012, *Speech corpora for language learning*, in "TALC10, 10th Teaching and Language Corpora Conference", University of Warsaw, 11-14/7/2012.

Azevedo R., Johnson A., Chauncey A. and Burkett C. 2010, *Self-regulated learning with MetaTutor: Advancing the science of learning with metacognitive tools*, in Khine M.S. and Saleh I.M. (eds.), *New Science of Learning. Cognition, Computers and Collaboration in Education*, Springer, New York, pp. 225-247.

Bax S. 2013. *The cognitive processing of candidates during reading tests: Evidence from eye-tracking*, in "Language Testing" 30 [4], pp. 441-465.

Bortoluzzi M. and Marenzi I. 2013, *YELLing for partnership: A social platform for sharing practice and reflection in teacher training for language learning*, in Riem Natale A., Dolce M.R., Mercanti S. and Colomba C. (eds.), *The Tapestry of the Creative Word in the Literatures in English*, Forum, Udine, pp. 249-262.

Bortoluzzi M. and Marenzi I. 2014, *YELLing for collaborative learning in teacher education: users' voices in the social platform LearnWeb2.0*, in "International Journal of Social Media and Interactive Learning Environments (IJSMILE)" 2 [2], pp. 182-198.

Brown A.L. and Day J.D. 1983, *Macrorules for summarizing texts: The development of expertize*, in "Journal of Verbal Learning and Verbal Behavior" 22, pp. 1-14.

Charters E. 2003, *The use of think-aloud methods in qualitative research. An introduction to think-aloud methods*, in "Broch Education Journal" 12 [2], pp. 68-82.

Day R.R. and Park J.-s. 2005, *Developing reading comprehension questions*, in "Reading in a Foreign Language" 17 [1], pp. 61-73.

Frazier D.W. 1993, *Transfer of college developmental reading students' textmarking strategies*, in "Journal of Reading Behavior" 25 [1], pp. 17-41.

Garzone G. 2000, *Textual analysis and interpreting research*, in "The Interpreter's Newsletter" 10, pp. 69-88.

Hatim B. 2013, *Translation studies: History, basic concepts and key issues in Research. Second Edition*, Routledge, London/New York.

Hatim B. and Mason I. 1997, *The Translator as Communicator*, Routledge, London.

Holdack-Janssen E. and Marenzi I. 2012, *Learning about literature on the Web in a German school*, in Vittorini P., Gennari R., Marenzi I., De La Prieta Pintado F. and Corchado Rodriguez J.M. (eds.), International Workshop on Evidence-based Technology Enhanced Learning, Springer-Verlag, Berlin/Heidelberg, pp. 57-65.

Kalina S. 2000, *Interpreting competences as a basis and a goal for teaching*, in "The Interpreters' Newsletter" 10, pp. 3-32.

Kamijo T. 2012, *Evaluating student reading strategies in a sophomore EAP course using student learning logs*, in "Ritsumeikan Higher Education Studies" 12, pp. 145-159.

Long P.D. and Siemens G. 2011, *Penetrating the fog: analytics in learning and education*, in "EDUCAUSEreview" 46 [5], pp. 31-40. Available at: http://er.educause.edu/~/media/files/article-downloads/erm1151.pdf (8.10.2016).

Marenzi I. 2014, *Multiliteracies and E-Learning2.0.*, Peter Lang, Frankfurt am Main.

Marenzi I., Bortoluzzi M. and Kalyani R. 2016, *YELL/TELL: online community platform for teacher professional development*, in Papadima-Sophocleous S., Bradley L. and Thouësny S. (Eds), *CALL communities and culture – Short papers from EUROCALL 2016*, Research-publishing.net, Dublin, pp. 307-312.

Marenzi I. and Kantz D. 2013, *ESP course design – a multiliteracies approach*, in *Proceedings of the 13th IEEE International Conference on Advanced Learning Technologies (ICALT 2013)*, IEEE Computer Society, Los Alamitos, CA, pp. 221-223.

Marenzi I. and Nejdl W. 2012, *I search therefore I learn. Supporting active and collaborative learning in language teaching*, in Okada A.L.P., Connolly T. and Scott P.J. (eds.), *Collaborative Learning 2.0: Open Educational Resources*, IGI Global, Hershei, PA, pp. 103-125. Available at: http://www.igi-global.com/chapter/search-therefore-learn-active-collaborative/64402 (8.10.2016).

Marenzi I. and Zerr S. 2012, *Multiliteracies and active learning in CLIL: the development of LearnWeb2.0*, in "IEEE Transactions on Learning Technologies (TLT)" 5 [4], pp. 336-48.

Moser-Mercer B. 1994, *Aptitude testing for conference interpreting. Why, when and how*, in Lambert. S and Moser-Mercer B. (eds.), *Bridging the Gap. Empirical Research in Simultaneous Interpreting*, John Benjamins, Amsterdam/Philadelphia, pp. 57-68.

Moser-Mercer B. 2000, *The Acquisition of interpreting skills*, in Gran L. and Kellet Bidoli C. (eds.), *L'interpretazione nelle lingue dei segni: aspetti teorici e pratici della formazione*. *Signed-language interpretation and training: Theoretical and practical aspects*, EUT, Trieste, pp. 57-61.

Ochoa Delarriva O. and Basabe E.A. 2015, *Reading logs and literature teaching models in English language teacher education*, in "HOW" 22 [2], pp. 37-53.

Park S.-M. and Cha K.-W. 2013, *Pre-service teachers' perspectives on a blended listening course using TED Talks*, in "Multimedia-Assisted Language Learning" 16 [2], pp. 93-116.

Peckham T. and McCalla G. 2012, *Mining student behavior patterns in reading comprehension tasks*, in Yacef K., Zaïane O., Hershkovitz H., Yudelson M. and Stamper J. (eds.), *Proceedings of the 5th international conference on educational data mining*, pp. 87-94. Available at: http://files.eric.ed.gov/fulltext/ED537180.pdf (8.10.2016).

Pöchhacker F. 2003, *Introducing Interpreting Studies*, Routledge, London/New York.

Rayner K. 1998, *Eye movements in reading and information processing: 20 years of research*, in "Psychological Bulletin" 124 [3], pp. 372-422.

Roberts L. 2012, *Psycholinguistic techniques and resources in second language acquisition research*, in "Second Language Research" 28 [1], pp. 113-127.

Rosenshine B., Meister C. and Chapman S. 1996, *Teaching students to generate questions: A review of the intervention studies*, in "Review of Educational Research" 66 [2], pp. 181-221.

Scott D.B. 2008, *Assessing text processing: A comparison of four methods*, in "Journal of Literacy Research" 40, pp. 290-316.

Shimizu H., Neubig G., Sakti S., Toda T. and Nakamura S. 2014, *Collection of a simultaneous translation corpus for comparative analysis*, in Calzolari N. *et al.* (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pp. 670-673. http://www.lrec-conf.org/proceedings/lrec2014/index.html (8.10.2016).

Siemens G. and Baker R. 2010, *Learning analytics and educational data mining: towards communication and collaboration.* https://pdfs.semanticscholar.org/c245/a3403d6fd4064e6fd7666914a29afc7fe5ea.pdf (8.10.2016).

Smith B. 1996, *Learning logs: a tool for cognitive monitoring*, in "Journal of College Reading and Learning" 228 [1], pp. 5-11.

Soldner I. 1998, *Reflection and developmental readers: facilitating metacognition with learning logs*, in "Journal of College Literacy and Learning" 29, pp. 18-24.

Sung S.-e. 2014, *Using TED in the Undergraduate Interpreting Classroom to Enhance Listening Competence*, in "STEM Journal". http://scholar.dkyobobook.co.kr/searchDetail.laf?barcode=4010023891873 (8.10.2016).

Taboada A. and Guthries J.T. 2006, *Contributions of student questioning and prior knowledge to construction of knowledge from reading information text*, in "Journal of Literacy Research" 38 [1], pp. 1-35.

Taibi D, Chawla S., Dietze S., Marenzi I. and Fetahu B. 2015, *Exploring TED talks as linked data for education*, in "British Journal of Educational Technology (BJET)" 46 [5], pp. 1092-1096.

Urlaub P. 2012, "*Reading strategies and literature instruction: teaching learners to generate questions to foster literary reading in the second language*", in "System" 40 [2], pp. 296-304.

Van Dijk T.A. and Kintsch W. 1983, *Strategies of discourse comprehension*, Academic Press, New York.

Wang F. and Hannafin M.J. 2005, *Design-based research and technology-enhanced learning environments*, in "Educational Research & Development" 53 [4], pp. 5-23.

Williams J.P. 1988, *Identifying main ideas: a basic aspect of reading comprehension*, in "Topics in Language Disorders" 8 [3], pp. 1-13.

Winograd P.N. 1983, *Strategic Difficulties in Summarizing Texts. Technical Report No. 274*. http://eric.ed.gov/?id=ED228616 (8.10.2016).

Winograd P.N. and Bridge C.A. 1986, *The comprehension of important information in written prose*, in Baumann J.B. (ed.), *Teaching Main Idea Comprehension*, International Reading Association, Newark, DE, pp. 18-48.

Zushi M., Miyazaki Y. and Norizuki K. 2015, *Analysis of learners' study logs: mouse trajectories to identify the occurrence of hesitation in solving word-reordering problems*, in *LAK '15 Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*. http://dl.acm.org/citation.cfm?id=2723645. (8.10.2016).

Zwaan R.A. and Rapp D.N. 2006, *Discourse comprehension*, in Traxler M.J. and Gernsbacher M.A. (eds.), *Handbook of Psycholinguistics. Second Edition*, Elsevier, London, pp. 725-764.