

Univerzitet u Sarajevu

Filozofski fakultet

Odsjek za anglistiku

ZAVRŠNI MAGISTARSKI RAD

The Role of Verbal Protocols in EFL Reading Comprehension

Uloga verbalnih protokola u razumijevanju teksta na engleskom kao stranom jeziku

Kandidat

Aida Mehmedić

Mentor

doc. dr. Nejla Kalajdžisalihović

Sarajevo, 2021.

Contents

Sažetak	2
Abstract	3
1 INTRODUCTION.....	4
2 VERBAL PROTOCOLS: THEORETICAL FRAMEWORK.....	6
2.1. Reading Comprehension.....	6
2.1.1. Reader	8
2.1.2. Text	9
2.1.3. Reading Comprehension Activity.....	9
2.2. Verbal Protocols.....	10
2.2.1. Think–Aloud Data	12
3 RESEARCH FRAMEWORK.....	14
3.1. Participants	15
3.2. Instrument.....	16
4 ANALYSIS AND INTERPRETATION OF RESULTS	17
4.1. EXCERPT 1: <i>Waggle Dance</i>	17
4.2. EXCERPT 2: <i>Language Universals</i>	22
4.3. EXCERPT 3: <i>The Hearer</i>	28
5 DISCUSSION OF THE FINDINGS IN RELATION TO READING EASE SCORES....	33
6 CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH.....	36
REFERENCES.....	38
APPENDIX 1.....	40
APPENDIX 2.....	50

Sažetak

Cilj ovog završnog magistarskog rada je istražiti i odrediti faktore koji utiču na razumijevanje teksta na engleskom kao stranom jeziku. Prvi dio rada tiče se teorijskog okvira i definicijom „čitanja s razumijevanjem“. Nakon navođenja elemenata najbitnijih za razumijevanje teksta, prelazi se na sam problem, odnosno kreće se sa istraživačkim dijelom rada. U istraživanju je učestvovalo deset studenata drugog ciklusa anglistike Filozofskog fakulteta u Sarajevu. Data su im tri kratka isječka od ukupno 395 riječi, iz udžbenika koji smo koristili za predmet Psiholingvistika na četvrtoj godini studija: *The Articulate Mammal*. To bi značilo da su im ti tekstovi bili poznati od prije. Njihov zadatak je bio da svaki isječak teksta pročitaju sami za sebe, te da ih parafraziraju. Tokom cijelog razgovora, uključena je opcija snimanja zadatka preko internet stranice online-voice-recorder.com. Potom su sakupljeni podaci transkribovani i kodirani te je analizirano ukupno 3418 riječi, što će pružiti direktan uvid u mentalni proces sudionika tokom izvršavanja zadataka koji se odnose na razumijevanje teksta, parafraziranje i dosjećanje. Rezultati istraživanja pokazuju da je metajezik bio veoma zahtjevan u procesu parafraziranja u odnosu na ostale uzorke teksta. Prema skali čitljivosti, prvi isječak teksta je „veoma lagan za čitanje“ (74.9), drugi isječak „veoma težak za čitanje“ (56.7), dok je treći isječak „standardan ili prosječan“ (60.7). Rezultati istraživanja odgovaraju skali čitljivosti kada govorimo o verbalnim protokolima. Možemo zaključiti i da, bez obzira što je tekst „standardan ili prosječan“, poteškoće u razumijevanju mogu biti povezane sa teškoćama u shvatanju datog poređenja u tekstu. To može biti prouzrokovano različitim faktorima, od općeg znanja, do faktora povezanih sa znanjem u određenom domenu. U svakom slučaju, s obzirom na to da je tekst zbog prisustva elemenata komparacije i usporedbe lakši za razumijevanje, predloženo je da se poređenja i metaforička poređenja više koriste u tekstu s ciljem podučavanja sadržaja.

Nakon završene analize, predloženo je i da se urade dalja istraživanja, te da bi bilo korisno da se u budućnosti pri pisanju udžbenika ili testiranju iz udžbenika čiji sadržaj usvajaju studenti kojima je engleski strani jezik obrati pažnja na pojave koje se odnose na procesiranje teksta na engleskom jeziku, a o kojima će biti govora u ovom završnom magistarskom radu.

Ključne riječi: čitanje s razumijevanjem, verbalni protokoli, procesiranje jezika, metajezik

Abstract

The aim of this master's thesis is to examine and determine factors that have an effect on students' comprehension of a text in English as a foreign language. The first part of the paper is concerned with the theoretical framework and definition of "reading comprehension". After outlining the main elements of the latter, we move on to the problem at hand, which is the research part of the paper. The research has been conducted with ten students from the master's degree program at the Department of English Language and Literature at the Faculty of Philosophy in Sarajevo who participated in the research. The participants were given three short excerpts to read, a total of 395 words. The excerpts were taken from a textbook that we used for the subject Psycholinguistics in our fourth year, *The Articulate Mammal*, meaning that the participants were familiar with the texts presented to them. Their task was to read the paragraphs and then paraphrase them. During the conversation, a recording application was turned on, using the website online-voice-recorder.com. The collected data was later transcribed and coded. A total number of 3418 words is analyzed, which will provide a direct insight into participants' mental processes during completion of the task related to reading comprehension. Results show that metalanguage was fairly difficult to paraphrase, in comparison with other text excerpts. According to the reading scores, Excerpt 1 is "fairly easy to read" (74.9), Excerpt 2 is "fairly difficult to read" (56.7) whereas Excerpt 3 is "standard or average" (60.7). The results correspond to the readability scores when it comes to verbal protocols. It can be concluded that, regardless of the text being "standard or average", difficulties with comprehension can be connected to the difficulty to understand the comparison given in the text. This can be caused by different factors, one of them being general knowledge, or factors related to domain-specific knowledge. In any case, since the comparison rendered the text easier to read, it is suggested that comparisons and metaphorical comparisons be used when teaching content via language.

After the analysis, further research is suggested. Furthermore, it would be useful if during the process of writing textbooks or testing from textbooks used by EFL learners, more attention is devoted to occurrences related to the processing of the text in the English language.

Key words: reading comprehension, verbal protocols, language processing, metalanguage

1 INTRODUCTION

To read without reflecting is like eating without digesting.

Edmund Burke

We acquire the ability to read at a very early stage of life, however, the learning process is more complicated than only being able to read the written material. As time goes by, we usually get more proficient as readers, however, we still face certain difficulties while comprehending what we are reading. Factors such as fluency and vocabulary can have an immense impact on one's ability to comprehend the material. Comprehension is both understanding and interpretation of what is read, and a big part of comprehension is having sufficient vocabulary and prior knowledge of the matter. Readers with strong reading comprehension skills are able to draw conclusions from the materials, distinguish what is important, what is a fact and what an opinion, and find hidden meanings which is also known as "reading between the lines". For readers of English as a foreign language, this task is even more challenging, as this reading process differs from reading in a native language. Sometimes readers spend a considerable amount of time translating words or phrases, mechanically focusing more on the meaning of separate units rather than on the sentence as a whole. This can result in the reader not quite understanding the point of the text, and in the process, the comprehension is lost.

Regardless of the fact that many different factors related to the reader's knowledge affect reading comprehension, we must also examine whether the type of the text being read plays a significant role in that process. As it was mentioned, vocabulary and prior knowledge on the matter are important, which means that, if the reader is faced with a piece of work that is more familiar to them, they will comprehend it more successfully than one they are presented with for the first time. Furthermore, if a reader is dealing with a text that is written in such a manner that they can easily understand, regardless of their previous knowledge on the matter, we can assume that they will process it rather successfully.

There are different types of text that a reader can come across, and we can assume that the efficiency of comprehension will depend on each type. Some types of texts are less challenging, while some require investing more time and attention in order to be understood. In a classroom

scenario, the teacher is able to provide help to students in order to improve their understanding. Unfortunately, teachers are limited with time, which means that, sometimes, it is their task to lead students to solve their problems on their own. In order to be able to do this as efficiently as possible, the teacher should first have an insight in what exactly are the major issues for students, and where their strengths lie. One way of gathering this information is through the use of *verbal protocols*.

According to Ericsson and Simon, the term verbal protocol is used to refer to participant's verbalization of their thoughts and successive behaviors while they are performing cognitive tasks. (1993, p. 1). This results in a direct insight into one's thinking process. Since verbal protocols are vast data sources that contain individual's spoken thoughts, it is safe to assume that using this method of research could be of importance in understanding EFL reading comprehension.

Having this in mind, the starting points of this paper will be the definition of reading comprehension and reading comprehension elements as well as verbal protocols and think-aloud data. When it comes to the research part, it will be conducted with ten students from masters' degree studies program at the Faculty of Philosophy in Sarajevo, Department of English Language and Literature.

The research will entail making of verbal protocols by recording conversations with the students who will be doing a simple task that will be described below. These recordings will then be transcribed and analyzed and serve as a corpus for this paper. The readability score for three text excerpts will be calculated as well. Therefore, the aim of this thesis is to examine and determine factors that affect a better understanding of content in English as a foreign language, the main focus being on proving that a text written using high-frequency words is easier to comprehend than the one that uses more complicated phrases and presents a certain idea in an abstract way. Based on this, the following hypotheses are proposed:

H1: In verbal protocols, paraphrases correspond to the reading ease score.

H2: Textbook material written in high-frequency language and using metaphors to convey new information is less difficult to comprehend than explanations using metalanguage.

H3: Verbal protocols can be used to demonstrate how reading comprehension works.

2 VERBAL PROTOCOLS: THEORETICAL FRAMEWORK

2.1. Reading Comprehension

It was once assumed that reading comprehension was simply a combination of decoding and oral comprehension skills (Hoover & Gough, 1990). That is, if readers were able to decode the words on a page, they would be able to monitor what was being read to themselves orally and understand what they were reading. However, contemporary research in reading comprehension has actually suggested that understanding what one reads involves more than decoding plus oral language comprehension (Duke & Pearson, 2002; Pressley, 2000). Essentially, just because readers can decode words does not mean that they have understood what they have read.

The RAND Reading Study Group (RRSG) defines reading comprehension as the process of simultaneously extracting and constructing meaning (Sweet & Snow, 2003). Both challenges are essentially recognized: figuring out how print represents words and accurately translating print to sound (extracting), and formulating a representation of the information being presented, which invariably necessitates building new meanings and integrating new with old information (constructing meaning).

Furthermore, Sweet and Snow (2003) suggest that comprehension entails three elements:

- 1 The reader who is doing the comprehension
- 2 The text that is to be comprehended
- 3 The activity in which comprehension takes part (p. 2)

While considering the reader, all abilities, information, and experiences that a person brings to the act of reading are included. The term “text” refers to anything that is read, whether it is printed or electronic. Three dimensions are considered while considering activity: purposes – why do readers read; processes – what mental activity do they participate in while reading; and outcomes – what do readers learn or experience as a result of their reading.

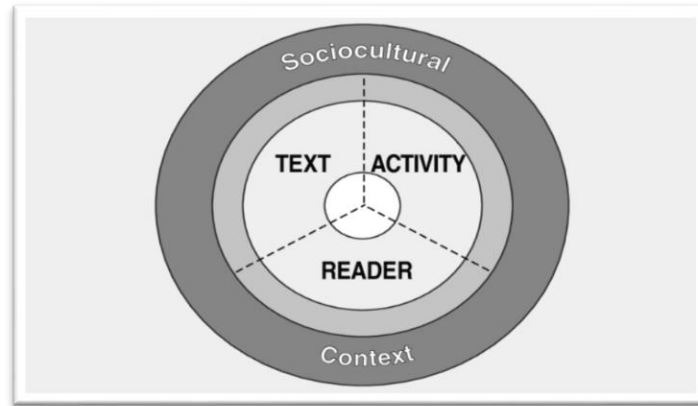


Fig. 1. A heuristic representation on thinking about reading comprehension (RAND Reading Study Group, 2002)

On the diagram in Fig. 1. we can see how the three elements are related. It is important to emphasize that this in no way suggests that they are independent elements. The reader's capabilities are related to a particular text, and there are no good or bad readers. However, some of us are good at reading novels, others are good at reading scientific texts or young adult literature. Moreover, some people are not particularly good at reading poetry, or non-fiction texts. Ultimately, it is "the interaction between the reader and the text that determines if the comprehension will be successful or not" (Sweet & Snow, 2003).

Additionally, readers' success varies with different activities. Sweet and Snow (2003) further elaborate that, as some people are easily engaged in escape fiction and others are not, some also learn easily from informational texts while some have to reread, outline and take notes in order to benefit from such reading. Also, some can automatically note and respond aesthetically to stylistic variation in text, while others read primarily for content. Evidently, the match of activity with reader determines if the comprehension will be successful or not (p.3).

To put it simply, the activity in the diagram in Fig. 1. occurs at the intersections between the segments, and not within the segments as separate units. Different texts require different activities. For example, recipes are intended for immediate application, not for long-term learning unless in some cases when they are analyzed for elements of culture, etymology, etc. On the other hand, history textbooks are meant to be comprehended in order to generate long-term learning (Sweet and Snow, 2003).

These three elements are also interrelated in ways that change over time. The readers' skills at the beginning of reading may not necessarily be the same once they complete it. For instance, a reader who picks up a certain text possesses cognitive, motivational and linguistic capabilities. It is possible for these characteristics to change during the reading process, due to the fact that the reader can learn new vocabulary, or they can simply become more engaged in the topic, so they do additional research.

2.1.1. Reader

Taking into consideration that not all students have the same amount of motivation when it comes to reading, it is natural to predict that some will come to comprehension tasks more prepared than others. Sweet and Snow suggest that good comprehenders have a wide range of abilities, such as attention, memory, critical analytic ability, inferencing, and visualization ability (p. 4). Besides being more motivated to read, good comprehenders differ from not-as-good ones due to their interest in the content, and overall in what they consider their purpose for reading. Some other abilities mentioned by Sweet and Snow (2003) are various types of knowledge; vocabulary, topic knowledge, linguistic and discourse knowledge, to name a few (p. 4).

All of the factors mentioned above affect one's comprehension and whether it will be successful or not. If the reader possesses such abilities, their comprehension process will be less demanding than of those who do not. In addition, the more the reader is fluent in a language the text is written in, the better they will comprehend the text itself. As it was mentioned above, it is possible for certain changes to occur in readers' abilities during the reading process. Fluency is encompassed by that change as well, as readers' knowledge in domains such as vocabulary, grammar, and discourse increases, so does their fluency. Teachers have a significant role in helping students improve their comprehension, and they are able to do that by focusing on appropriate instruction and adjusting content to fit students' needs, all the while guiding them towards becoming autonomous readers.

2.1.2. Text

Apart from reader's capabilities, another important factor in the process of reading comprehension are the features of the text itself. If the text is structured in a way that is hard to understand, using complex vocabulary and sentence structure, and is based on information that requires background knowledge from the reader, it will be quite difficult to comprehend.

During the reading process, readers focus on a number of things, beginning with translating the words into their native language and visualizing the ideas presented in the text. Reading in a foreign language is a challenging activity that requires concentration and skills and reading texts that are written in a way that only confuses the reader more can only result in unsuccessful comprehension. For example, there is a case of metalanguage, which is by definition language used to talk about language. The term was originally used by a linguist Roman Jakobson in 1956 (p. 113). The language under study is called the object language and the language being used to make assertions about it is the metalanguage. When English speakers examine the English language, both the object language and metalanguage are the same (Nordquist, 2020). For EFL learners, using metalanguage for examining English can be a difficult task, as English is a foreign language for them, which complicates the process of learning.

As it was mentioned before, just because one understands the meaning of the words, it does not mean that they understand what they have read. A reader may be familiar with the content presented, however, if the linguistic structure and discourse style are more complex, they will face difficulties in comprehending it. If the text is more relatable to their own interests and common knowledge, it will be less challenging for them to grasp the meaning. This is why the relationship between the reader and the text is significant. Besides, the relationship between the text and the reading activity is significant as well. As it was previously mentioned, some types of texts require a specific reading activity to be applied, thus influencing the comprehension.

2.1.3. Reading Comprehension Activity

The third element of the reading comprehension is the reading activity. Reading is done for a certain purpose, and it occurs in different conditions. That purpose can be externally (e.g., completing a class assignment) or internally (e.g., wanting to program a VCR) motivated (Sweet and Snow, 2003). The purpose is influenced by one's motivation, interest, etc., and can be changed

during the reading process depending on previously mentioned variables. Of course, the success of comprehension depends on the purpose of reading. When one is doing the reading upon instruction by someone, they might approach it with the main purpose in mind being simply getting the task done. This can result in them completing the task mechanically without successfully comprehending the matter. On the other hand, while reading for personal enjoyment, the reader is more engaged and dedicated to the task, thus being more successful in comprehension. Namely, one can pick up a historical novel for the purpose of enjoying a book but learn a lot of interesting facts along the way, increasing their knowledge. The above-mentioned example of a recipe is a simple example of application consequence, and it is related to the initial goal of the reader. And last, but not least, engagement is the consequence of many other reading activities. An interesting text can keep a reader involved in reading even in inconvenient situations.

It is evident that reading activities can have long-term “consequences”, as any knowledge on how to do something, acquisition of new vocabulary and historical facts are all a result of reading with comprehension and is new knowledge that a reader brings to new reading experience. Reading comprehension is a significant skill that furthers the development of reader’s other skills that are beneficial for their overall improvement. It is also helpful while performing tasks such as analyzing texts in foreign language, explaining and elaborating on themes covered in said texts, but it also helps reader’s express their own opinions and give ideas about certain materials.

2.2 Verbal Protocols

According to Gass and Mackey (2000), verbal protocols are defined as the data one gets by “asking individuals to vocalize what is going on through their minds as they are solving a problem or performing a task” (p. 13). This is a data source that contains one’s spoken thoughts that occurred during working on a task. While working on a task, individuals vocalize their thoughts either immediately, or in intervals specified by the researcher. As it was mentioned by Leow and Morgan-Short (2004), “all verbal reports are not equal” (p. 36). There are different types of verbal protocols, and they differ in whether they reflect a think-aloud data, introspection, or retrospection (Cohen, 2013). In the case of think-aloud data, participants vocalize their immediate thoughts without any further explanation or analysis of what they are doing, while introspection leads to an implication that they report on what they think they are doing in order to accomplish a task.

These two terms can be distinguished by describing think-aloud as self-revelation, while introspection in contrast constitutes self-observation (Cohen, 2000). Consequentially, self-revelation is a more accurate representation of the actual thought-process. While think-aloud and introspection data is gathered while participants are performing a task, retrospective verbal reports are obtained after the completion of the task. We can consider the first two as more trustworthy, as the second one is gathered by memory, which is susceptible to modification. Below is a more detailed description of the three broad types of verbal report data (Cohen, 1987, 1991):

1. *Self-report*, where learners provide descriptions of what they do, characterized by generalized statements, for instance, about their language use strategies. For example: “When I read in an L2, I try to look for clues in the immediate context when I encounter words I don’t know.” Self-reports are retrospective and mentalistic in nature.

2. *Self-observation*, which is the inspection of specific, not generalized, language behavior, either introspectively (i.e., within 20 seconds of the mental event) or retrospectively. A further distinction would be between immediate retrospection (i.e., within an hour or so of the mental event) and delayed retrospection (say, up to a week or more after the event). Here is an example of immediate retrospection: “What I just did was to quickly skim the next several paragraphs to see if there were any clues as to the meaning of the word that I’m stumped on here. And sure enough, I found a handy synonym for that word, one that I did know.”

3. *Self-revelation*, typically referred to as think-aloud, is characterized by the stream-of-consciousness disclosure of thought processes while the information is being attended to. For example: “Hmm . . . I wonder if these three words are actually a paraphrase of that word in the previous paragraph that I didn’t recognize.” Self-revelation is introspective and non-mentalistic in nature and is seen as most accurately reflecting learners’ cognitive processes (Ericsson & Simon, 1993; Cohen, 2000). Researchers need to be mindful of the fact that verbal reports may comprise a combination of these different types (Radford, 1974; Cohen & Hosenfeld, 1981; Cohen, 1987), and so, if the intention is to use only one type, the data collected need to be classified based on the types mentioned.

Self-report data can be gathered by oral interviews or written questionnaires that prompt participants to elaborate on the way they learn language. They can speak about how they usually learn, or about how they perform what they have learned. In the case of self-observation, the data is a reference to an actual instance of language behavior, for example recollection of the process of L2 readers finding out what a certain word means in a given text. On the other hand, self-revelation or think-aloud is data that is available only at the moment when reading comprehension event is taking place. This means that the reader is sharing their immediate thoughts, as well as showing their struggle and difficulties by placing certain words in proper context.

2.2.1. Think–Aloud Data

Think-alouds are a form of verbal protocol (Ericsson & Simon, 1993) useful for examining the mental strategies (e.g., processes) in which participants engage. This data is comprised of stream-of-consciousness responses and spontaneous behavior, as the researcher requests from the participant to verbalize what they are thinking at an exact moment. The advantage of think-aloud data is the fact that it provides the most accurate information regarding readers' working memory during task execution. They differ from introspection and retrospection as they do not require from participants to dwell on the cognitive process of task completion, but to only voice their initial thoughts while doing their task which allows the researcher a better view of the mental process and the struggles that the reader faces.

Researchers can use think-aloud data for various purposes and with different goals. Collecting think-aloud data is rather simple, and the only requirement is that the participant has the ability to speak. The research can examine different variables, depending on the purpose of the study, and it can be done using different strategies. The think-aloud data is recorded and transcribed. In order to avoid any bias, the researcher is supposed to refrain from providing any kind of feedback, other than encouraging participants to verbalize their thoughts at a specific moment.

The analysis of the think-aloud data is done by coding, that is, applying labels to the verbalization that align with categories of interest. Furthermore, there are two ways of analyzing the transcripts: inductive and deductive. In the inductive method, codes are decided based on thorough reading of transcripts with no prior notion of what the categories might be. On the other hand, the deductive method is when a researcher has a pre-established plan for coding based on the research goal.

Bohn-Gettler and Kendeou (2014) were interested in techniques that lead to good comprehension, thus they used a deductive method based on previous research.

When it comes to participants' response, it can contain multiple thoughts and ideas. In this case, it is important to determine how to phrase each statement. The term "parsing" in this case refers to instances when a researcher splits a participant's response into shorter statements in order to be able to assign codes more specifically. There are multiple possible ways of assigning codes. First, the researcher could code the entire response with just one code that represents the majority of the responses, focusing on the dominant process that the participant used. Also, the researcher could code only the first idea of the response, addressing the first, initial thoughts of the participant. And the third way of assigning codes is by parsing the participant's response into multiple shorter statements and coding each idea in the response separately. Bohn-Gettler and Kendeou (2014) parsed each response into ideas or events, which usually represented subject-verb phrases.

Once the codes are assigned, the researcher can move on to the analysis. As participants read multiple texts they give different answers, meaning that one participant may have given a one-word answer while another may have stated a few sentences. The comparison between participants can also be done in multiple ways. The researcher could calculate the total number of each code that the participants provided and divide that by the total number of codes in the entire text. This would result in percentages of how often each participant engaged in each process. Secondly, statistical analyses could be performed on each category separately, or related categories could be combined.

As can be concluded, think-aloud data is a very specific and thorough analysis of one's thinking process, and it is of great help in examining how students process different types of texts.

3 RESEARCH FRAMEWORK

In this research part of the paper, the aim is to determine factors that affect a better understanding of a text, or rather the content in English as foreign language. Founded upon the theoretical framework of studies in psycholinguistics and reading comprehension, as well as the process of obtaining think-aloud data, this research is concerned with examining which type of text, out of the three types given, do the participants comprehend most successfully in a given timeframe. The participants will be asked to paraphrase the excerpts they have read, and the recording will be transcribed and analyzed according to the process mentioned above and coding scheme presented below.

Code	Definition	Text excerpt	Participant's response
Paraphrases	Restating the sentence in different words	Colleges and universities could develop programs aimed at training female politicians.	<i>Colleges could create training for female politicians.</i>
Connecting inferences	Referring to textual information presented in the immediately preceding sentence	Sentence 2: Instead of fumbling for change, they drive right through. Sentence 3: The device sends a radio signal that records their passage.	<i>So that people don't just drive through without paying...</i>
Reinstatement inferences	Referring to textual information presented earlier than the immediately preceding sentence	Sentence 12: Airplanes regularly carry several hundred passengers. Sentence 15: They usually carry only two to five passengers.	<i>It's smaller than an airplane.</i>
Elaborations	Utilizing prior knowledge to help explain what was just read	Virginia has reduced its minimum age to 14.	<i>... and they wouldn't be able to punish them, unless they were tried as adults.</i>

Predictions	Anticipating what will occur next	... and they wouldn't be able to punish them, unless they were tried as adults.	<i>Uh sounds like the confederacy will soon be formed.</i>
Opinions	Providing an opinion about what was read	Virginia has reduced its minimum age to 14.	<i>Uh, that seems really low to be tried as an adult.</i>
Statements of accuracy	Statements reflecting not understanding or not knowing	On December 23, 1947, three scientists at Bell Telephone Laboratories demonstrated their new invention of the point contact transistor amplifier	<i>I don't know what that is.</i>
Associations	Stating information and/or prior knowledge and experiences that do not serve to increase understanding of the text	President Eisenhower sent federal troops and the National Guard to intervene on behalf of the students.	<i>Um, this randomly reminded me of Remember the Titans.</i>
Other	Responses that do not fit any other category, and non-responses	In September, formerly all-white Central High School learned that integration is easier said than done.	<i>Integration ... so yeah.</i>
Invalid statement	Statements reflecting a misunderstanding of the text. (Coded in addition to at least one of the other codes above.)	In 1950, President Truman ordered the Atomic Energy Commission to make the hydrogen bomb.	<i>January is 1850 ...</i>

Table 1. Think-aloud codes and examples, Bohn-Gettler and Kendeou (2014)

3.1. Participants

The total number of individuals who participated in this study is 10. The participants are students enrolled in master's degree programs at the Department of English Language and Literature (9

students are enrolled in MA in Teacher Education and one student is enrolled in MA in Linguistics program). All of the participants attended the course the textbook of which has been included in the verbal protocol research using familiar text.

All of the participants are native in the Bosnian language and speak English as a foreign language. The study was conducted via Zoom and recorded using an online website online-voice-recorder.com (see Appendix 2), as the recording of the participants' response is the data that represents the verbal protocols. The voice recording of each participant has been saved but the participants' identities and personal information has been anonymized. The participation in the research was voluntary.

3.2. Instrument

As the aim of this research is to determine factors that affect better understanding of a text in English as a foreign language based on three paragraphs used as sample material. The participants are first shown the first paragraph and given one minute to read it for themselves. Then they are asked to paraphrase what they have read. After they paraphrase the first paragraph, they will be shown the second paragraph, and then after that, the third one. The procedure is the same for all three paragraphs/excerpts. All the while, online-voice-recorder is turned on, documenting the thought process of the participants. To an extent, they are familiar with the texts used for research, since the excerpts are taken from *The Articulate Mammal*, a textbook studied during the course in Psycholinguistics that the students attended in their fourth year. In this way, it was ensured that they were all working with texts they had encountered before.

The first excerpt that the participants of this research were asked to read is a rather common type of text, comprising high frequency words. The second paragraph is a definition that is written in a rather complicated manner containing abstract terms and more formal vocabulary. The writing style is significantly different from the first type of the text, making the comprehension more difficult. The third excerpt explains something by using a comparison and vocabulary that students are familiar with.

4 ANALYSIS AND INTERPRETATION OF RESULTS

The data gathered in the process of the research has been analyzed via using strategies developed and used by Bohn-Gettler and Kendeou (2014). The goal of the research was to use verbal protocols in order to determine which type of three given texts is the most difficult for students to comprehend, and which one they understood without difficulty and why.

As was mentioned before, the transcribed data first must be coded, and the results are much more accurate if the researcher is using the deductive method, meaning that the coding is done with a pre-established plan that aligns with the researcher's goals. The text excerpts comprised between 120 and 140 words, and the think-aloud data contained multiple think-alouds and mental processes. In order for results to be more specific, the process of assigning codes was not limited only to successful and unsuccessful paraphrasing, but the participant's response was first divided into smaller ideas and thoughts through the so-called parsing, as mentioned earlier. In this way, we have a more detailed insight in the overall process, and it is possible to combine the results.

The coding was conducted after the parsing stage. Each think-aloud was assigned a code that fits that idea the best. In Table 1. we have seen codes and their definitions, and the same table was used as a coding scheme for this research. It is important to note that Bohn-Gettler and Kendeou (2014) wanted to reduce the number of analyses and considered the fact that some codes are co-related and can be combined in larger categories.

After the transcription and assigning codes, the analysis was focused on four categories Bohn-Gettler and Kendeou (2014) created. After participants' responses, a table with total number of codes for each category is presented. Appendix 1 at the end of this paper contains a detailed table containing participants' responses divided into codes and codes sorted into above-mentioned categories.

4.1. EXCERPT 1: *Waggle Dance*

As stated above, the first excerpt that the participants of this research were presented with comprises high frequency words:

When a worker bee finds a source of nectar it returns to the hive to perform a complex dance which informs the other bees of its location. It does a 'round dance', which involves turning round in circles if the nectar is close to the hive, and a 'waggle dance' in which it

wiggles its tail from side to side if it is far away. The other bees work out the distance by noting the tempo of its waggles and discover what kind of flower to look for by smelling the scent on its body. Bees, incidentally, are not deaf, as was once assumed. As a forager bee dances, it beats its wings. The bees in a dark hive can hear and interpret the wingbeats even when they cannot see the dance. (Aitchison, 2008, p. 31)

The participants' responses are given below:

Participant 1	Uhm, okay... the... I read the uhm... small excerpt regarding the uhm... bee dance uhm... and how they, uhm... they usually start to uhm... have this uhm... wiggle dance, I think uhm..., which also, uhm with the way they notify the other bees to come and collect the nectar uhm... They have this kind of uhm... circling around, uhm... Maybe spinning around the flower, or uhm... and so uhm... pretty much the... the excerpt is uhm... how bee do the dance in collecting the nectar.
Participant 2	The text is about a ... uh is about uhm... working bees. So, when a working bee uhm... finds uhm... nectar, a place uhm... where they can collect nectar, this bee comes to hive and uhm it dances and if, uhm... I believe if the nectar is close to the hive, then she... it does this round dance and if uhm... if the nectar is far from the hive it does this waggle dance. Previously it had been believed that uhm... bees are deaf, but they are not since they can hear, uhm, since they can hear moves in the hive even if the hive is dark.
Participant 3	Basically, when a worker bee finds a source of nectar it returns to the hive and it does a dance to inform the other bees of where it is. It can do a round dance if its close by or a waggle dance if it's farther away. Um... depending on how, like the dance goes, the bees can tell exactly how far away it is. And, they don't necessarily need to see it, because bees can hear the beat of the worker bees wings as its dancing, and that can also help them figure out even if they cannot see the dance. so that's it.
Participant 4	Uhm... basically as once it was thought that bees could not hear, that they are deaf, but uhm... in fact that is not true because bees can communicate uhm... with their bodies. Uhm... basically if they find nectar, if one bee finds the nectar, she comes to other bees and uhm... she does a dance which indicates where and how far the nectar is and other bees understand her uh... in means that they can hear her wings flapping. Basically, if the nectar is close uh..., she would do the round dance, and if the nectar is not that close she would do the waggle dance. And the bees would recognize where the nectar is, which flower the nectar is from by smelling uhm... the bee. And basically, bees can communicate uhm... this way.
Participant 5	Okay, so uhm... when bees find a source of nectar, they go back to their hive and perform a complex dance. They either do a round dance or a waggle dance. Uhm... in case of round dance, they do that kind of dance uhm... if the nectar is near, and in waggle dance in which they shook their tails side to side so if the source is far away. The bees in the hive can judge distance by the tempo and the type of flower by the scent of the bee. Uhm... previously bees have

	been perceived as deaf, but they aren't. They can produce like a sound, the way they beat their wings... I think that's it.
Participant 6	So, basically the text is about bees. And the text says that when bees find the nectar suiting for them, they um... they... uhm... The bee comes to their friends and coworkers and informs them by performing a dance. This was dancing in circles if the nectar is near... near them. This can happen even in the dark. So basically bees have a let's say navigation system that allows them to collect the nectar.
Participant 7	Alright, so this text is about bees. So basically, when bees find nectar, they are going to do their wiggle dance... uhm... to inform other bees that nectar is near... so, if they are wiggling rapidly, that means that nectar is uhm... very close so the other bees are going to mimic this so-called dance they're going to inform other bees in the hive where the nectar is.
Participant 8	Okay so it's a text about how bees communicate and that they aren't really deaf as it was once thought. Uhm, once a bee finds a nectar source it returns to the hive and it does either a round dance if a flower or nectar is nearby, or a waggle dance if the nectar or flower is further away. That's what I gathered.
Participant 9	So, uhm... it has been proved that bees are actually not deaf and that uhm... when a bee which is responsible for looking for honey, when she finds honey she informs her... she returns back to the hive and informs her bee friends, uhm... that she has found a bee and she has two ways of doing it. One is by round dance which signifies that the bee is close to them and the other one is the wiggle dance or something like that, uhm... which signifies that the bee is far away. And the other bees... uhm... the... recognize the distance by analyzing the tempo of her wiggles and uhm... they find out which flower to look for... by... by looking at the scent of her body. Basically, that's it.
Participant 10	Well, this is interesting text. It's about bees and them wagging their tails in order for other bees to find the nectar and flowers. Uhm... as once was assumed, bees are not deaf so they perform a dance in order to tell other bees where its the flower and where is the nectar they need to find. And also uhm... there was one interesting thing I wanted to remember but now I forgot it. Ah yeah, by the way they're wagging their tail, they show other bees how far is the flower and how far is the nectar. I think that was the most important part of the text.

Table 2. Verbal Protocol Responses to Excerpt 1

The four categories that Bohn-Gettler and Kendeou (2014) created are paraphrases, text-based inferences, knowledge-based inferences and non-coherence processes. As it was mentioned above, the participants' responses have been parsed into smaller chunks, which were later categorized according to codes they correspond with. In the table below, the results of the coding are presented:

TOTAL NUMBER OF CODES	117
Paraphrases	86
Text-based inferences	18
Non-coherence processes	13

Table 3. Total Number of Codes for Excerpt 1

In Fig.2, we can see a chart of the results for the first excerpt:

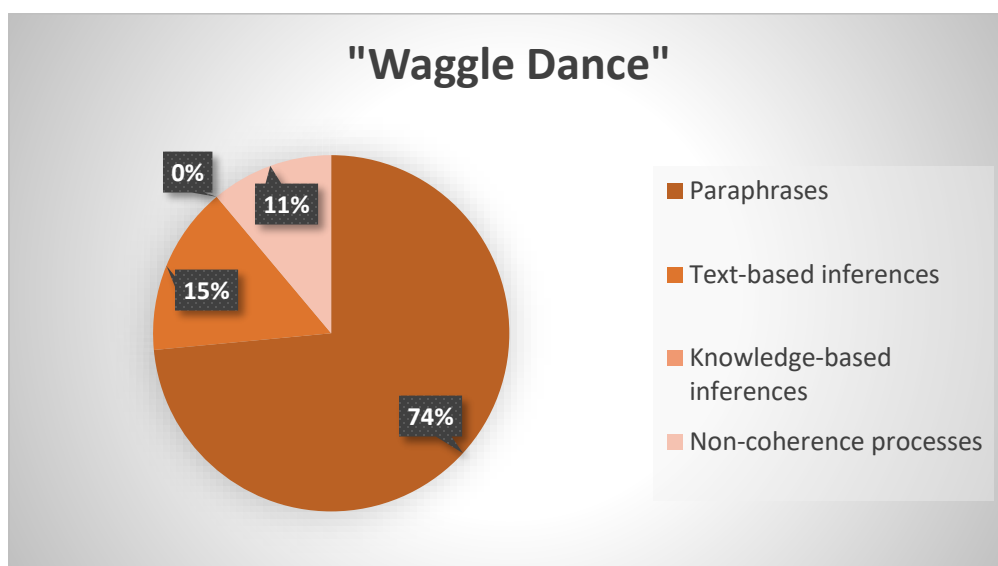


Fig. 2. *Waggle Dance* Coding Results

As for the most frequent strategy, paraphrases help with the memory of the text, as the process requires restating a certain sentence using different words. For example:

Excerpt from the text: It does a ‘round dance’, which involves turning round in circles if the nectar is close to the hive, and a ‘waggle dance’ in which it wiggles its tail from side to side if it is far away.

Participant’s response: /It can do a round dance if it’s close by or a waggle dance if it’s farther away. /

This essentially means that if one wants to paraphrase something successfully, they have to recall the matter from their memory, and not everyone is able to do this without extensive reading or studying. That is where the issue of the type of a text one is reading comes to play. This chart is evidence that if a text that is written using simple vocabulary and word phrases that students are familiar with, they will remember that text fairly quickly and comprehend the essence of the excerpt without many difficulties. Out of the total number of think-alouds in this text excerpt, which is 117, paraphrases take up 74%. It is the highest percentage out of all three paragraphs, and we can conclude that this type of text has content that is the least difficult one to sum up and paraphrase.

The next category presented by Bohn-Gettler and Kendeou (2014) is the so-called *text-based inference*, and it refers to statements that are related strictly to the text itself. This category combines codes *connecting inferences* and *reinstatement inferences*. Connecting inferences are codes that refer to information presented in the preceding sentence, while reinstatement inferences are codes that refer to information presented earlier than the preceding sentence.

As an example of connecting inferences code in this text excerpt, we have the following:

Sentence 4: Bees, incidentally, are not deaf, as was once assumed

Sentence 5: As a forager bee dances, it beats its wings.

Participant's response: /Uhm... previously bees have been perceived as deaf/, but they aren't/. They can produce like a sound/, the way they beat their wings. /

Out of 117 think-alouds for this excerpt, text-based inferences hold 15%, or a total of 18 codes, which is essentially not a high number. However, participants greatly succeeded in paraphrasing, which makes this type of text the one that was comprehended most successfully. Furthermore, Bohn-Gettler and Kendeou (2014) formed a third category that is a combination of elaborative and predictive inferences. These codes require activating prior knowledge in order to understand the text, and assuming what will happen next. Despite the fact that students were familiar with the textbook and this particular text, out of 117 codes, there were none that could have been sorted in this category.

The last on the list is the fourth category by Bohn-Gettler and Kendeou (2014). In this category are the processes that are not quite likely to help the participants with their task. These are phrases used for stating uncertainty, unelaborated opinions, and making loose associations. The results in the Fig. 2. show that the number of non-coherence processes is quite low; 13 codes, or in other words: 15%.

Here is an example of these codes in the participant's response for this text excerpt:

Excerpt from the text: When a worker bee finds a source of nectar it returns to the hive to perform a complex dance which informs the other bees of its location. It does a 'round dance', which involves turning round in circles if the nectar is close to the hive, and a 'waggle dance' in which it wiggles its tail from side to side if it is far away.

Participant's response: /They have this kind of uhm... circling around, /uhm... maybe spinning around the flower, / or uhm... and so uhm... / pretty much the... the excerpt is uhm... how bee do the dance in collecting the nectar. /

After assigning codes, the ones that signify long pauses, broken-up thoughts that make no sense, and unfinished statements are all sorted in this category as well. The percentage of this category shows us the general picture of how difficult the text was for participants to comprehend. In this case, that percentage is very low, which brings us to a conclusion that the comprehension of this text excerpt was successful.

4.2. EXCERPT 2: *Language Universals*

The results of the coding for Excerpt 2 and Excerpt 1 differ vastly. The second excerpt (i.e., Excerpt 2) is given below, as well as the participants' responses.

Language universals, Chomsky suggested (1965), are of two basic types, substantive and formal. Substantive universals represent the fundamental 'building blocks' of language, the substance out of which it is made, while formal universals are concerned with the form or shape of a grammar. An analogy might make this distinction clearer. If, hypothetically, Eskimos were born with an innate knowledge of igloo-building they would have two kinds of knowledge. On the one hand, they would know in advance that the substance out of which igloos are made is ice and snow, just as thrushes automatically know that their nests are made of twigs, not bricks or worms or glass. On the other hand, their innate knowledge

of igloo-building would include the information that igloos are round in shape, not square or diamond-shaped or sausage-like, just as thrushes instinctively build round nests, not ones shaped like bathtubs. (Aitchison, 2008, p.101)

The participants think-alouds for Excerpt 2:

Participant 1	Uhm... okay... a little... this excerpt is a little uhm... demanding... uhm... It's regarding the uhm... Chomsky's hypothesis, uhm... and innate uhm... knowledge. Uhm... so uhm... it uhm... it is based on the research regarding... uhm... not sure if it's research from him, but so... it's about uhm... two way thinking, uhm... if... Eski... so... Eskimos can... If their innate knowledge is uhm... two sides. So one would be that they uhm... build their igloos from uhm... ice and snow and another would be regarding that uhm... it's the round shape uhm... so uhm... not sure... So, I mean interesting fact, uhm... Something new for me to learn. But definitely, uhm... we would have this uhm... two way of thinking... so we need to... uhm... to split... uhm... from... from... what the igloo is made, and what... what... kind of shape it is, and what is... uhm... what is..., what is its uhm... main... what is its main uhm... concept, I mean uhm... it's pretty much house for Eskimos...
Participant 2	So, according to Chomsky, there are two types of language universals, substantive and informative. Substantive universals are the ones that present building blocks of language and uhm... formative—formative universals are concerned with form and shape of grammar and there... there was this analogy presented, this distinction between substantive and formative universals uhh was uhm... was compared to the knowledge that is innate to Eskimos. For example Eskimos... if Eskimos were born with two types of knowledge that would be... substantive knowledge would be that they know that igloos are made of uhm... made of snow and ice, and formative knowledge would be that they know that igloo is round in shape and not like sausage like, or whatever...
Participant 3	Uhm... so... Chomsky said that language universals are basically divided into two categories. Substantive and formal. Substantive are those that are considered the building blocks of language or what its made of, and formal are the words which tell us like how to make formations out of those building blocks or at least how grammar is formed. An analogy can be made, or a sort of a metaphor if we say that Eskimos have an innate knowledge of how to build igloos, then you would say that they have two types of knowledge. The first one is the knowledge that igloos are made of ice, and the second is the knowledge that like, that there are a round shape, that how to shape the ice. In the same same way that uhm... it was mentioned at the end uhm... thrushes birds, they do round nests and not nests in the shape of a bath.
Participant 4	Uhm... To start with... Basically, there's... uhm... Chomsky talks about how there is some substantial and fundamental knowledges, and the first thing is that some knowledge is innate, which means for example if Eskimos were to be born with innate knowledge of knowing how to build uhm... igloo, they would know uhm... that they have to do it, but on the other hand they would also be born with knowledge that the igloos are built round and with ice... That's it.

Participant 5	<p>Okay, this one is a bit more difficult... Chomsky suggests that two... Ooh... okay, oh my God... uhm... I forgot.. So, we have... oh Oh God, ... okay okay... I'll just say how I remember. There's substantive and forma... formative grammar... In which uhm... this substantial consists of building blocks of language so the thing that its made of, and the formative is, ... oh it's formative... The formative is about the form of grammar... so analogy used are Eskimos when they build igloos. So they know the substance of igloos they are building, which is ice, but they don't know... Wait, wait, no, no. They know it's built from ice, but they don't have the innate knowledge from when it comes to building its shape, the same way thrushers know that their... oh God, this one is difficult... aa... so basically igloos know that... Eskimos know igloos are made out of ice, but they don't know, they don't have an innate knowledge about shape. Okay, this one was difficult.</p>
Participant 6	<p>So the second text is about meta language and uhm... according to Chomsky there are two basic types of language universals, formal and substantive. Formal.. uhm... universal is concerned with form and shape of the grammar, while substantive is uhm... about fundamental building blocks of the language. And uhm... the example given was of the Eskimos and how they build igloos. According to language universals uhm... Igloos... if we apply the uhm... substantive form, Eskimos were born with the innate knowledge on how to build an igloo, and which material they should use and how the shape would be. So, basically that's it. Well, you know on the other hand, formal universals would be that they have... that they had to have been thought how to build an igloo and what material to use and how the igloo would be.</p>
Participant 7	<p>Alright, so this one is about Chomsky uhm... and him saying that there is a substantive and formative type of something... uhm... so he gave us an example of Eskimos, of their let's say innate uhm... language abilities but uhm... he brought it in a way that if Eskimos had innate knowledge of building igloos uhm... they know that they're built from ice, but on the other hand uhm... they do not, I think that they do not know which shape it needs to be, does it need to be square, rounded, or any other shape.</p>
Participant 8	<p>Uhm... it's a text about language universals presented by Chomsky. Two types, formal and I'm not sure what the other one's called. Uhm... Formal one deals with... I have no idea. I remember the uhm... Eskimos. It deals with innate knowledge of language of whether or not humans are born with some innate language knowledge, I know that. Chomsky gives an example of Eskimos and how they build igloos. If they had innate knowledge, they would know that igloos are made out of ice and in particular shape, so I think... circular. That's what I remember.</p>
Participant 9	<p>Uh, well according to Chomsky, language universals are of two types, substantive and um... one more, I forgot. Uhm, might be fundamental, or something like that. Anyways, the first one is concerned with the uhm... building blocks of language, while the other one is uhm... more like uhm... uhm.... has to do more with grammar and he made an analogy with Eskimos and their igloo building and if the knowledge of igloo building for them was innate, they would know that, it was one of two kinds of knowledge, one would be</p>

	similar to substantive, which would suggest them that their igloos are made of snow and ice and the other one would be fundamental... and I forgot...
Participant 10	This one was a little bit more difficult. But, it was about language universals. Which was proposed by Chomsky. He said that there are two types, formal and substantive, or substantial, I'm not sure. One of them represents the building blocks and the other represents, in a way, form and shape how the language is formed and the sentences. He used Eskimos and them building uhm... an igloo. Uhm... it's like they have two kinds of knowledge. Uhm... one of the... one type of knowledge helps them to build uhm... the igloos which is in a round shape, and the other... Honestly, I don't remember. Yeah, and uhm... they, yeah... And they have some kind of innate knowledge which helps them to shape the igloo in that shape in which it is. I think that it was something like that.

Table 4. Verbal Protocol Responses to Excerpt 2

Table 5. shows the overall number of codes for this excerpt:

TOTAL NUMBER OF CODES	172
Paraphrases	77
Text-based inferences	18
Non-coherence processes	77

Table 5. Total Number of Codes for Excerpt 2

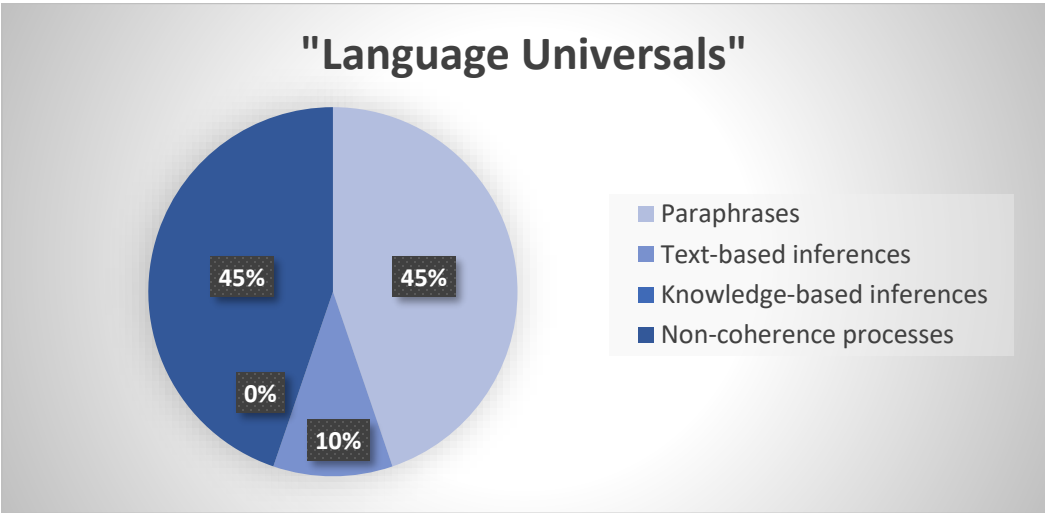


Fig. 3. Language Universals Coding Results

First of all, it is important to emphasize that we have the same percentage for two categories: Paraphrases and Non-coherence processes. These two groups are very different, and it is a bit of a surprise that they are the ones that contain the same number of codes.

For example, we have a few instances of the paraphrasing task done successfully, or partially so:

Excerpt from the text: Substantive universals represent the fundamental ‘building blocks’ of language, the substance out of which it is made, while formal universals are concerned with the form or shape of a grammar.

Participant’s response: /Formal... uhm... universal is concerned with form and shape of the grammar/, while substantive is uhm... about fundamental building blocks of the language. /

Excerpt from the text: On the one hand, they would know in advance that the substance out of which igloos are made is ice and snow, just as thrushes automatically know that their nests are made of twigs, not bricks or worms or glass. On the other hand, their innate knowledge of igloo-building would include the information that igloos are round in shape, not square or diamond-shaped or sausage-like, just as thrushes instinctively build round nests, not ones shaped like bathtubs.

Participant’s response: /Chomsky gives an example of Eskimos/ and how they build igloos. / If they had innate knowledge, they would know that igloos are made out of ice/ and in particular shape,/ so I think... circular./

There are just as many unsuccessful instances, many containing unelaborated opinions and statements of uncertainty:

Excerpt from the text: On the one hand, they would know in advance that the substance out of which igloos are made is ice and snow, just as thrushes automatically know that their nests are made of twigs, not bricks or worms or glass. On the other hand, their innate knowledge of igloo-building would include the information that igloos are round in shape,

not square or diamond-shaped or sausage-like, just as thrushes instinctively build round nests, not ones shaped like bathtubs.

Participant's response: But definitely, uhm... we would have this uhm... two way of thinking... /so we need to... uhm.. to split.. uhm... from... from.. what the igloo is made/, and what... what... kind of shape it is/, and what is.../ uhm... what is.../, what is its uhm.. main.../ what is its main uhm... concept/, I mean uhm... it's pretty much house for Eskimos.../“

Excerpt from the text: Language universals, Chomsky suggested (1965), are of two basic types, substantive and formal.

Participant's response: /Okay, this one is a bit more difficult.../ Chomsky suggests that two... /Ooh... okay, oh my God.../ uh... I forgot... /So, we have... /oh... Oh... God,.../

After further analysis of this excerpt has been finished, the conclusion that can be drawn is that we have two opposing situations. The participants either understood the excerpt and were able to paraphrase it, or they failed to comprehend the information presented, which resulted in a significant number of incoherent statements.

Out of the total number of 172 codes, there are noted only 18 codes that can be sorted into the category of the text-based inferences. For example:

Sentence 5: On the one hand, they would know in advance that the substance out of which igloos are made is ice and snow, just as thrushes automatically know that their nests are made of twigs, not bricks or worms or glass.

Sentence 6: On the other hand, their innate knowledge of igloo-building would include the information that igloos are round in shape, not square or diamond-shaped or sausage-like, just as thrushes instinctively build round nests, not ones shaped like bathtubs.

Participant's response: /...if Eskimos were born with two types of knowledge/ that would be... /substantive knowledge would be that they know that igloos are made of uhm... /made of snow and ice,/ and formative knowledge would be that they know that igloo is round in shape/ and not like sausage like/...

Low number of text-based inferences is another piece of evidence of low comprehension of this text. Students were unable to connect multiple statements into one coherent thought. It is important to mention that even think-alouds that were sorted into this category were not completely successful in their intent, as they were borderline paraphrasing of multiple ideas. However, they were closer to connecting inferences rather than paraphrasing. Namely, the essence of an idea from two consecutive sentences was paraphrased into one statement. Hence why it can be considered a connecting inference instead of a mere paraphrase. Nonetheless, the overall number of codes in this category show how difficult it was for the participants to follow the text flow and connect its ideas.

Taking into consideration these results, it can be concluded that a text written in this manner is rather challenging for students to comprehend as they get confused and lose track of thought.

4.3. EXCERPT 3: *The Hearer*

The third excerpt that the participants of this research were presented is given below:

At the beginning of the twentieth century, psycholinguists assumed that the process of understanding speech was a simple one. The hearer was envisaged, metaphorically, as a secretary taking down a dictation. She mentally wrote down the sounds she heard one by one, then read off the words formed by them. Or, taking another metaphor, the hearer was envisaged as a detective solving a crime by matching fingerprints to known criminals. All the detective had to do was match a fingerprint found on the scene of the crime against one on his files, and see who it belonged to. Just as no two people's fingerprints are the same, so each sound was regarded as having a unique acoustic pattern.(Aitchison, 2008, p.206)

The participants' think-alouds for Excerpt 3 are as follows:

Participant 1	Okay... uhm... Okay, I managed to read it uhm... twice, but uhm... I still couldn't... uhm... remember as... as much as I could. So, it's... it's... regarding the research that scientists done in twenties. Uhm... and it connected uhm... somehow with uhm... investigating uhm... Actually, they tried to, uhm... what they hear uhm they tried to put it on the paper, uhm... any sound I suppose, uh, even saying A, or... or... any articulated sound, they tried to put it on the paper uhm... and then read it. Uhm..., I don't... I'm not sure if that is, but I assume it is, because they investigated language, I think. Uhm..., so its also uhm... mentioned something regarding this detective search uhm... and uhm... how they uhm... managed to uhm..., the tried to... with the fingerprints uhm... to
---------------	---

	uhm... find uhm... I think uhm, ... find something uhm... crime scene.. not sure.. I didn't catch it and figure it out what it's for... This one is the hardest...
Participant 2	It is... At the beginning of the twentieth century psycholinguists assumed that the underst... that the process of understanding speech was a simple one and there were two comparisons. First one said that they compared uhm... the hearer to a uhm... secretary who would just be taking down notes while a person is speaking and then uhm... after that person is done speaking uhm... the hearer would just read off the... uh... the words that she or he took down and then uh... understand them. And the second comparison was that hearer... hearer was envisaged as a uhm... detective who would uh... match uhm.. fingerprints found on a crime scene to .. uhm to all the persons in his file and then find out who.. whose fingerprints match the person in his file and in that way they compared the words of a hearer to a hearer.
Participant 3	At the beginning of the twentieth century they thought that the process of hearing and understanding speech was a simple one. Basically, the hearer would, uhm... it's like a person taking dictation, like you would simply record the words mentally and then read them off one by one. Uh.. Alternatively, it can also be described as like a... detective, where they just found fingerprints at the scene of the crime and then match the fingerprints with files they already have. In other words, each sound would have a unique corresponding word, as far as I understand.
Participant 4	Uhm... Psycholinguists uh... at the beginning of the twentieth century saw speech understand,... us people understanding speech as something not so complex but yet simple. Uh.. and uh... to describe .. to describe how they saw it.. uh... they saw the hearer, the person who hears something as a detective who matches the fingerprints to a person it belongs to. Basically that's how we, uh.. how we ... uh.. basically that's how people.. uh... when they hear a sound that's how they uh... recognize it and then later use it. Just as detectives can pinpoint that one fingerprint to the person it belongs to. I would say that.
Participant 5	Okay, so in the twentieth century psycholinguists assumed that um.. aa.. language... a... uhm... that... oh God... that language... uhm... actually... um... so basically was simple. It was envisioned as a secretary taking down a dictation and then after she read the words one by one. Uhm... the same... it was also as a detective who finds a fingerprint on a crime scene and then matches the fingerprint to the cases he has. Um... oh God.. so, each sound... each sound is.. uh basically each word has a unique acoustic sound.. That's it.. This was.. My brain is kind of lagging.
Participant 6	The last text is about a comparison and it's about speech. And the appliance of pyscholinguistics and how we understand speech and it basically uh... even... every word every sound, we tend to wrote... write mentally down and its compared to the job of detective, to fingerprints. Meaning that every sound is like every fingerprint, unique, different from other fingerprints or sounds. So basically every sound we use is different from the others.
Participant 7	Alright, so uh.. the text is about psychologists who thought that uh... either letter or songs were simple, so they envi... envi... envisaged, they saw them as people who could... as uh... I'm sorry... Uh... Can we do this again, please... So, they

	saw people for example as detectives who would find clues uh.. like.. fingerprints uh as every fingerprint is not the same, so each sound is not the same, it has uh.. they are different. Uhm.. Just the first one, I can't remember the word. It's dictation.. uh.. The person who writes down the words. .. uh so they write each and every word, and when they write them they memorize them so they are able to pronounce them again.
Participant 8	This text deals with the understandment... understanding of speech and how psycholinguists perceive it. Uhm, in one example they say its like, that a hearer is like a some type of a secretary that basically mentally writes down whatever a speaker says and then uhm... pairs certain words to other to form meaningful uh.. sentence. Another example would be that the hearer is the type of a detective and that different sounds uhm... can be attributed to like fingerprints, and no fingerprint is uhm.. identical so different sounds mean different word patterns and all the hearer has to do is match those patterns to form uh... meaningful sentences.
Participant 9	So, at the beginning of the twentieth century uhm... psycholinguists assumed that the process of uh...recording...processing speech is simple one and they compared hearer to a uhm... secretary at who someone dictates the words and she simply uhm... puts letter uhm... puts sounds one by one and forms a word or they would compare a hearer to a uhm..., detective who is decoding fingerprints. Just as each fingerprint is different, so are the sounds and the only thing a hearer had to do was to match the fingerprint found on the place of the crime to the fingerprints available and it was like a game of matching.
Participant 10	Well, this one was interesting also. Uhm... it was about psycholinguists assuming that language and understanding of a language was a simple one. They used metaphors. The first one was about hearer being a secretary writing down all the sounds they hear and then in a way uhm... decoding them. And then the other metaphor was a detective solving a crime uhm.... when they match fingerprints they find on a crime scene to a person to whom it belongs. And, the last sentence was an interesting one where they say that there are no two... there are no two the same fingerprints as there are no two same acoustic sounds. Something like that “

Table 6. Verbal Protocol Responses to Excerpt 3

Table 7. shows the overall results for this excerpt:

TOTAL NUMBER OF CODES	155
Paraphrases	80
Text-based inferences	15
Non-coherent processes	60

Table 7. Total Number of Codes for Excerpt 3

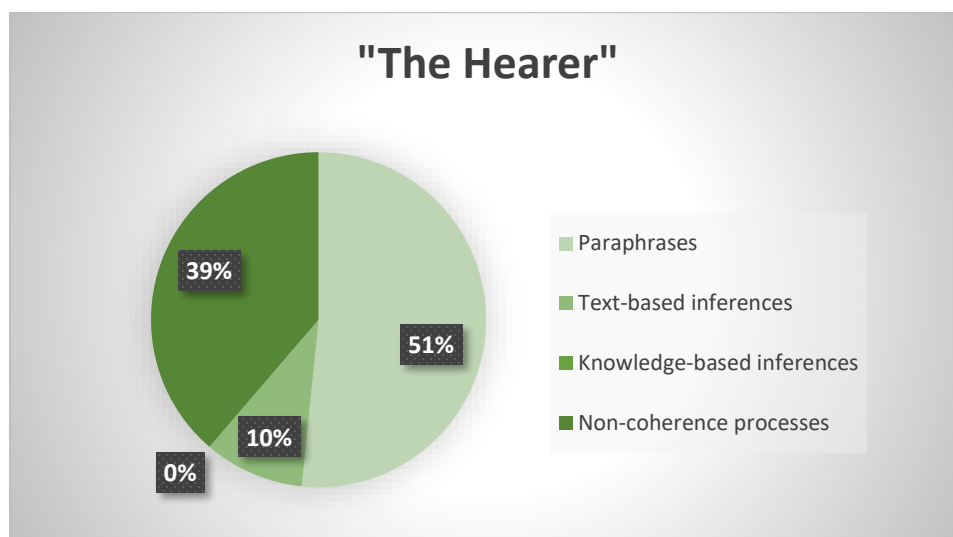


Fig. 4. *The Hearer* Coding Results

The third text presents new information by using metaphorical comparisons in conveying new information, which is why it should be easier for the participants to understand this excerpt than the previous one. It was unexpected that the percentage of the non-coherent processes is not as low as it was assumed it would be. The vocabulary used in the text is common, and it is a surprise that 39% of the participant's thoughts were non-coherent processes. Here are some of the examples from the data collected:

Excerpt from the text: Or, taking another metaphor, the hearer was envisaged as a detective solving a crime by matching fingerprints to known criminals. All the detective had to do was match a fingerprint found on the scene of the crime against one on his files, and see who it belonged to. Just as no two people's fingerprints are the same, so each sound was regarded as having a unique acoustic pattern

Participant's response: /I think. Uhm.../, so it's also uhm... mentioned something regarding this detective search uhm.../ and uhm.. how they uhm.. managed to uhm../ the tried to... /with the fingerprints uhm... /to uhm... find uhm... I think uhm,.../ find something uhm... crime scene.../ not sure.../ I didn't catch it /and figure it out what it's for.../ This one is the hardest.../“

In the participant's response, there is an obvious struggle to find the connection between the metaphor of a detective who matches fingerprints from a crime scene to the ones in his file and the hearer who is in the process of understanding language or that sounds are compared to. The results show that a significant number of participants was unable to comprehend this comparison.

However, more than a half of participant's responses were successful paraphrases. In the following example we can see one response that can be considered as successfully comprehended:

Excerpt from the text: Or, taking another metaphor, the hearer was envisaged as a detective solving a crime by matching fingerprints to known criminals. All the detective had to do was match a fingerprint found on the scene of the crime against one on his files, and see who it belonged to.

Participant's response: /Uh... they saw the hearer/, the person who hears something/ as a detective who matches the fingerprints to a person it belongs to/. Basically that's how we/, uh... how we... / uh.. basically that's how people... /uhm... when they hear a sound/ that's how they uh... recognize it/ and then later use it/.

Additionally, only 10% of the 155 codes are text-based inferences. This shows that participants failed to understand the text enough to connect ideas from multiple sentences.

Sentence 4: Or, taking another metaphor, the hearer was envisaged as a detective solving a crime by matching fingerprints to known criminals.

Sentence 6: Just as no two people's fingerprints are the same, so each sound was regarded as having a unique acoustic pattern.

Participant's response: /Another example would be that the hearer is the type of a detective/ and that different sounds uhm... can be attributed to like fingerprints/, and no fingerprint is uhm... identical/ so different sounds mean different word patterns/ and all the hearer has to do is match those patterns/ to form... uhm... meaningful sentences./

In this example we can see a successful case of Reinstatement inference, which has rarely been the case during this research. Not many participants used this method in their responses. The number of non-coherent processes cannot be ignored, however, an overall conclusion regarding this text excerpt is that participants were fairly successful.

5 DISCUSSION OF THE FINDINGS IN RELATION TO READING EASE SCORES

The process of collecting think-aloud data was an extensive one, taking approximately eight to ten minutes per participant. During this time, the procedure was explained to each participant, and they were given one minute for each paragraph, after which they moved on to paraphrasing what they read. Out of total eight, three minutes were spent on the participants' reading, meaning that they took approximately five minutes for paraphrasing. The amount of time spent on each paragraph is around five minutes. It is important to note that not every participant spent the same amount of time on the same paragraph. In addition, verbal protocol data from this research also revealed some additional information regarding similarities and differences when it comes to students' personalities as well.

The first participant had many difficulties while paraphrasing either of the text. The total number of non-coherence processes being 48 is evidence enough that the texts were demanding. The text about Waggle Dance was the one they comprehended for the most part, while the second and the third turned out to be more demanding. During the task, one participant commented on the difficulty of the texts, saying what the numbers later confirmed, which is that the second and the third text were challenging.

On the other hand, the second participant focused on paraphrasing the texts without saying out loud anything other than statements related to the text. This way they ensured zero possible distractions. Comparing just these two participants, we can see how different mental processes of two different participants are. Some participants find it easier to express themselves if they speak aloud every thought that occurs to them, while others gathered their thoughts and utter fully formed sentences. Each student is a unique person, hence different mental processes for essentially the same task. It is important to take into consideration these differences and help students express themselves in the best way they know how to. This is what teachers need to pay attention to when they are doing reading exercises in classes, as not every student is able to understand and comprehend texts in a foreign language during the same amount of time as others.

As much as there are differences among students, there are similarities as well. The third and fourth participant from this research both quite successfully paraphrased most of the matter in the texts. Neither of them lost any time on additional comments, mainly focusing on recalling as much of the text as they could. From this example we can see that there are students that work in a similar way, which is also an explanation why group work in the classroom sometimes works very well, and sometimes not. Students who have similar techniques for finishing their task can easily come to an understanding and compromise.

The fifth participant of this research was an example of a persistent student who gives their best to finish the task, no matter the obstacles they face. Regarding the Language Universals paragraph, this participant was the one with the highest number of non-coherent processes which include comments and opinions on the difficulty of the text. The struggle to remember what was written in the text was evident, and the student was quite disappointed for not being able to recall the information. After finishing the task, the participant also stated that it was frustrating to understand the text, but not being able to understand the point. This particular situation has been mentioned at the beginning of the paper, and it was explained that it is possible for an EFL learner to understand the words and vocabulary but failing to comprehend the point behind those words. That is why the way texts in the reading material are written is significant.

The sixth participant left out a major part of the texts during the paraphrasing stage, however, the responses were quite satisfactory as they covered the essence of the texts. From this response we can conclude that while paraphrasing, students do not necessarily have to keep the majority of information from the text in order to comprehend the text as a whole.

The responses from remaining participants were mostly concise, straight to the point, and without major issues during completing the task. Nonetheless, almost all of the participants stated that the Language Universals text excerpt was the most challenging one (which also corresponds to the reading ease score as well), and that even though they did not have difficulties with vocabulary or even paraphrasing, they were still confused and not quite sure what the text was actually about.

The non-coherence processes in participants' responses were both their comments vocalized, as well as invalid statements, only further supporting the argument regarding the difficulty of the texts they read, which can be related to understanding the subject matter actually as the readability scores for the three excerpts show that Excerpt 1 is "fairly easy to read" (74.9), Excerpt 2 is "fairly

difficult to read” (56.7) whereas Excerpt 3 is “standard or average” (60.7) The results correspond to the readability scores and it can be concluded that, regardless of the text being “standard” or “average”, the difficulty of being able to fully comprehend it may be found in the difficulty to understand the comparison given, which may be caused by different factors, one of them being general knowledge or factors related to domain-specific knowledge (e.g. how sounds are “unique”). In any case, as the comparison rendered the text easier to read, it is suggested that comparisons and metaphorical comparisons be used when teaching about content via language, which the author of the book tried to do by using numerous comparisons throughout the textbook.

6 CONCLUSION AND SUGGESTIONS FOR FUTURE RESEARCH

The aim of this master's thesis was to explore and determine factors that affect reading comprehension. The strategy for doing so was making and analyzing verbal protocols. The verbal protocols were collected owing to the participation in the study of 10 students of the master's degree studies at the Department of English Language of Literature, Faculty of Philosophy, University of Sarajevo who attended an elective course in Psycholinguistics. The participants were given a timeframe in which they had a task to read a short paragraph, and afterwards paraphrase it. The task was to read three paragraphs for three different types of texts in terms of their reading scores calculated online.

One paragraph was a text written using high-frequency language and phrases familiar to participants. The second paragraph contained metalanguage, an excerpt written in a more complex way, while the third one was using a metaphor to convey its message. The participants' responses were recorded via a web page (Online-voice-recorder.com). The recordings were transcribed by the author of this thesis and coded using the coding scheme by Bohn-Gettler and Kendeou (2014). The results of the study showed that the text written using high-frequency language and simple vocabulary was the one with the highest percentage of successful paraphrases. The text that proved to be the most challenging one was the metalanguage excerpt which created confusion among students, and its complex style was difficult to comprehend. The third text had somewhat more encouraging results, leading us to a conclusion that, after high-frequency language and simple instances in definitions, using metaphors and comparison is the next most plausible approach.

As can be concluded from the results above, reading comprehension is a difficult, yet important skill, the lack of which can have a quite negative impact on the reader. After the analysis of the results, we can see that the definition containing abstract phrases and a language that cannot be heard in everyday life was more difficult for students to comprehend in the given timeframe. If we were to update school textbooks so that definitions resembled more the first excerpt from this research rather than the second, in this way readers/learners would be able to comprehend the matter faster, which would furthermore improve their skills. It is not suggested here that the texts

should be simplified but rather adapted or even boosted by additional metaphorical comparisons or even illustrations to accompany them. As we have seen, verbal protocol data is a thorough procedure that reveals one's mental processes and, therefore, can be used for exploring the way students' minds' work while performing different reading comprehension tasks, with the final aim being improvement of classes to fit students' needs better. The research from this paper was done with university students who were familiar with the texts they were given to read in their task.

After the analysis I came to a conclusion that one type of text was more difficult for students to comprehend, while the other two were less challenging. Since all the three of the texts were taken from a textbook students used previously in their course, I cannot help but wonder if the overall comprehension of the subject matter would have been more successful if the textbook had been written in the similar manner as the two texts that proved to be less difficult for them? What if they would have been able to acquire more knowledge if they had been presented with definitions that were easy to learn, rather than confusing examples that take up more time and effort to remember? Speaking from personal experience, I have encountered a considerable number of textbooks that were not only difficult process, but also made me lose interest in the subject as the matter was too time-consuming. Regardless of the fact that the professors were doing their best to pass their knowledge onto us, a lot of us simply were not engaged in learning more than it was necessary and exploring beyond the planned curriculum, as the material we had to study was demanding.

Taking into account the presented evidence and the hypotheses confirmed, it is safe to assume that verbal protocol data shows that this method of research may provide an insight into the mental processes important for reading comprehension, which is a significant skill that can either ease the learning experience for students or complexify it. If similar research were to be done with younger students, using texts appropriate for their age and level of knowledge, the results could show which types of texts students have trouble comprehending. Later on, teachers could adjust their teaching material so it corresponds to students' needs. In this way, teachers would be able to present students content that they would be able to comprehend easily, which would eventually motivate them to be more engaged in their studies and study materials.

REFERENCES

- Aitchison, J. (2008). *The Articulate Mammal*. New York: Routledge.
- Bohn-Gettler, C. M., & Kendeou, P. (2014). The interplay of reader goals, working memory, and text structure during reading. *Contemporary Educational Psychology*, 39(3), pp. 206-219.
- Brian Scott. (n.d.). Readability Formulas. Retrieved from <https://readabilityformulas.com/free-readability-formula-tests.php>.
- Cohen, A. D. (1987). Using verbal reports in research on language learning. In C. Faerch & G. Kasper (Eds.), *Introspection in second language research*. Clevedon, England: Multilingual Matters, pp. 82-95.
- Cohen, A. D. (2000). Exploring strategies in test taking: Fine-tuning verbal reports from respondents. In G. Ekbatani & H. Pierson (Eds.), *Learner-directed assessment in ESL* Mahwah, NJ: Erlbaum, pp. 131–45.
- Cohen, A. D. Verbal Report (2013). In C.A. Chapelle (General Ed). *The Encyclopedia of Applied Linguistics*. Oxford: Wiley-Blackwell.
- Cohen, A. D., & Hosenfeld, C. (1981). Some uses of mentalistic data in second-language research. *Language Learning*, 31(2), pp. 285–313.
- Duke, N., & Pearson, P. D. (2002). Effective practices for developing reading comprehension. In A. Farstrup & S. J. Samuels (Eds.), *What research has to say about reading instruction*, pp. 205–242. Newark, DE: International Reading Association.
- Ericsson, K. & Simon, H. (1993). *Protocol analysis*. Cambridge, MA: MIT Press.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (rev. ed.). Cambridge, MA: MIT Press.
- Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Green, A. (1998). *Verbal Protocol Analysis in Language Testing Research: A Handbook*. Studies in Language Testing, 5. Cambridge: Cambridge University Press.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and writing: An interdisciplinary journal*, 2, pp. 127–160.

Jakobson, R. (1956). *Metalanguage as a Linguistic Problem. Selected Writings, VII*. Edited by S. Rudy. Mouton, 1985, pp. 113-121.

Leow, R. P., & Morgan-Short, K. (2004). To think aloud or not to think aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26(1), pp. 35–57.

Measure the Readability of Text - Text Analysis Tools - Unique Readability Tools to Improve Your Writing! App.readable.com. (2021). *Readable*, <https://app.readable.com/text/?demo>.

Nordquist, R. (2020, August 26). *Metalanguage in Linguistics*. Retrieved from <https://www.thoughtco.com/what-is-metalanguage-1691382>

Pressley, M. (2000). *What should comprehension instruction be the instruction of?* In M. Kamil, P. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research*, Vol. III. Mahwah, NJ: Lawrence Erlbaum, pp. 545–561.

Pressley, M., Afflerbach, P. (1995). *Verbal Protocols of Reading: The Nature of Constructively Responsive Reading*. Erlbaum: Hillsdale, NJ.

Radford, J. (1974). Reflections on introspection. *American Psychologist*, 29(4), pp. 245–50.

Robinson, P., ed.(2001). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Snow, C. E., & Sweet, A. P. (2003). *Reading for Comprehension*. In A. P. Sweet, & C. E. Snow (Eds.), *Rethinking Reading Comprehension*. New York: The Guilford Press.

APPENDIX 1

First Participant's Response	Coding Statistics	
<p>/Uhm, okay... the... /I read the uhm... small excerpt regarding the uhm.. bee dance /uhm... and how they, uhm... they usually start to uhm... have this uhm... wiggle dance/, I think uhm... /, which also, uhm with the way they notify the other bees / to come and collect the nectar uhm... /They have this kind of uhm... circling around, /uhm..., maybe spinning around the flower, / or uhm... and so uhm... / pretty much the... the excerpt is uhm... how bee do the dance in collecting the nectar./</p>	Total number of codes in the response	9
	Paraphrases	5
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	4
<p>/Uhm... okay.../ a little... this excerpt is a little uhm... demanding.../ uhm.. It's regarding the uhm.. Chomsky's hypothesis/, uhm... and innate uhm... knowledge./ Uhm... so uhm... it uhm.../ it is based on the research regarding... uhm.../ not sure if it's a research from him,/ but so... its about uhm... two way thinking, /uhm... if... Eski... / so... Eskimos can.../ If their innate knowledge is uhm... two sides./ So one would be that they uhm... build their igloos from uhm... ice and snow/ and another would be regarding that uhm.../ its the round shape uhm.../ so uhm... not sure./ So, I mean interesting fact, /uhm... Something new for me to learn./ But definitely, uhm... we would have this uhm... two way of thinking... /so we need to... uhm.. to split... uhm... from... from... what the igloo is made/, and what... what... kind of shape it is/, and what is.../ uhm... what is.../, what is its uhm... main.../ what is its main uhm... concept/, I mean uhm.. it's pretty much house for Eskimos.../</p>	Total number of codes in the response	25
	Paraphrases	5
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	20
<p>/Okay... uhm.../ Okay, I managed to read it uhm... twice,/ but uhm... I still couldn't... /uhm.. remember as... as much as I could./ So, it's.. /it's... regarding the research that scientists done in twenties. /Uhm... and it connected uhm... somehow with uhm... investigating/ uhm...</p>	Total number of codes in the response	29
	Paraphrases	5

<p>Actually, they tried to, uhm.../ what they hear uhm / they tried to put it on the paper, /uhm... any sound I suppose, / uh, even saying /a/, or... or... any articulated sound/, they tried to put it on the paper /uhm.. and then read it. / Uhm..., I don't... /I'm not sure if that is,/ but I assume it is,/ because they investigated language, /I think. Uhm.../, so its also uhm... mentioned something regarding this detective search uhm.../ and uhm... how they uhm... managed to uhm.../ the tried to... /with the fingerprints uhm... /to uhm... find uhm... I think uhm..., / find something uhm... crime scene.../ not sure.../ I didn't catch it /and figure it out what it's for.../ This one is the hardest.../</p>	Text-based inferences	4
	Knowledge-based inferences	0
	Non-coherence processes	24

Second Participant's Response	Coding Statistics	
<p>/The text is about a... uhm... is about uhm... working bees./ So, when a working bee uhm... finds uhm... nectar, /a place uhm... where they can collect nectar/, this bee comes to hive/ and uhm it dances/ and if, uhm... I believe if the nectar is close to the hive then she.../ it does this round dance/ and if uhm... if the nectar is far from the hive it does this waggle dance./ Previously it had been believed that uhm.. bees are deaf/, but they are not/ since they can hear,/ uhm, since they can hear moves in the hive/ even if the hive is dark./</p>	Total number of codes in the response	12
	Paraphrases	7
	Text-based inferences	5
	Knowledge-based inferences	0
	Non-coherent processes	0
<p>/So, according to Chomsky, there are two types of language universals, substantive and informative./ Substantive universals are the ones that present building blocks of language/ and uhm.. formative—formative universals are concerned with form and shape of grammar/ and there... there was this analogy presented/, this distinction between substantive and formative universals uhh was uhm..was compared to the knowledge that is innate to Eskimos/. For example Eskimos.../ if Eskimos were born with two types of knowledge/ that would be...</p>	Total number of codes in the response	14
	Paraphrases	5
	Text-based inferences	8
	Knowledge-based inferences	0

/substantive knowledge would be that they know that igloos are made of uhm... /made of snow and ice,/ and formative knowledge would be that they know that igloo is round in shape/ and not like sausage like/, or whatever.../	Non-coherent processes	1
/It is... /At the beginning of the twentieth century psycholinguists assumed that the underst.../ that the process of understanding speech was a simple one/ and there were two comparisons/. First one said that they compared uhm... the hearer to a uhm... secretary/ who would just be taking down notes/ while a person is speaking/ and then uhm... /after that person is done speaking uhm... / the hearer would just read off the.../ uhm... the words that she or he took down/ and then uhm... understand them/. And the second comparison was that hearer.../ hearer was envisaged as a uhm... detective/ who would uh... match uhm... /fingerprints found on a crime scene to .../ uhm... to all the persons in his file/ and then find out who... /whose fingerprints match the person in his file/ and in that way they compared the words of a hearer to a hearer./	Total number of codes in the response	20
	Paraphrases	15
	Text-based inferences	3
	Knowledge-based inferences	0
	Non-coherence processes	2

Third Participant's Response	Coding Statistics	
/Basically when a worker bee finds a source of nectar it returns to the hive/ and it does a dance to inform the other bees of where it is/. It can do a round dance if its close by/ or a waggle dance if it's farther away./ Uhm... depending on how, like the dance goes/, the bees can tell exactly how far away it is/. And, they don't necessarily need to see it/, because bees can hear the beat of the worker bees wings as its dancing/, and that can also help them figure out/ even if they cannot see the dance/... so that's it./	Total number of codes in the response	11
	Paraphrases	6
	Text-based inferences	4
	Knowledge-based inferences	0
	Non-coherence processes	1
/Uhm... so.../ Chomsky said that language universals are basically divided into two categories/. Substantive and formal/. Substantive are those that are... are considered the building blocks of language/ or what it's made of/, and formal are the words which tell us like how to	Total number of codes in the response	19
	Paraphrases	17

<p>make formations out of those building blocks/ or at least how grammar is formed/. An analogy can be made/, or a sort of a metaphor/ if we say that Eskimos have an innate knowledge of how to build igloos/, then you would say that they have two types of knowledge/. The first one is the knowledge that igloos are made of ice/, and the second is the knowledge that like/, that there are a round shape/, that how to shape the ice/. In the same same way that /uh.. it was mentioned at the end /uhm... thrushes birds/, they do round nests/ and not nests in the shape of a bath./</p>	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherent processes	2
<p>/At the beginning of the twentieth century they thought that the process of hearing and understanding speech was a simple one/. Basically, the hearer would, / uhm.. i's like a person taking dictation/, like you would simply record the words mentally/ and then read them off one by one/. Uhm... Alternatively, it can also be described as like a... detective/, where they just found fingerprints at the scene of the crime/ and then match the fingerprints with files they already have/. In other words, each sound would have a unique corresponding word/, as far as I understand./</p>	Total number of codes in the response	10
	Paraphrases	6
	Text-based inferences	3
	Knowledge-based inferences	0
	Non-coherent processes	1

Fourth Participant's Response	Coding Statistics	
<p>/Uh.. basically as once it was thought that bees could not hear/, that they are deaf/, but uhm... in fact that is not true /because bees can communicate /uh.. with their bodies/. Uhm... basically if they find nectar/, if one bee finds the nectar/, she comes to other bees/ and uhm... she does a dance/ which indicates where and how far the nectar is/ and other bees understand her /uhm... in means that they can hear her wings flapping/. Basically, if the nectar is close/ uhm..., she would do the round dance/, and if the nectar is not that close/ she would do the waggle dance/. And the bees would recognize where the nectar is/, which flower the nectar is from by smelling uhm... the bee/. And basically, bees can communicate uhm... this way./</p>	Total number of codes in the response	19
	Paraphrases	18
	Text-based inferences	1
	Knowledge-based inferences	0
	Non-coherent processes	0

/Uhm... To start with... /Basically, there's... uhm.../ Chomsky talks about how there is some substantial and fundamental knowledges/, and the first thing is that some knowledge is innate/, which means for example if Eskimos were to be born with innate knowledge of knowing how to build uhm... igloo/, they would know uhm... /that they have to do it/, but on the other hand they would also be born with knowledge that the igloos are built round and with ice/... That's it./	Total number of codes in the response	9
	Paraphrases	6
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherent processes	3
/Uhm... Psycholinguists /uh...at the beginning of the twentieth century saw speech/ understand,.../ us people understanding speech/ as something not so complex but yet simple/. Uhm... and uhm.... to describe ... /to describe how they saw it... /uhm... they saw the hearer/, the person who hears something/ as a detective who matches the fingerprints to a person it belongs to/. Basically thats how we/, uhm... how we .../ uhm... basically that's how people... /uhm... when they hear a sound/ that's how they uh... recongize it/ and then later use it/. Just as detectives can pinpoint that one fingerprint to the person it belongs to/. I would say that./	Total number of codes in the response	19
	Paraphrases	13
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherent processes	6

Fifth Participant's Response	Coding Statistics	
/Okay, so uhm... /when bees find a source of nectar/, they go back to their hive/ and perform a complex dance/. They either do a round dance/ or a waggle dance/. Uhm... in case of round dance/, they do that kind of dance /uhm... if the nectar is near/, and in waggle dance/ in which they shook their tails side to side/ so if the source is far away/. The bees in the hive can judge distance by the tempo/ and the type of flower by the scent of the bee/. Uhm,.. previously bees have been perceived as deaf/, but they aren't/. They can produce like a sound/, the way they beat their wings/... I think that's it./	Total number of codes in the response	18
	Paraphrases	14
	Text-based inferences	2
	Knowledge-based inferences	0
	Non-coherent processes	2
/Okay, this one is a bit more difficult.../ Chomsky suggests that two... /Ooh... okay, oh my God.../ uhm... I forgot... /So, we have... /oh Oh God,.../	Total number of codes in the response	31

<p>okay okay... I'll just say how I remember/. There's substantive/ and forma.../ formative grammar.../ In which uhm.../ this substantial consists of building blocks of language/ so the thing that its made of/, and the formative is/,... oh it's formative... /The formative is about the form of grammar/... so analogy used are Eskimos/ when they build igloos/. So they know the substance of igloos they are building/, which is ice/, but they don't know/... Wait,wait, no, no. /They know it's built from ice/, but they don't have the innate knowledge from when it comes to building its shape/, the same way thrushers know that their /... oh God, this one is difficult... /aa... so basically igloos know that... /Eskimos know igloos are made out of ice/, but they don't know/, they don't have an innate knowledge about shape/. Okay, this one was difficult./</p>	Paraphrases	0
	Text-based inferences	5
	Knowledge-based inferences	0
	Non-coherent processes	26
<p>/Okay, so in the twentieth century psycholinguists assumed that uhm... aa... language.../ a... uhm.. that.../ oh God.../ that language.../ uhm... actually.../ uhm... so basically was simple/. It was envisioned as a secretary taking down a dictation/ and then after she read the words one by one/. Uhm... the same /... it was also as a detective/ who finds a fingerprint on a crime scene/ and then matches the fingerprint to the cases he has/. Uhm... oh God... /so, each sound... /each sound is. uh basically each word has a unique acoustic sound/... That's it.. /This was... /My brain is kind of lagging.../</p>	Total number of codes in the response	17
	Paraphrases	6
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	11

Sixth Participant's Response	Coding Statistics	
<p>/So, basically the text is about bees/. And the text says that when bees find the nectar suiting for them/, they um ... /they... uhm... /The bee comes to their friends and coworkers/ and informs them by performing a dance/. This was dancing in circles if the nectar is near... near them/. This can happen even in the dark/. So basically bees have a let's say navigation system that allows them to collect the nectar./</p>	Total number of codes in the response	9
	Paraphrases	4
	Text-based inferences	3
	Knowledge-based inferences	0
	Non-coherence processes	2

<p>/So the second text is about meta language/ and uhm... according to Chomsky there are two basic types of language universals/, formal and substantive/. Formal... uhm... universal is concerned with form and shape of the grammar/, while substantive is uhm... about fundamental building blocks of the language/. And uhm... the example given was of the Eskimos/ and how they build igloos. /According to language universals uhm... /Igloos.../ if we apply the uhm... substantive form/, Eskimos were born with the innate knowledge on how to build an igloo/, and which material they should use/ and how the shape would be/. So, basically that's it. /Well, you know on the other hand/, formal universals would be that they have... /that they had to have been taught how to build an igloo/ and what material to use/ and how the igloo would be./</p>	Total number of codes in the response	17
	Paraphrases	12
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	5
<p>/The last text is about a comparison/ and it's about speech/. And the appliance of psycholinguistics/, and how we understand speech/ and it basically uhm.../ even... every word every sound, we tend to wrote.../ write mentally down/ and it's compared to the job of detective/, to fingerprints/. Meaning that every sound is like every fingerprint/, unique, different from other fingerprints or sounds/. So basically every sound we use is different from the others./</p>	Total number of codes in the response	12
	Paraphrases	11
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	1

Seventh Participant's Response	Coding Statistics	
<p>/Alright, so this text is about bees/. So basically, when bees find nectar they are going to do their wiggle dance/... uhm... to inform other bees that nectar is near/... so, if they are wiggling rapidly/, that means that nectar is uh.. very close/ so the other bees are going to mimic this so called dance/ they're going to inform other bees in the hive where the nectar is./</p>	Total number of codes in the response	7
	Paraphrases	7
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	0
<p>/Alright, so this one is about Chomsky uhm.../ and him saying that there is a substantial and formative type of something/... uhm... so he gave us an example of Eskimos/, of their let's say innate uhm.. language abilities/ but uhm... he</p>	Total number of codes in the response	9
	Paraphrases	6
	Text-based inferences	0

brought it in a way that if Eskimos had innate knowledge of building igloos/ uhm... they know that they're built from ice/, but on the other hand uhm... they do not/, I think that they do not know which shape it needs to be/, does it need to be square, rounded, or any other shape./	Knowledge-based inferences	0
	Non-coherence processes	3
/Alright, so uhm... the text is about psychologists who thought that uhm... either letter or sounds were simple/, so they envi... envi... envisaged/, they saw them as people who could/... as uh.. I'm sorry... Uhm.../ Can we do this again, please/... So, they saw people for example as detectives who would find clues uhm... like... fingerprints/ uh as every fingerprint is not the same/, so each sound is not the same/, it has uhm.../ they are different/. Uhm... Just the first one, I can't remember the word/. It's dictation/... uhm.. The person who writes down the words/... uhm so they write each and every word/, and when they write them they memorize them/ so they are able to pronounce them again./	Total number of codes in the response	14
	Paraphrases	0
	Text-based inferences	3
	Knowledge-based inferences	0
	Non-coherence processes	11

Eighth Participant's Response	Coding Statistics	
/Okay so it's a text about how bees communicate/ and that they aren't really deaf as it was once thought/. Uhm, once a bee finds a nectar source it returns to the hive/ and it does either a round dance if a flower or nectar is nearby/, or a waggle dance if the nectar or flower is further away/. That's what I gathered./	Total number of codes in the response	6
	Paraphrases	5
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	1
/Uhm... it's a text about language universals presented by Chomsky/. Two types, formal and I'm not sure what the other one's called/. Uhm.. Formal one deals with... I have no idea/. I remember the uhm... Eskimos/. It deals with innate knowledge of language/ of whether or not humans are born with some innate language knowledge/, I know that/. Chomsky gives an example of Eskimos/ and how they build igloos/. If they had innate knowledge, they would know that igloos are made out of ice/ and in particular shape/, so I think... circular/. That's what I remember./	Total number of codes in the response	14
	Paraphrases	8
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	6
/This text deals with the understandment.. understanding of speech/ and how	Total number of codes in the response	10

<p>psycholinguists perceive it/. Uhm, in one example they say it's like, that a hearer is like a some type of a secretary/ that basically mentally writes down whatever a speaker says/ and then uhm... pairs certain words to other to form meaningful uhm... sentence/. Another example would be that the hearer is the type of a detective/ and that different sounds uhm... can be attributed to like fingerprints/, and no fingerprint is uhm... identical/ so different sounds mean different word patterns/ and all the hearer has to do is match those patterns/ to form uhm... meaningful sentences./</p>	Paraphrases	4
	Text-based inferences	6
	Knowledge-based inferences	0
	Non-coherence processes	0

Ninth Participant's Response	Coding Statistics	
<p>/So, uhm... it has been proved that bees are actually not deaf/ and that uhm... when a bee which is responsible for looking for honey/, when she finds honey she informs her.../ she returns back to the hive/ and informs her bee friends/, uhm... that she has found a bee/ and she has two ways of doing it/. One is by round dance/ which signifies that the bee is close to them/ and the other one is the wiggle dance/ or something like that/, uhm... which signifies that the bee is far away/. And the other bees... uhm... the... recognize the distance/ by analyzing the tempo of her wiggles/ and uh.. they find out which flower to look for/...by... by looking at the scent of her body/. Basically that's it./</p>	Total number of codes in the response	17
	Paraphrases	14
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	3
<p>/Uh, well according to Chomsky, language universals are of two types/, substantive and uh... one more/, I forgot/. Um, might be fundamental/, or something like that/. Anyways, the first one is concerned with the uhm... building blocks of language/, while the other one is uhm... more like uhm.. uhm... /has to do more with grammar/ and he made an analogy with Eskimos/ and their igloo building/ and if the knowledge of igloo building for them was innate/, they would know that it was one of two kinds of knowledge/, one would be similar to substantive/, which would suggest them that their igloos are made of snow and ice/ and the other one would be fundamental/... and I forgot./</p>	Total number of codes in the response	16
	Paraphrases	10
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	6

/So, at the beginning of the twentieth century/ uhm... psycholinguists assumed that the process of uhm... recording... processing speech is simple one/ and they compared hearer to a uhm... secretary/ at who someone dictates the words/ and she simply uh.../ puts letter uhm.../ puts sounds one by one/ and forms a word/ or they would compare a hearer to a uhm..., detective who is decoding fingerprints/. Just as each fingerprint is different, so are the sounds/ and the only thing a hearer had to do was to match the fingerprint found on the place of the crime to the fingerprints available/ and it was like a game of matching./	Total number of codes in the response	12
	Paraphrases	10
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	2

Tenth Participant's Response	Coding Statistics	
/Well, this is interesting text/. It's about bees and them wagging their tails in order for other bees to find the nectar and flowers/. Uhm... as once was assumed, bees are not deaf/ so they perform a dance in order to tell other bees where it's the flower/ and where is the nectar they need to find/. And also uhm... there was one interesting thing I wanted to remember but now I forgot it/. Ah yeah, by the way they're wagging their tail, they show other bees how far is the flower/ and how far is the nectar/. I think that was the most important part of the text./	Total number of codes in the response	9
	Paraphrases	6
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	3
/This one was a little bit more difficult/. But, it was about language universals/. Which was proposed by Chomsky/, He said that there are two types: formal and substantive/, or substantial, I'm not sure/. One of them represents the building blocks/ and the other represents, in a way, form and shape how the language is formed and the sentences/. He used Eskimos/ and them building uhm... an igloo/. Uhm... it's like they have two kinds of knowledge/. Uhm... one of the.../ one type of knowledge helps them to build uhm... the igloos/ which is in a round shape/, and the other/... Honestly, I don't remember/. Yeah, and uhm they, yeah.. /And they have some kind of innate knowledge/ which helps them to shape	Total number of codes in the response	18
	Paraphrases	12
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	6

the igloo in that shape in which it is/. I think that it was something like that./		
/Well, this one was interesting also/. Uhm... it was about psycholinguists assuming that language and understanding of a language was a simple one/. They used metaphors/. The first one was about hearer being a secretary writing down all the sounds they hear/ and then in a way uhm... decoding them. / And then the other metaphor was a detective solving a crime/ uhm... when they match fingerprints they find on a crime scene to a person to whom it belongs. /And, the last sentence was an interesting one /where they say that there are no two.../ there are no two the same fingerprints/ as there are no two same acoustic sounds./ Something like that./	Total number of codes in the response	12
	Paraphrases	10
	Text-based inferences	0
	Knowledge-based inferences	0
	Non-coherence processes	2

APPENDIX 2

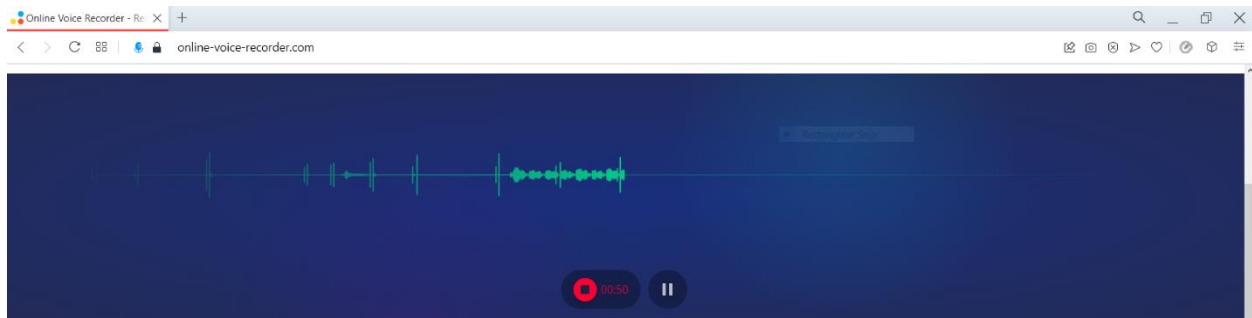


Fig. 5. Online voice recorder, screenshot