University of Sarajevo

Faculty of Philosophy

English Language Department

EVALUATION AND ASSESSMENT OF LANGUAGE COMPETENCIES IN ELEMENTARY SCHOOLS: BASIC PRINCIPLES AND THEIR APPLICATION

(Vrednovanje i praćenje razvoja jezičkih kompetencija u osnovnim školama: osnovni principi i njihova primjena)

Diploma thesis

Student: Irma Vehab   Mentor: Prof.dr. Larisa Kasumagić-Kafedžić

Sarajevo, 2019

CONTENTS

ABSTRACT

The purpose of this research is to determine which types and methods of evaluation and assessment are used by teachers in Bosnian schools, how effective these methods are, and to what extent the current education system in Bosnia and Herzegovina allows and supports the use of traditional and alternative ways of evaluating and assessing the development of language competences. The initial step of the research involves compiling and analyzing relevant literature with the aim of defining language assessment, distinguishing between different types of assessment and underlining the importance of using a variety of assessment methods and techniques, particularly in the context of Bosnian education system. The final step of the research involves an anonymous survey conducted among English language teachers at the level of Sarajevo primary schools. Results of the survey indicate that Bosnian education system does not fully support the use of alternatives in assessment. According to the survey, Sarajevo teachers use traditional means of assessment for the majority of the school year and rarely employ alternatives in assessment due to lack of time, finances and motivation. This research will offer valuable information on the importance of assessment in education, as well as highlight the need for future reforms in education that would offer more freedom for the teachers.

KEY WORDS: *assessment, evaluation, education, alternatives, young learner*

SAŽETAK

Svrha ovog istraživanja je utvrditi koje vrste i metode vrednovanja i praćenja razvoja jezičkih kompetencija koriste nastavnici u bosanskohercegovačkim školama, koliko su one učinkovite i u kojoj mjeri trenutačni obrazovni sistem u Bosni i Hercegovini dopušta i podržava upotrebu tradicionalnih i alternativnih načina vrednovanja i praćenja razvoja jezičnih kompetencija. Početna faza istraživanja obuhvata prikupljanje i analizu relevantne literature s ciljem definiranja pojmova vrednovanja, praćenja i ocjenjivanja, razlikovanja vrsta ocjenjivanja i naglašavanja važnosti korištenja različitih metoda i tehnika vrednovanja i praćenja, naročito u kontekstu bosanskohercegovačkog obrazovnog sistema. Završna faza istraživanja uključuje anonimno istraživanje provedeno među nastavnicima engleskog jezika na razini osnovnih škola u Sarajevu. Rezultati ankete pokazuju da bosanskohercegovački obrazovni sistem u potpunosti ne podržava upotrebu alternativa u vrednovanju, praćenju i ocjenjivanju.. Prema rezultatima istraživanja, nastavnici u Sarajevu koriste tradicionalne metode praćenja i ocjenjivanja tokom većinskog dijela školske godine i rijetko koriste alternativnije metode praćenja i ocjenjivanja zbog nedostatak vremena, finansija i motivacije. Ovo istraživanje će ponuditi korisne informacije o važnosti vrednovanja i praćenja razvoja kompetencija u obrazovanju, kao i naglasiti potrebu za budućim reformama u obrazovanju koje će pružiti više kreativne kao i finansijske slobode za nastavnike.

KLJUČNE RIJEČI: *vrednovanje, praćenje, obrazovanje, alternative, mladi učenik*

1. INTRODUCTION

In this day and age we are very well aware of the different ways in which children acquire languages. We see children as individuals with their own strengths and weaknesses, likes and dislikes, affinities, capabilities, styles of learning as well as their cognitive, social, and emotional skills. Gardner's multiple intelligence theory (1983, according to McKay, 2006) suggests that human intelligence can be divided into multiple specific abilities rather than it being a single general ability. There are eight types of intelligence – linguistic, musical, logical-mathematical, spatial, bodily-kinesthetic, interpersonal, intrapersonal, and naturalistic. In addition to this, we can further argue that there are multiple ways in which children learn and we cannot restrict them to a single learning style.

Unfortunately, it seems as if teachers nowadays spend too much time on introducing a variety of different teaching styles and methods, but not merely enough time on introducing new methods of assessment. According to various reports (Dragunić, 2013; Ibrahimović, 2015; Centar za Politike i Upravljanje, 2003) teachers in Bosnian primary schools seem to be heavily relying on testing. Douglas H. Brown (2007) defines testing as a proven method of measuring someone's knowledge in a given domain. However, testing is only one of the many methods that can be used for assessing someone's knowledge. Assessment is a continuing process that cannot be defined by a simple test, especially with young learners. A good teacher will keep track of every time a learner has uttered a new structure, answered a question, made a comment, and assess their knowledge subconsciously.

McKay (2006) argues that in order to assess language competence of young learners a teacher has to be familiar with the overall characteristics of young learners as well as the overall characteristics of the language learning process. Without proper understanding of the said matter a teacher risks unfair and invalid assessment.

Having this in mind, three major starting points of the literature overview will be the definition of the term 'young learner', the definition of the term 'assessment' and related terms (as well as detailed analysis of assessment in general, assessment of the young language learner, and types of assessment), and lastly, the overview of the educational system in Bosnia and Herzegovina in relation to assessment practices in teaching langauges. The main aim of

the theoretical part is to offer information from relevant sources that would support or reject the idea that traditional methods of evaluation are not enough in providing effective evaluation of language competence.

When it comes to the research part of the paper, the main aim is to determine which types and methods of evaluation and assessment are used by teachers in Bosnian primary schools, how effective these methods are, and to what extent the current education system in Bosnia and Herzegovina allows and supports the use of standard and alternative ways of evaluating and assessing the development of language competences. Essentially, this part will offer a detailed look into the methods, techniques and instruments used for the realization of the research, as well as a look into the end results of the research.

This entire thesis is intended to serve as a small contribution to the discussion about the assessment practices in Bosnian primary schools. It is intended to shed the light on the existing issues in the very education system, especially in relation to the assessment practices. Lastly, it should serve as a starting point of a bigger research and a bigger discussion about the importance of assessment, and the importance of learning and implementing authentic assessment practices in the classroom.

2. LITERATURE REVIEW

Three major starting points of this part of the paper are the definition of the term 'young learner', the definition of the term 'assessment' and related terms (as well as detailed analysis of assessment in general, assessment of the young language learner, and types of assessment), and lastly, the overview of the educational system and assessment practices in Bosnia and Herzegovina in comparison with educational systems and assessment practices in other European countries.

2.1. YOUNG LANGUAGE LEARNERS

It is important to clarify the very definition of the term 'young learner' or even 'young language learner'. There is an obvious difference between these two terms: a 'young learner' is simply a child that is going through the process of formal schooling, while a 'young language learner' is a child learning a foreign or a second language while going through the said process (McKay, 2006). There are a vast number of sources offering different interpretations of these two terms. This is understandable since educational systems differ from country to country, which essentially means that children in different countries start their educational journey at different ages. On top of that, children can also start learning new languages at different ages, either that be at kindergarten, primary school or secondary school levels.

According Gail Ellis (2013), a British Council adviser for young learners, a child is defined as anyone who has not reached their 18th birthday (as established by the United Nations Convention on the Rights of the Child in 1989). Furthermore, Ellis offers her understanding of the correct young learner terminology. A preschooler, age two to five, should be addressed as 'early years'. A primary school pupil aged six to ten should be addressed as 'primary'. A secondary school pupil aged eleven to fourteen should be addressed as 'lower secondary', while a secondary school pupil aged fifteen to seventeen should be addressed as 'upper secondary'. Lastly, a university student aged eighteen to twenty five should be addressed as young adult. This definition, alongside with all of the terminology, can be seen as problematic because it cannot be applied to every context (Ellis, 2013, pg. 77).

Pinter (2006) tries to include most contexts in which children might start learning a foreign language and defines the age group of a young language learner from five to fourteen years of age.

On the other hand, McKay (2006) goes into more detail explaining the reasoning behind choosing the age group from five to twelve as a representative of the young language learner. 'Young language learners are those who are learning a foreign or a second language and who are doing so during the first six or seven years of formal schooling' (McKay, 2006, pg. 1). She goes on to explain that in most countries, young learners are those attending primary schools or elementary schools.

It would not be appropriate to claim one of these definitions is more accurate than the other, especially since all of them could be applied successfully in different contexts. In order to stay true to Bosnian educational system, young language learners will be defined as those who are learning a foreign language (English language, in particular) while attending primary schools. There are kindergartens which offer English language classes to children aged three to six, however, this paper will focus on primary school children only.

2.1.1. Characteristics of young language learners

Teaching children is a completely different experience to teaching older learners. Most teachers would agree that teaching children is more challenging (contrary to the popular belief) but also more rewarding than teaching any other age group. There is a variety of characteristics that are typical of young language learners in comparison to older language learners.

The majority of consulted sources (McKay, 2016; Brown, 2004; Brown, 2007; Pinter, 2006) agree that the most prominent characteristics of young language learners are intellectual development, limited language skills (even in their first language), limited attention span, vulnerability, need for meaningful language, as well as need for all their senses to be stimulated.

When it comes to intellectual development we can agree that primary school children are going through major changes that influence their abilities and performance. McKay argues

that 'children are in a state of constant cognitive, social, emotional and physical growth' (McKay, 2006, pg. 6). What this means is that their perception of the world is completely different to that of an adult learner. They are still learning to organize information, to manipulate thoughts and ideas, and even to use their own language. Understanding different stages of their development is crucial for effective assessment. Brown D. (2007) also argues that children are usually at the stage where they cannot grasp abstract aspects of language or linguistic concepts which implies that these matters require mindful approach in teaching and assessing.

Children also have limited literacy skills in both their first and their second/foreign language. Pinter (2006) suggests that older learners already have developed skills of reading, writing and speaking in their first language which makes it easier for them to focus on learning a completely different language. Younger learners, on the other hand, are usually learning both at the same time which is a fertile ground for language transfer. McKay (2006) agrees with this notion and suggests that the literacy development in both languages does not have to occur in parallel; in fact, young learners might be learning a huge chunk of their literacy in their second language. She also explains why teachers need to be aware of this matter in order to effectively assess language competence. Teachers need to be able to select appropriate tasks for younger learners and to adjust their expectations of the performance on said tasks.

Brown D. (2007) argues that it would be incorrect to assume that children have limited attention span. If you give them their favorite cartoon to watch they will sit in front of the television for hours. However, if you give them something they consider boring, difficult or pointless they will lose their interest in the matter of minutes. Unfortunately, language learning can be difficult and frustrating at times which only means that teachers have to be able to make it as fun and lively as possible.

Another very important aspect to keep in mind is that older learners are not the only ones who are vulnerable to criticism and failure. Children are usually at the stage where they are confident only if they receive reassurance and support from their surroundings. McKay (2006) suggests that children are sensitive to praise, criticism and approval in the classroom as these experiences at school majorly influence their perception of the world. Children have to feel

secure, loved and nurtured in order to succeed. This especially applies to language learning because young learners might become overly insecure and refuse to participate in class. This is something that Brown D. (2007) also mentions and suggests that children are much more fragile than adults in that respect. If they are afraid of being mocked or criticized by their peers or by the teacher they will refuse participation and the overall quality of their performance will decrease.

Meaningful language is something that both Brown D. (2007) and Pinter (2006) discussed as an important trait of a young language learner. Pinter (2006) argues that children in general have a holistic approach to language. Since they are still unable to grasp difficult linguistic concepts and discuss language as such, they can only understand meaningful messages. What is meant by meaningful language is language that is contextual, connected to real life situations, backgrounds and stories. Brown D. (2007) even goes as far as to state that children can sense language that is not authentic. They will only be interested in language they can use then and there, and therefore will reject anything that does not have a purpose in that moment.

Lastly, young learners need all their senses to be stimulated in class. Teachers usually rely on visual and auditory stimulation, however, there are many other ways in which you can cater to every learner's needs. Brown D. (2007) suggests physical activities (TPR or Total Physical response[1] activities, for example), hands-on activities in which they can use meaningful language, sensory aids like realia where children get to experience language through touch or smell, as well as non-verbal language which is also essential in teaching any age group.

---

[1] TPR (Total Physical Response) is a language teaching method created by Dr. James J. Asher to mimic the way in which children learn their mother tongue. According to this method, the teacher plays the role of the parent and initiates 'language-body conversations'. The teacher instructs and the child physically responds to that instruction. For example, the teacher says 'Look at the board!' or 'Jump!' and the child does so. After some time, the child is able to reproduce language spontaneously ('TPR-Total Physical Response', 2004).

2.1.2. First and second language acquisition

The ability of acquiring a language has been a topic of interest to linguists and psychologists all over the world. The very process of language acquisition differs depending on whether we are talking about first language acquisition or second language acquisition.

First language acquisition is the process by which children acquire their first language or languages (Lightbown and Spada, 2006).

Second language acquisition is the process by which people learn a language other than their first language, whether that be as a result of living in a country where that language is spoken, or as a result of learning it in a classroom through instruction (Ellis, 1997).

First and second language acquisition are two different phenomena. Although they do have some similarities the very process of learning is different. Stephen Krashen argues that "adults can 'acquire,' which is the way children 'get' their first language, subconsciously, through informal, implicit learning. Once you have acquired something you're not always aware you have done it. It just feels natural; it feels as if it has always been there. Quite distinct from acquisition is conscious learning. This is knowing about language, explicit, formal linguistic knowledge of the language" (1982; according to Blair, 1982; pg. 17).

First language acquisition is a process that happens naturally and instinctively. It is very interesting to notice that this process is extremely similar with children all over the world, regardless of what their first language is. First language acquisition is very instinctive, rapid and complete. First sounds and vocalizations made by babies are the initial traces of a first language. Cooing and gurgling sounds that babies make are their first attempts of communication. By the age of two, most children already produce a variety of different words which they try to put in order to make meaningful sentences. This is known as telegraphic speech. By the age of four most children can ask questions, give commands, and tell stories while using grammatically correct constructions most of the time. During the pre-school and school years most children also develop metalinguistic awareness which can be defined as 'the ability to treat language as an object separate from the meaning it conveys' (Lightbown and Spada, 2006; pg. 8). This means that children at this age would be able to understand why

certain constructions do not make sense although they are technically grammatically correct, or that certain words and sentences can have ambiguous meanings. During the school years the child develops extensive vocabulary in various different registers and is able to produce fully correct grammatical structures (Lightbown and Spada, 2006).

Second language acquisition, on the other hand, is a process that happens through conscious effort and instruction. It is important to mention that 'second' and 'foreign' are not meant to be contrasted in this context. It can 'refer to any language that is learned subsequent to the mother tongue. Thus, it can refer to the learning of a third or a fourth language. Whether you are learning a language naturally as a result of living in a country where it is spoken, or learning it in a classroom through instruction, it is customary to speak generically of 'second' language acquisition (Ellis, 1997; pg. 3). When learning a second language children do not have to go through the 'telegraphic speech phase'; in fact, they can start by using full sentences depending on the learning conditions or instruction. Lightbown and Spada (2006) argue that all second language learners have already acquired at least one language; therefore, they are familiar with how language works. Although this could be an advantage, it could also create confusion and learner errors since no language functions in the same exact manner. In addition to this, although second language learners have already learned at least one language and have some knowledge of how language works, they still seem to be less successful in comparison with first language learners.

Language acquisition has been a topic of interest since the dawn of the time, and various linguists have developed theories on how language is learned and acquired. Piaget's Cognitive Development Theory, Vigotsky's Social Development Theory, Skinner's Behaviorist Theory, Chomsky's Innatist Theory, Critical Period Hypothesis and Krashen's Monitor Model are only some of the most famous ones.

2.1.3. *Important theories and perspectives*

Cognitive Development Theory
Developmental psychologists argue that language acquisition is dependent on the environment of the child. In their opinion, language acquisition is no different from

acquisition of any other skills or knowledge; in fact, it is influenced by them rather than it being a separate process. Jean Piaget, a Swiss psychologist, is one of the earliest advocates of the idea that children's language is influenced by their cognitive development. Piaget spent some time observing children while they were playing or interacting with other children or objects. According to his findings, children first become aware of a concept or an object and then they acquire language needed to convey the meaning. Therefore children's cognitive development influences the way they use language (Lightbown and Spada, 2006).

Piaget also devised a framework which suggests that there are four stages of development that all children go through. He argues that each child follows these stages in the same order, and 'development unfolds as a result of the biological processes of growth, and the development of the child's brain' (Pinter, 2006; pg. 6). The four stages in question are: sensory-motor stage (from birth to two years of age), pre-operational stage (from two to seven years of age), concrete operational stage (from seven to eleven years of age), and formal operational stage (from eleven years onwards) (Pinter, 2006).

Social Development Theory

This theory argues that social interaction plays an essential role in the cognitive development of a child. While observing interaction among children in their schools, Lev Vygotsky, a Russian psychologist, concluded that language develops primarily from social interaction. He argued that there is a metaphorical concept in which children would perform better than they would ever be capable of independently. A supportive and interactive environment would ensure peak performance. Vygotsky named this metaphorical concept as the zone of proximal development (Lightbown and Spada, 2006).

Pinter (2006) argues that this zone of proximal development represents the difference between the current knowledge of the child and the potential knowledge that the child would be able to achieve with the help of a more knowledgeable or experienced person. Zone of proximal development is a fertile ground for learning because it builds on child's immediate needs based on the existing knowledge.

Behaviorist Theory

B.F. Skinner, and American psychologist, is the best known advocate of behaviorism with regards to language learning. He describes language learning as behavior evoked by outside

stimuli that can come in forms of positive reinforcement or negative reinforcement. This would imply that children learn languages based on imitation, repetition as well as reinforcement of successful use. Behaviorists argue that when children imitate the language used by those around them they are usually met by praise and encouragement. This is positive reinforcement. What positive reinforcement does is encourage children to keep imitating and producing sounds and patterns which form correct language habits. Negative reinforcement is used to correct the improper uses of language (Lightbown and Spada, 2006).

This theory is also applied to second language acquisition. According to this approach, second language learners imitate sentences and patters in order to create correct language habits. However, learners also try to relate the knowledge of their first language and apply it to the second language. This could lead to 'positive transfer' or 'negative transfer' in language. Positive transfer is a result of similarities between learners' first language and their second language while negative transfer is a result of differences between the two languages. This is why behaviorism is associated with Contrastive Analysis Hypothesis (CAH) which states that learners should acquire target language with ease if their first language and their target language are similar, and vice versa (Lightbown and Spada, 2006).

Innatist Theory

Noam Chomsky, an American linguist, proposed the idea that all languages are innate and that all children are biologically programmed for language. In his opinion, it would be impossible for a child to learn correct grammatical structures and complex syntax solely by imitating language around them. Young children are usually exposed to broken language as adults tend to use simplified language when talking to children. In addition to this, a child's brain is not a blank slate ready to be filled with knowledge. Instead, they are born with the ability to acquire and understand language. This ability is called 'universal grammar' and it prevents children from developing incorrect language habits (Lightbown and Spada, 2006).

Chomsky defines universal grammar as 'the system of principles, conditions and rules that are elements or properties of all human languages' (Lightbown and Spada, 2006; pg. 35). This system allows children to acquire any language they are exposed to during the critical period of their development. Although Chomsky never clarifies whether this theory can be applied to

second language acquisition, other linguists have made it clear that it offers a great way of understanding second language acquisition.

Critical Period Hypothesis

Chomsky's ideas are also linked to the Critical Period Hypothesis (CPH) which claims that 'animals, including humans, are genetically programmed to acquire certain kinds of knowledge and skills at specific times in life. Beyond those critical periods it is either difficult or impossible to acquire those abilities' (Lightbown and Spada, 2006; pg. 17).

According to this theory, children should be exposed to language throughout their childhood until puberty for the best acquisition results. If children are not exposed to language during this period it would be impossible for them to fully acquire it. Since the majority of children are exposed to at least some type of language it would be impossible to prove this theory as incorrect (Lightbown and Spada, 2006).

The idea of the 'critical period' can also be applied to second language acquisition. Younger children can acquire more than one language simultaneously (childhood bilingualism), as well as learn a new language easier than older children. This theory is also supported by the fact that it is very difficult to achieve native-speaker fluency in a second language if learned after puberty (Lightbown and Spada, 2006).

Language acquisition is a unique process which cannot be confined to only one truth, or one theory. All of these theories serve to offer different perspectives on the very process of acquisition which is a crucial part of understanding the importance of language assessment. Language assessment must be seen as a unique process which requires different methods and techniques in order to produce reliable results. Just like language teaching, language assessment must be focused on the needs of the learner.

2.2. ASSESSMENT

With the rise of communicative, holistic and humanistic approaches, language teaching has been increasingly focused on the needs of the learner. However, assessment is not a priority to the majority of language teachers. A huge number of studies conducted in a variety of countries across Europe indicate that language teachers are either not skilled in language assessment, or do not see assessment as a valid focal point of language teaching (Edelenbos

and Johnstone, 1996; Nikolov, 2000; Hill, 2000; Low, Brown, Johnstone and Pirrie, 1995; according to McKay, 2006). However, assessment is a crucial part of the teaching process which offers insight into learners' achievements and progress, as well as information on what needs to be improved and done by the teacher.

2.2.1. Why is assessment important?

'Assessment has the power to change people's lives' (Shohamy, 2001; according to McKay, 2006; pg. 18). The effect that assessment has on students, especially young learners, is crucial for their language learning experience. Shaaban (2005) defines assessment as a tool that helps teachers determine learners' linguistic abilities and place them in appropriate proficiency levels. The success of assessment is determent by the appropriate selection of tools and strategies, as well as the interpretation of the learners' performance. Pinter (2006; pg. 131) defines assessment as a 'process of data analysis' used by teachers to track their learners' progress and achievement.

According to McKay (2006), effective assessment is that which is designed to provide valid and fair information about the learners' abilities and progress. Receiving this information is the main purpose of assessment. Assessment can yield both positive and negative impact, depending on various factors Effective assessment provides the teacher with the insight into learners' achievements and progress that they have made over the course of time. It also helps teachers track the attainment of their goals, aims and objectives. Parents receive valuable feedback on their child's performance and progress. The administration receives feedback on the effectiveness of the curriculum. Learners receive feedback on what they have been doing well, or what they need to work on to achieve even better results. Effective assessment also provides learners with motivation to work harder, and confidence in their abilities.

The majority of definitions of assessment offer similar explanations but the most essential point to be taken is that assessment is an ongoing process. Very often teachers seem to confuse assessment with simply grading a learner's performance, or testing it through written or oral format. There is a significant difference between all these concepts.

Assessment is an ongoing process that never stops. During a single lesson the majority learners will make comments, use new words and structure, ask questions and overall interact with the teacher and the rest of the class in the target language. A good teacher never stops assessing students, whether that be incidental or intentional. Teachers are always making assessments of what is going on in the classroom and keeping track of every learner's progress (Brown, 2007).

Evaluation is a process similar to assessment which is why the two terms have been treated as synonymous over the course of time. Assessment helps us determine whether any achievement has been made while evaluation determines the degree of the said achievement. Evaluation could be defined as a process of measuring achievement based on standard criteria (Brown, 2004)

Testing is a method of measuring learners' abilities and achievements in a given domain. A test is an instrument that is carefully designed to serve the purpose of measuring achievement. Tests are 'prepared administrative procedures that occupy identifiable time periods in curriculum when learners muster all their faculties to offer peak performance, knowing that their responses are being measured and evaluated. (Brown, 2007)

Grading can be seen as the end product of the entire process of assessment and evaluation. Grades are often seen as the ultimate indicator of learners' knowledge, however, that is not the case. There are many other factors that might go into the final grade and they cannot be seen as the reflection of knowledge or as the end goal of learning (Brown, 2004).

2.2.2. *Assessment principles*

Having in mind that assessment in itself is a process, there are many things that go into its realization. H. Douglas Brown (2004; 2007) offers five main principles which serve as a guideline for constructing tests or other assessment strategies to ensure maximum efficiency. As Brown mentions, he selected these five principles because various assessment specialists list them as a priority for the design of language assessment. The five principles in question are practicality, reliability, validity, authenticity and washback.

Table 1
Assessment principles (Brown, 2004)

ASSESSMENT PRINCIPLES
Practicality
reliability
validity
authenticity
washback

Assessment needs to be *practical.* What this means is that a good test (or any other assessment strategy) should not take too much time to be completed or to be evaluated. It is also important for it to be within the financial or any other limitations of the school or the teacher. For example, if the school does not possess computers students cannot be expected to undergo tests that require the use of the computer (Brown, 2004; 2007).

Assessment needs to *reliable*. There are a variety of factors that might influence an unreliable test and they might be related to the test itself, to the teacher, to the learner, to the administration of the test or to the scoring of the test. If the classroom where assessment is taking place is noisy, dark, or stifling it could influence learners' performance which means that the results could not be considered reliable. Learners themselves might be going through private issues that influence their performance. The teacher in charge of grading the tests might also be going through private issues, or might be running late on grading everyone on time. All of these factors could influence assessment reliability (Brown, 2004; 2007).

Assessment needs to be *valid*. Validity is one of the most complex principles of assessment, as it represents 'the degree to which the test actually measures what it is intended to measure' (Brown, 2004; pg. 448). Validity is very difficult to determine. Brown argues that although there are standard measures that could be used to determine validity, one can only rely on observation and theoretical justification. Ultimately, it is up to the teacher to recognize whether the test conveyed accurate degree of learners' knowledge. Brown also draws attention to three types of validity: content validity, face validity and construct validity. Content validity means that the test requires the learner to perform the exact action that is being measured. If a teacher is trying to measure learners' conversational skills by using a test

that requires written answers on multiple-choice questions, the test does not achieve content validity. Face validity is closely related to content validity. In order to achieve face validity learners taking the test have to be convinced that the content of the test truthfully represents what is being measured. If they feel like the results of the test will not show their true knowledge of a particular topic, the test does not achieve face validity. Lastly, construct validity requires the test to be able to define particular theoretical constructs. As Brown suggests, every issue in language learning can be seen as a theoretical construct (proficiency, communicative competence, self-esteem etc.). In other words, in order for a test to achieve construct validity every single component of the test needs to be justified in theory (Brown, 2004; 2007).

Authenticity suggests that tasks given in the test are likely to be enacted in the 'real world'. Contrary to the popular belief, assessment should be as enjoyable as possible. Boring, unrelatable topic and tasks will not motivate learners; quite on the contrary, they will not provide reliable results. The language of the test should be as natural as possible and tasks should be contextualized and relevant to the learner age group (Brown, 2004; 2007).

Washback can be seen as feedback that learners get after completing a test. Washback also includes the effects of the preparation for assessment, not only of the assessment itself. When it comes to informal assessment, feedback is a natural effect of interaction between the teacher and learners. However, with formal assessment (especially with written assignments) feedback seems to be either missing or reduced to a number, letter or a simple phrase. By giving learners proper feedback on their performance, the teacher is enhancing their motivation, autonomy, confidence, language ego and overall helping them understand what they did well and what they should work on in the future (Brown, 2004; 2007).

It is important to notice that not all authors agree or emphasize these five principles. Jeremy Harmer (2005) suggests that validity and reliability are two most important characteristics when judging the effectiveness of any test. W. James Popham (2003) agrees, as he also emphasizes these two principles.

Lastly, it is up to teacher to decide which of these principles are to be considered when choosing the right assessment method or technique.

2.2.3. Assessment types

There are many reasons why learners might be assessed. These reasons offer a criteria based on which we differentiate a variety of types of assessment. Since assessment is an ongoing process there are many ways in which learners are assessed, whether that be informally or formally.

Informal assessment refers to assessment that is incidental, unplanned, and carried out in the classroom during the course of teaching and learning (McKay, 2006). Informal assessment can be carried out in multiple ways. Simple comments and feedback given by the teacher during the lesson are a form of assessment because they enable the teacher to instruct learners on how to proceed with their learning. Douglas Brown (2004) also argues that a great deal of informal assessment is carried out through various tasks designed by the teacher, not to grade learners, but to simply monitor their progress and give them feedback.

Formal assessment, on the other hand, refers to assessment that is carefully planned out. Formal assessment follows formal procedures which are designed to estimate learners' achievements and knowledge. Formal assessment happens periodically and is usually announced in time for learners to prepare for it (McKay, 2006). It is important to mention that although tests fall under the category of formal assessment they are not synonymous. All tests are formal assessment but not all formal assessment is carried out by tests (Brown, 2004).

Formative assessment 'includes evaluating students in the process of 'forming' their competencies and skills with the goal of helping them continue that growth process' (Brown, 2004; pg. 6). Formative assessment is mainly informal and it gives the teacher information on how well learners are doing. McKay (2006) mentions that formative assessment also includes diagnostic assessment and on-the-run assessment. Diagnostic assessment refers to instances where teacher analyzes learners' specific strengths and weaknesses. On-the-run assessment refers to evaluation carried out on the spot with immediate feedback by the teacher.

Summative assessment aims to provide a summation of what a learner has learned, usually performed at the end of a unit, a course, or a semester. The main issue with summative assessment is that it usually evokes negative connotations. Learners prepare for summative tests looking forward to forgetting all the information after the test. It is important to keep in

mind that not all tests are summative, and not all tests have to be overwhelming. It is up to the teacher to figure out how to make tests less intimidating (Brown, 2004).

Norm-referenced assessment refers to the type of assessment where 'the grading of each pupil's performance is related to the performance of others' (Kyriacou, 2007; pg. 107). Norm-referenced tests are usually designed to measure global language abilities. James D. Brown (1996) argues that the main purpose of such tests is to spread out learners along a continuum of different scores so that those who have lower abilities in a certain area are placed at one end of the spectrum, while those who have higher abilities are placed at the other end. A typical example of a norm-referenced test is TOEFL (Test of English as a Foreign Language) which is a standardized test intended for assessment of the knowledge of English language of a large number of test-takers. Tests like this one have already fixed and determined responses, and are made in a way that can be quickly and efficiently scored with minimum expanse. Money and efficiency are main concerns of norm-referenced tests (Brown, 2004).

Criterion-referenced assessment refers to the type of assessment where learners are graded based on fixed criteria. This means that all learners who meet the criteria would be graded accordingly regardless of the performance of other learners (Kyriacou, 2007). In addition to this, Douglas Brown (2004) suggests that criterion-referenced tests are also designed to give feedback on specific course, whether that be in form of a grade or a comment. Such tests require more time and effort from the teacher.

Performance-based assessment is centered on the learner. Rather than assessing learners solely based on written tests, 'performance-based assessment typically involves oral production, written production, open-ended responses, integrated performance (across skill areas), group performance, and other interactive tasks' (Brown, 2004; pg. 11). Naturally, these types of tests are more expensive and time consuming; however, they pay off because learners get to show off their knowledge in real-world situations and contexts. This is a much more authentic type of assessment (Brown, 2004).

Standards-based assessment is based on a set of pre-defined standards which determine different levels of learner achievement. These standards are usually disclosed to learners so that they are aware against which criteria their performance is going to be judged. One of the most prominent characteristics of this type of assessment is the feedback. Learners do not get

a grade, a score or a percentage of the test they completed. Instead, they are offered an explanation of the expectation of their performance based on the said standard. The main idea behind standards-based assessment is to connect the achievement if learning to the learning outcomes (Brown, 2004).

Classroom assessment, as the name implies, is assessment carried out and prepared in the classroom. There are instances when such assessment is carried out to obtain information needed for those outside the classroom; meaning, parents or the administration (McKay, 2006). Kyriacou (2007) refers to it as internal assessment and implies that such assessment is usually a part of teacher's own program of teaching.

External assessment is that which is prepared outside the classroom (but not necessarily carried out outside the classroom). It is usually prepared and graded by examiners outside the school, or graded by the teacher but double-checked by the examiners (Kyriacou, 2007).

Alternative assessment is a contrast to what is known as traditional assessment. While traditional assessment is more focused on standardized tests and pen-and-paper examinations, alternative assessment relies on performance, the ability to use knowledge to create something. Brown and Hudson (1998) argue that using the term 'alternative assessment' is counterproductive as it implies that something new is being used rather than what is required for responsible test construction. They suggest using the term 'alternatives in assessment' instead. Some of suggested alternative techniques of assessment are nonverbal responses, oral interviews, written narratives, presentations, student-teacher conferences, self-assessment, peer and group assessment, and student portfolios (Shaaban, 2005).

McKay (2006) also mentions assessment types based on their purpose. *Assessment for placement purposes* refers to instances where assessment is used to place newly transferred children to appropriate classes or groups. When it comes to young learners, she argues, assessment of this kind might involve an interview, a short reading session and a written task. *Assessment to encourage and motivate* includes assessment procedures that are structured in a way to encourage children by showing them what they have learned and giving them positive feedback. Since tests can very often discourage and demotivate children it is very important to devise tasks which will do the opposite. *Assessment for research*, as the name suggests, is

carried out for the purposes of research. Researchers usually want to find out details about language acquisition or the assessment decisions that teachers make (McKay, 2006).

Kyriacou (2007) mentions a couple of contrasting pairs of assessment. *Continuous assessment* is assessment made over a period of time consisting of different pieces of assessment, while *terminal assessment* is assessment made at the end of a course, or a semester, or a year. *Objective assessment* is assessment solely made on concrete evidence of accurate performance based on a particular standard or criteria, while *subjective assessment* is influenced by the teacher's subjective judgment of the learner or the performance (Kyriacou, 2007).

There are so many different types of assessment that it might be difficult for the teacher to decide which of these would be the best choice. Assessment types should be chosen based on the needs of the learner, the course or even the teacher. Depending on what the teacher wants to achieve with a particular method, it should be fairly easy to choose the right method if he or she is familiar with all of these different types. It would be best if all of these types of assessment could be combined throughout the course so that learners could truly show their full potential and knowledge.

2.2.4. Alternatives in assessment

The difference between 'traditional' and 'alternative' assessment is not always clear. As James Brown and Thom Hudson (1998) stated in their article, the term 'alternative assessment' is problematic as it implies something new and different rather than something that follows all the rules of proper assessment construction. Traditional assessment, on the other hand, normally refers to methods that have been used and standardized over the course of time. When put in this dichotomy of 'traditional' versus 'standardized', traditional assessment usually carries negative connotations as it reminds of something that is old and ineffective. Douglas Brown (2004) argues that it is impossible to even differentiate which methods of assessment fall into which category. Not only that, he also argues that it would be completely incorrect to assume that one is superior to the other.

Brown (2004) offers a summary of characteristics of traditional and alternative assessment in order to illustrate their differences (adapted from Armstrong, 1994, and Bailey, 1998;

according to Brown, 2004). The table should be read with caution as it contains overgeneralizations and bias towards alternative assessment.

Table 2
Traditional and alternative assessment (Brown, 2004; pg. 13)

TRADITIONAL ASSESSMENT	ALTERNATIVE ASSESSMENT
One-shot, standardized exams	Continuous long-term assessment
Timed, multiple-choice formal	Untimed, free-response format
Decontextualized test items	Contextualized communicative tasks
Scores suffice for feedback	Individualized feedback and washback
Norm-referenced scores focus on the "right" answer	Criterion-referenced scores
	Open-ended, creative answers
Summative	Formative
Oriented to product	Oriented to process
Non-interactive performance	Interactive performance
Fosters extrinsic motivation	Fosters intrinsic motivation

Nowadays, traditional assessment methods are usually associated with conventional testing. This may include standardized tests, multiple-choice tests, true or false tests, matching tests, fill-in-the-blank tests and other types of pen-and-paper tests (Brown, 2004). These tests are very efficient, inexpensive and easy to correct. However, traditional methods cannot be the only ones used in evaluating language competencies, especially when working with young language learners. In the Table 1, Brown (2004) makes a comparison between what we stereotypically consider traditional assessment and alternative assessment. He associates traditional assessment with standardized tests comprising of exercises constructed with no real-life context, tests centered around the 'right answer', and grade-oriented motivation. On the other hand, he presents alternative assessment as being long-term, contexualized, individualized, and overall centered around the needs of the learner. Such assessment fosters intrinsic motivation and offers valuable feedback to both the teacher and the learner. However, even Brown (2004) states that this comparison should be taken lightly as it suggests that one is better and more efficient than the other. Traditional and alternative assessment

should not be separated; they should be combined and used depending on the needs of the learner or the course.

2.2.5. *Assessing young learners*

Assessing young learners is quite different from assessing any other age group. Their knowledge of a language is often expressed through songs and chants, random sentences and phrases, stories and games. It would be impossible to assess them solely based on a pen-and-paper test, especially when many of them are not yet good at writing. It would be counterproductive to reduce their knowledge to a quantifiable test result (Pinter, 2006). Shaaban (2005) also argues that because young learners are poor test-takers teachers are running at risk to assign false labels to them and affect their confidence and motivation. Tests cause a great deal of anxiety because they seem too formal and forced. An anxiety-free environment could be achieved by helping children perceive assessment 'as an integral component of the learning/teaching process rather than an independent process whose purpose is to pass judgment on their abilities in relation to their classmates' (Shaaban, 2005; pg. 36). Alternative assessment methods which seem to be more performance based could be a great way of achieving this goal. Some of alternative methods of assessment that can be used in a classroom are nonverbal responses, oral interviews, written narratives, role-play, observations, projects and presentations, student-teacher conferences, self-assessment, peer and group assessment, and student portfolios.

Nonverbal responses are a great way of assessing very young learners. Nonverbal responses include the use of physical performance and pictorial products and are overall easy to carry out. This technique is great for lowering the anxiety with young learners as it seems like an extension of the learning process rather than evaluation (Shaaban, 2005).

Oral interviews refer to situations where the teacher interviews a student with the purpose of assessment. Interviews are a great way of assessing speaking skills but they can have various other purposes such as establishing learners' needs, discovering their styles and preferences, and requiring feedback. Interviews are great for lowering the anxiety because they can be designed in a way to seem like simple, informal conversations (Brown, 2004). Shaaban (2005) adds that interviews can also be very convenient when working with young

learners. Learners might be asked to choose a picture to talk about, and the teacher can ask question encouraging them to use related vocabulary (Shaaban, 2005).

Written narratives are a great way of assessing writing skills while allowing children to enjoy and express themselves creatively. Teachers can get creative with topics they assign as long as they are purposeful and authentic. Writing a letter to a friend, invitations to a birthday party or writing a short story about something that is interesting to them will keep them motivated (Shaaban, 2005).

Role play combines oral performance and physical activity. It is a great way of illustrating the topic of learning (whatever it may be). Children all ages typically enjoy such activities because they make them feel comfortable and motivated. It is a great way for teachers to get creative with topics and use the time to observe learners' language use (Shaaban, 2005).

Observations are a great way of assessment because learners are usually not even aware that they are being assessed. Depending on the purpose of assessment, teachers can observe children while doing various different activities whether that be individually, in pars or in groups. They are also a great way of assessing non-linguistic skills (such as engagement, motivation and interest) (Pinter, 2006). McKay (2006) argues that observations can be incidental and planned. Incidental observations simply occur as a part of the teaching process. 'During story time, for example, the teacher scans the listeners for facial expressions and body language and listens for verbal responses indicative of enjoyment, language development and comprehension…There are innumerable incidental observations inherent in day-to-day interactions with children. These incidental observations provide valuable information about what individual students are feeling, thinking, understanding, and guide the responsive teacher in setting appropriate expectations and experiences for them' (Puckett and Black; 2000, pg. 217; according to McKay, 2006). Planned observations, on the other hand, are recorded systematically and over time, typically on checklists or rating scales. This way all changes in student performance are noted (McKay, 2006).

Projects and presentations are a great way of assessing team work skills as well as oral and written performance. Projects offer an opportunity for children to work together and learn from each other (if the members of the group are carefully selected). Project work is a great way for children to show their non-linguistic skills (drawing, singing, acting) which spikes up

their motivation and confidence. However, project work is difficult to grade fairly as the teacher has to take into consideration individual work and group effort. This is why it would be better to use this form of assessment to give feedback and praise (Pinter, 2006). Presentations give the teacher insight in learners' interests and abilities as well as their creative side. Nowadays, they are even more popular because of all of the technological advances (Shaaban, 2005).

Conferences refer to focused discussions between the teacher and learners. These discussions usually focus on a particular piece of work or a portfolio. Conferences are great for tracking learners' progress and enabling them to reflect on their work (Pinter, 2006). Shaaban (2005) argues that conferences are best and most effective if done after observations. This way they concentrate directly on the learning process and its results in the real time.

Self-assessment might seem inappropriate for young language learners but it can lead to great discoveries if done correctly. Shaaban (2005) argues that it can be done by using one of the two techniques: K-W-L charts or learning logs. K-W-L charts provide examples of what learners 'know, wonder, or what they have learned'. A learning log is a record of learners' experience with language outside the classroom. Self-assessment is a great way for learners to reflect on their own experience with language, to give their own opinion on their progress, and to give them confidence in their own understanding of the process of learning (Pinter, 2006). Brown (2004) argues that self-assessment might seem like an 'absurd reversal of politically correct power relationships' (Brown, 2004; pg. 270) especially when talking about young learners, but self-assessment has been proven as beneficial and necessary in the learning/teaching process. It derives its justification from a number of principles of second language acquisition, such as autonomy and intrinsic motivation.

Group and peer-assessment develop learners' ability to work cooperatively with others in a group. Students could be encouraged to write positive evaluations of their peers' work, along with constructive criticism or advice on how to improve what they consider needs improving. Such assessment could be graded as a group effort to emphasize the importance of their team work (Shaaban, 2005). When it comes to young language learners it might be challenging to introduce self- and peer-assessment; however, with some effort put into teaching them how to comment on their work and their friends' work it would be possible.

One of the best ways for trying out this method with younger learners would be through assessment charts and sheets. They are easy to follow and convenient for younger learners (McKay, 2006).

Portfolios are borrowed from the field of fine arts where they were used as a collection of artist's best works (Brown, 1998; according to Shaaban, 2005). Their purpose in language teaching is to demonstrate learner's samples of oral and written work in the target language. Since these samples are collected over time they are a great way of tracking learner's progress and language development. Portfolios as such can be shown to students, teachers, parents and the administration as a systematic collection of student's work (Shaaban, 2005). McKay (2006) argues that in order for portfolios to become an assessment strategy, there needs to be a plan to select tasks for assessment and collection, and to collect materials systematically. Portfolios have already formed a strong component of assessment in primary schools as they have been proven as convenient and effective when it comes to young learners.

2.2.6. *Assessing speaking, listening, reading and writing*
Assessing speaking

Young language learners practice their language skills primarily through spoken interaction; which, if planned correctly, can be done through various activities, imaginative plays, storytelling, rhymes and songs, and participation in simple discussions. Teachers need to be able to assess oral language in tasks such as interviews, or even pair and group work comprising of such activities (McKay, 2006).

As children develop their oral language skills, teachers need to be familiar with main characteristics of successful performance in order to be able to assess them. McKay (2006) suggests using Bachman and Palmer's model of language knowledge to describe these characteristics. The model suggests that learner's development of oral language should include the following areas of language knowledge: grammatical knowledge (knowledge of vocabulary, syntax and phonology), textual knowledge (ability to speak in a cohesive and organized manner, both in conversation interactions and in extended speaking turns), functional knowledge (ability to use language for many different functions), and sociolinguistic knowledge (ability to use oral language appropriate to the language use

situation that they are in). These characteristics can be used as a checklist or as objectives of assessment (McKay, 2006).

Brown (2004), on the other hand, suggests microskills and macroskills of speaking as an objective for language assessment. 'Microskills refer to producing the smaller chunks of langauge such as phonemes, morphemes, words, collocations, and phrasal units. Macroskills imply the speaker's focus on the larger elements: fluency, discourse, function, style, cohesion, non verbal communication, and strategic options' (Brown, 2004; pg. 142). He offers a list of sixteen different objectives to assess speaking:

Table 3
Micro- and macroskills of oral production (adapted from Richards, 1983; according to Brown, 2004; pg. 142)

MICROSKILLS	MACROSKILLS
1. Produce differences among English phonemes and allophonic variants.	12. Appropriately accomplish communicative functions according to situations, participants, and goals.
2. Produce chunks of language of different lengths.	13. Use appropriate styles, registers, implicature, redundancies, pragmatic conventions, conversation rules, floor-keeping and -yielding, interrupting, and other sociolinguistic features in face-to-face conversations.
3. Produce English stress patterns, words in stressed and unstressed positions, rhythmic structure, and intonation contours.	14. Convey links and connections between events and communicate such relations as focal and peripheral ideas, events and feelings, new information and given information, generalization and exemplification.
4. Produce reduced forms of words and phrases. 5. Use an adequate number of lexical units words) to accomplish pragmatic purposes.	15. Convey facial features, kinesics, body language, and other nonverbal cues along with verbal language.
6. Produce fluent speech at different rates of delivery.	16. Develop and use a battery of speaking strategies, such as emphasizing key words, rephrasing, providing a context for interpreting the meaning of words, appealing for help, and accurately assessing how well your interlocutor is understanding you.
7. Monitor one's own oral production and use various strategic devices- pauses, fillers, self-corrections, backtracking to enhance the clarity of the message.	
8. Use grammatical word classes (nouns, verbs, etc.), systems (e.g., tense, agreement, pluralization), word order, patterns, rules, and elliptical forms.	
9. Produce speech in natural constituents: in appropriate phrases, pause groups, breath groups, and sentence constituents.	
10. Express a particular meaning in different grammatical forms.	
11. Use cohesive devices in spoken discourse.	

Lastly, it is important to be able to select appropriate tasks that would be best suited for oral language assessment. McKay (2006) suggests news telling (children presenting various

different news to the rest of the class), storytelling, picture talks (children describing various different pictures), categorization task (children sorting and finding patterns), and oral presentations.

Assessing listening

Listening assessment is very often ignored as it is usually combined with speaking assessment. However, listening is a big part of the language learning experience, and there are many instances in which listening needs to be assessed individually, especially in school contexts. Listening is more challenging to assess as it is normally an 'invisible' skill. Listening should be assessed indirectly, with the aim of finding evidence of listening comprehension in children's reactions and subsequent tasks (McKay, 2006).

Brown (2004) offers a list of micro- and macroskills that could serve as a checklist for assessment, or as objectives of assessment.

Table 4

Micro- and macroskills of listening (adapted from Richards, 1983; according to Brown, 2004, pg. 121)

MICROSKILLS	MACROSKILLS
1. Discriminate among the distinctive sounds of English.	12. Recognize the communicative functions of utterances, according to situations, participants, goals.
2. Retain chunks of language of different lengths in short-term memory.	13. Infer situations, participants, goals using real-world knowledge.
3. Recognize English stress patterns, words in stressed and unstressed positions, rhythmic structure, intonation contours, and their role in signaling information.	14.From events, ideas, and soon, described, predict outcomes, infer links and connections between events, deduce causes and effects, and detect such relations as
4. Recognize reduced forms of words.	main idea, supporting idea, flew information, given
5. Distinguish word boundaries, recognize a core of words, and interpret word order patterns and their significance.	information, generalization, and exemplification.
6. Process speech at different rates of delivery.	15. Distinguish between literal and implied meanings.
7. Process speech containing pauses, errors, corrections, and other performance variables.	16. Use facial, kinesic, body language, and other nonverbal clues to decipher meanings.
8. Recognize grammatical word classes (nouns, verbs, etc.), systems (e.g., tense, agreement/pluralization), patterns, rules, and elliptical forms.	17. Develop and use a battery o listening strategies, such s detecting key words, guessing the 'meaning o words from context, appealing for help, and signaling comprehension or lack thereof.
9. Detect sentence constituents and distinguish between major and minor constituents.	
10. Recognize that a particular meaning may be expressed in different grammatical forms.	
11. Recognize cohesive devices in spoken discourse.	

When preparing listening tasks for children, teachers need to be aware that children will not listen just because they were told to. Listening tasks need to have a purpose because children need to know why they are listening to something (example: 'Listen to the story and find out why the monster ran away!'). Teachers should be able to create a situation in which children could show that they understood the story. They should be asked to draw a picture, perform an action, and respond to a question in order to show their understanding. (McKay, 2006).

Some of tasks suitable for listening assessment are action tasks (activities like *Simon Says* were children are asked to listen and perform a task), total physical response tasks (children are asked to listen and perform a physical response to a request or command), true/false tasks, aural cloze (a written passage with deleted words where children need to listen and fill in the blanks), noting specific information, matching tasks, spot the mistake tasks, and dictations (McKay, 2006).

Assessing reading

Reading demands understanding of the text. According to McKay (2006), 'readers employ three main cuing systems when they read; they rely on graphophonic cues at the word level (i.e. cues from the way a word is written and how it 'sounds out'), syntactic cues at the sentence level (i.e. cues that give information about the role of any one word within a sentence or clause of words), and semantic cues at the whole text level (i.e.cues that relate to the meaning of a word or words in relationship with the whole text)' (McKay, 2006; pg. 223).

Reading assessment is heavily determined by the curriculum. The curriculum determines texts or types of text that children should be reading, as well as child's expected growth in reading ability. However, there are also frameworks that teachers could use to help them define their objectives for reading assessment. McKay (2006) suggests using Bachman and Palmer's model of language knowledge to define those objectives. The model suggests that learner's development of reading abilities should include the following areas of language knowledge: grammatical knowledge (ability to decode letters and words, and recognize letters and words by sight), textual knowledge (ability to read a range of texts, for the range of purposes, and to understand the meaning if the text), functional knowledge (ability to

understand purposes of the language), and sociolinguistic knowledge (sufficient knowledge and experience of the target language culture) (McKay, 2006).

Brown (2004) offers a list of micro- and macroskills that could serve as a checklist for assessment, or as objectives of assessment.

Table 5
Micro- and macroskills for reading comprehension (adapted from Richards, 1983; according to Brown, 2004, pg. 187)

MICROSKILLS	MACROSKILLS
1. Discriminate among the distinctive graphemes and orthographic patterns of English. 2. Retain chunks of language of different lengths in short-term memory. 3. Process writing at an efficient rate of speed to suit the purpose. 4. Recognize a core of words, and interpret word order patterns and their significance. 5. Recognize grammatical word classes (nouns, verbs, etc.), systems (e.g., tense, agreement, pluralization), patterns, rules, and elliptical forms. 6. Recognize that a particular meaning may be expressed in different grammatical forms. 7. Recognize cohesive devices in written discourse and their role in signaling the relationship between and among clauses.	8. Recognize the rhetorical forms of written discourse and their significance for interpretation. 9. Recognize the communicative functions of written texts, according to form and purpose. 10. Infer context that is not explicitly using background knowledge. 11. From described events, ideas, etc., infer links and connections between events, deduce causes and effects, and detect such relations as main idea, supporting idea, new information, given information, generalization, and exemplification. 12. Distinguish between literal and implied meanings. 13. Detect culturally specific references and interpret them in a context of the appropriate cultural schemata. 14. Develop and use a battery of reading strategies, such as scanning and skimming, detecting discourse markers, guessing the meaning of words from context, and activating schemata for the interpretation of texts.

Lastly, it is important for teachers to be able to select appropriate tasks for reading assessment. McKay (2006) suggests observations, interviews with parents (to find out children's reading habits), teacher-student reading conferences, oral reading, informal reading procedures (tests of contributing skills such as visual discrimination, listening comprehension, alphabet recognition, letter-sound correspondence etc.), miscue analysis (recording of differences between children's response and actual words on the page), portfolios and self-assessment.

Assessing writing

Writing assessment is challenging, not only because it is mainly determined by the curriculum, but also because it is determined by the purpose and the goal of the assessment. The teacher needs to decide whether he/she wants to assess vocabulary, grammar and spelling or writing abilities, creativity, and development of the main idea (Brown, 2004).

The teacher also has to be familiar with the characteristics of the development of writing ability. McKay (2006) suggests using Bachman and Palmer's model of language knowledge to define those characteristics. The model suggests that learner's development of writing abilities should include the following areas of language knowledge: grammatical knowledge (ability to form letters and characters with correct spelling), textual knowledge (ability to write across the range of topics and texts in the target language), functional knowledge (ability to write to achieve the purpose they wish to achieve), and sociolinguistic knowledge (knowledge and experience of the target language culture in order to convey their message to the audience). McKay suggests using these characteristics as objective of writing assessment.

Brown (2004), on the other hand, offers a list of micro- and macroskills of writing that could serve as a checklist when assessing writing abilities.

Table 6
Micro- and macroskills of writing (adapted from Richards, 1983; according to Brown, 2004, pg. 221)

MICROSKILLS	MACROSKILLS
1. Produce graphemes and orthographic patterns of English. 2. Produce writing at an efficient rate o speed to suit the purpose. 3. Produce an acceptable core o words and use appropriate word order patterns. 4. Use acceptable grammatical systems (e.g., tense, agreement, pluralization), patterns, and rules. 5. Express a particular meaning in different grammatical forms. 6. Use cohesive devices in written discourse.	7. Use the rhetorical forms and conventions o written discourse. 8. Appropriately accomplish the communicative functions o written texts according to form and purpose. 9. Convey links and connections between events, and communicate such relations as main idea, supporting idea, new information, given information, generalization, and exemplification. 10. Distinguish between literal and implied meanings when writing. 11. Correctly convey culturally specific references in the context o the written text. 12. Develop and use a battery o writing strategies, such as accurately assessing the audience's interpretation, using prewriting devices, writing with fluency in the first drafts, using paraphrases and synonyms, soliciting peer and instructor feedback, and using feedback for revising and editing.

Tasks that could be used for writing assessment in a classroom setting are observations, writing conferences, portfolios, self-assessment, writing in speech bubbles (for younger learners to fill in the speech bubbles in a cartoon story), open response writing (children are required to write about someone or something), and re-forming a text (children are asked to re-form a text in a new genre) (McKay, 2006).

When assessing language competencies it is very helpful to keep track of each specific skill (listening, reading, writing and speaking). Children very often develop these skills unevenly, meaning that some children might excel at speaking and listening but might face difficulties with reading and writing. It is important for a teacher to be able to distinguish between these skills, teach and assess them separately, as well as be able to devise tasks that are going to be best suited to assess or evaluate particular skills.

2.2.7. *Evaluation and marking*

It is quite ironic to conclude that the entire learning process discussed in the previous pages can be, and is, reduced to a simple letter on the paper. For the majority of learners grades are the most important part of the process; their entire academic achievement and sense of self-worth are summed up in a single grade. In addition to this, Brown (2004) argues that learners' educational lives are controlled by grades. 'Educational systems define honors students, marginal students, college-bound students, exceptional students (on either end of the scale), failing students, and average students not so much by the quality of their performance(s) and not necessarily by demonstrated skills that have been observed, but rather by grades' (Brown, 2004; pg. 282).

There are a lot of teachers and professors that add to this distorted image of the value of grades. Unfortunately, there are cases where a teacher's worth is seen through statistics of how many students are failing their subject. Very often the main difference between 'easy' teachers and 'tough' teachers are their grading standards (Brown, 2004). So what should grades reflect?

This is a difficult question to answer. When dealing with language learning in particular, there are many different aspects of the learning process that could and should be assessed and graded. The final grade could comprise of language performance of the student as formally demonstrated on tests, quizzes, and other explicitly scored procedures, teacher's intuitive, informal observation of the student's language, oral participation in class, improvement, behavior in class, effort, and/or motivation (Brown, 2004).

There have been many debates over the course of time discussing whether a final grade should include all of these aspects, or whether it should only include formal achievement. One of the best and widely known assessment specialists, Norman Gronlund, suggests: 'base grades on student achievement, and achievement only. Grades should represent the extent to which the intended learning outcomes were achieved by students. They should not be contaminated by student effort, tardiness, misbehavior, and other extraneous factors: If they are permitted to become part of the grade, meaning of the grade as an indicator of achievement is lost' (Gronlund, 1998; pg. 174-175; according to Brown, 2004; pg. 284).

Stripping the grade of all of the 'external' aspects and leaving it solely based on formal achievement seems like good practice that would help achieve fair assessment. However, if assessment is an ongoing process with the goal of showing true evaluation of learners' achievement, it cannot be stripped of all the important factors that led to the final result. Motivation, behavior, improvement and effort are all a part of the process. If the final grade is based solely on test achievement the true purpose of assessment will be lost.

McKay (2006) argues that in order to evaluate learner's performance in the best way possible teachers need to come up with a scoring method. 'The scoring method consists of (1) the criteria by which students' responses are evaluated and (2) the procedures followed to arrive at a score' (Bachman and Palmer, 1996; according to McKay, 2006; pg. 266). Criteria typically refer to the definitions of what constitutes a successful performance. Criteria can be written in various different ways. It can be written in headings, statements or questions. It can be written broadly or specifically. Teacher chooses how to write criteria depending on what is being assessed. The more specific the criterion, the more objective the final assessment can be. However, when it comes to language assessment it would not be informative to go into specific details of a child's performance. When writing language assessment criteria, McKay (2006) suggests accepting a lesser degree of specificity. Language involves various different elements that indicate whether or not the performance was successful. Since the criteria are rarely explicit and precise, the final decision of the performance depends on the teacher.

When it comes to procedures that help decide on the final grade or a score, teachers have to decide whether there are going to only be 'correct' and 'incorrect' answers, or are there going to be varying degrees of correctness to each answer. Teachers typically go for the 'correct' and 'incorrect' route and come up with a final grade by adding the correct answers together (McKay, 2006).

Contrary to the popular belief, learners can also be graded in many different ways. The most popular ones are letter grades and number grades, but learners can also be graded with words, phrases, comments, feedback and many other alternatives in grading (Brown, 2004). Meaning behind letter grades and number grades is typically identical:

Table 7
Letter and number grades meanings (Brown, 2004; pg. 294)

LETTER GRADE	NUMBER GRADE	MEANING
A	5	Excellent
B	4	Good
C	3	Adequate
D	2	Inadequate/Unsatisfactory
F	1	Failing/ Unacceptable

However, there are some controversies regarding letter and number grades and the meaning behind them. Many argue that simple phrases such as 'excellent' or 'unacceptable' do not serve the true purpose of assessment. They do not provide sufficient information about learners' strengths and weaknesses; neither do they suggest alternatives and offer help. Since grades have such a huge impact on a learner's educational life, such overgeneralizations should not be permitted. If letter and number grades are used, the lack of information and feedback should be compensated.

Brown (2004) suggests some alternatives to letter and number grades. For formative assessment alternatives could be 'a teacher's marginal and/or end comments, a teacher's written reaction to a student's self-assessment of performance, a teacher's review of the test in the next class period, peer-assessment of performance, self-assessment of performance, and a teacher's conference with the student' (Brown, 2004; pg. 295). For summative assessment alternatives could be 'a teacher's marginal and/or end of exam/paper/project comments, a teacher's summative written evaluative remarks on a journal, portfolio, or other tangible product, a teacher's written reaction to a student's self-assessment of performance in a course, a completed summative checklist of competencies, with comments, narrative evaluations of general performance on key objectives, a teacher's conference with the student' (Brown, 2004; pg. 295).

As these alternatives suggest, it is crucial for a teacher to keep in mind that the goal is not to simply write the final grade on the paper. Learners deserve information on their

performance, explanations of their strengths and weaknesses and guidance on what they should do to improve.

2.3. ASSESSING YOUNG LEARNERS: EUROPEAN PERSPECTIVE VERSUS BOSNIAN PERSPECTIVE

European countries are known to have carefully crafted vocational certification and education systems for young people. It is not uncommon to hear positive stories from many European countries regarding their education systems and the success with which they operate.

In general, schools in Europe are categorized into five levels: pre-school, primary level (usually grades 1-4 or 5), secondary level I (usually grades 5-9 or 10), secondary level II (grades 10-13), tertiary level (includes universities and other schools of higher education). Starting at the end of secondary level I, nationally standardized exit exams are required to graduate from one level and to enter the next level.

The majority of European countries have certain things in common when it comes to their education systems. First of all, they administer national exit exams at a number of levels and those exams are based on a nationally mandated curriculum. Secondly, at age 16, for those who work toward acceptance at a university or other institutions of higher education, specialization choices exist. At age 18/19, nationally mandated exit exams and the grades received in the major subject areas will determine the chances of being accepted at the school and discipline of his/her choice. Thirdly, curricula are standardized across the individual nations and within the European Community. The educational reforms during the '60s, '70s, and '80s of this century in Europe reflect a greater accessibility to universally recognized career training options, but they also reflect a strict and high standard of general education during the compulsory education years (to age 16). Students, teachers, school administration and cooperating business and industry leaders, as well as the rest of the population, know they have a responsibility in helping achieve the standards (Berger-Proßdorf, n.d.).

Moreover, there are many positive examples coming from a variety of European countries showing their dedication to improving their education systems. In Germany, for example, 'private industry and government are jointly responsible for education and training in fields of pre-professional specialization. The federal government sets the standards for the education, while the respective federal states are responsible for the education in schools of pre-professional specialization. There are about a half-million companies that are qualified for training. After usually three years of a combination of practical training in the workplace and theoretical training in school, the student must take an exit exam. Larger companies often maintain their own pre-professional specialization schools which must be accredited by the government. In turn, the government supports these schools financially. A lot of other countries have similar education systems, but all of them have some kind of a national standard set in coopoeration with business and industry associations' (Berger-Proßdorf, n.d.).

European countries also have a great record of professional development opportunities for teachers. 'In 2006, the Dutch government implemented the Education Professions Act, which dealt with competence standards for all education-related professions. This act required all schools to establish a support program for new staff and to formulate teacher training and induction programs in conjunction with training institutions. The Ministry of Education also furnishes a fund to provide individual training grants to teachers for them to obtain higher or specialist qualifications. This fund also covers the cost of substitute teachers that replace teachers who are in training' (Netherlands: Teacher and Principal Quality', n.d.).

Furthermore, in Finland at most levels of education 'teachers are required to participate in in-service training every year. Finnish teachers consider in-service training to be a privilege and therefore participate actively. In-service training is offered by different providers. The state funds in-service training programmes, primarily in areas important for implementing education policy and reforms. Education providers can also apply for funding to improve the professional competence of their teaching personnel ('Finnish National Agency for Education -Teacher Education', n.d.).

When talking about positive European practices one cannot forget about the Common European Framework of Reference for Languages. The Common European Framework of Reference for Languages (CEFR) is a great example of a guideline created to 'provide a

common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe' ('Use of the CEFR', n.d.). It was created by the Council of Europe as a part of the project called 'Language Learning for European Citizenship' which lasted from 1989 to 1996. In November 2001, a European Union Council Resolution recommended using the CEFR to set up systems of validation of language ability ('Use of the CEFR', n.d.).

Essentially, the CEFR organizes learners into six levels, A1 to C2, which can be regrouped into three broad levels: Basic User, Independent User and Proficient User. These levels can be further divided depending on the needs of the user. All levels are defined through 'can-do' descriptors. These descriptors were created without reference to any specific language, which means they can be applied to any context. The descriptors specify the mastery of each skill which is represented on a six-level scale (A1, A2, B1, B2, C1, C2). However, for an average user these specifications may appear extremely broad, especially since individual languages are not addressed. The Reference Level Descriptions (RLD) for national and regional languages, which provide detailed content specifications for different CEFR levels, have been developed to address this issue. CEFR marked a major turning point as it can be adapted and used for multiple contexts and applied for all languages ('The CEFR Levels', n.d.).

Another amazing invention brought to us by the Council of Europe is the European Language Portfolio (ELP). It can be defined as an instrument based on the CEFR with which individual learners can reflect upon their learning and record their progress. This paper has already touched upon the importance of learner autonomy and fostering such skills from a very young age, as well as mentioned portfolios as an alternative method of assessment. The ELP is definitely one of the best examples of an alternative method of assessment. What the ELP does is support learner autonomy by allowing learners to record their language learning process, progress and experience. It also fosters the development of intercultural competence and awareness, as well as the development of plurilingualism ('What is the ELP?', n.d.).

The ELP consists of three components: a language passport, a language biography, and a dossier. It should be noted that the ELP has also been created for learners of all ages and all languages of the member states of the Council of Europe. This means that there is also a

Junior version of the ELP which was adapted specially for young learners. It consists of the same three components but it was made in a way to help young learners learn to reflect on their progress and achievements when it comes to learning a new language ('What is the ELP?', n.d.).

Furthermore, it also very important to mention that one of the essential components of the ELP is the self-assessment grid which basically illustrates the levels of language proficiency described in the Common European Framework of Reference for Languages. It presents thirty-four scales of listening, reading, spoken interaction, spoken production and writing activities ('Self-assessment grids – CEFR', n.d.).

Autobiography of Intercultural Encounters is another great example of an alternative assessment tool. Autobiography of Intercultural Encounters is designed specifically to develop intercultural competencies and intercultural awareness. It is a tool which helps people think about and learn from specific intercultural encounters, whether that be face to face or through visual media. There are two separate tools that can be used for this purpose: Autobiography of Intercultural Encounters (AIE), and Images of Others: An Autobiography of Intercultural Encounters through Visual Media (AIEVM). Both of these tools also come in two versions. There is a standard version which can be used by older learners and adults, as well as a version for younger learners specifically designed to be used by children who need help from an adult when it comes to reading, writing or even thinking of specific encounters. The focus is on intercultural encounters that have made a strong impression or had a long-lasting effect on the people who use the AIE or AIEVM. In discovering what underlies these encounters, users become more aware of their experience and reflect on their reactions, thereby developing their intercultural competences ('Autobiography of Intercultural Encounters', n.d.).

All of these alternative tools can be adapted to any language or any context. This means that they were designed to be implemented in any classroom with as little issues as possible. Unfortunately, there is no evidence that European Language Portfolio or the Autobiography of Intercultural Encounters have been implemented in any Sarajevo schools, or any Bosnian schools for that matter.

When it comes to Bosnian education system and practices, there are many issues that need to be discussed and addressed. According to the official webpage of the Ministry of Civil Affairs of Bosnia and Herzegovina ('Osnovne informacije o obrazovanju u BiH', 2017), institutional image of the education system in Bosnia and Herzegovina is a reflection of state regulation, defined by the Constitution of Bosnia and Herzegovina, entity and cantonal constitutions, and the Brcko District Statute. The Republic of Srpska, the ten cantons in the Federation of Bosnia and Herzegovina and the Brcko District all have full and undivided jurisdiction in the field of education.

'The Republic of Srpska has a centralized government and one ministry of education. Federation of BiH has a decentralized government and consists of ten cantons where each canton has their own ministry of education. There is also Federal ministry of education, but this ministry has only coordinative role. And Brcko district of BiH has a government with departments. One of those departments is The Department of Education. In accordance with that there are twelve responsible institutions of education in BiH. There are three constituent peoples and three official languages: Bosnian, Croatian and Serbian. They are taught in schools as a mother tongue' ('Key Features of the Education System in Bosnia and Herzegovina', 2017). The grading scale used in Bosnian schools is the one already mentioned in the theoretical part of the paper. Grades are expressed by a number with a one word meaning attached to it. In particular, these grads are: 1 (insufficient/ nedovoljan), 2 (sufficient/ dovoljan), 3 (good/ dobar), 4 (very good/ vrlo dobar), and 5 (excellent/ odličan).

Moreover, education is organized into four basic levels, namely preschool education and care, primary education, secondary education, and higher education. Preschool education and care is intended for children from six months old until they start with primary education. It comprises three levels: nursery (for children from six months to three years old), kindergarten (for children from three years to six years old) and preschool preparatory program. Primary education is compulsory and it lasts nine years. Public primary education is free and the only admission criterion is the age – all children have to be enrolled in primary school between five and a half and six and a half years old. Secondary education is available in forms of general secondary education schools (grammar schools), vocational secondary education schools, art schools and other, less frequent types. General secondary education lasts four

years and vocational secondary education lasts three or four years. Students are usually fifteen years old when they enter secondary education and they are admitted to schools on the basis of primary school achievements and final exam results. Higher education in Bosnia and Herzegovina is organized in three cycles. The first cycle leads to the academic title of completed undergraduate studies (the title of Bachelor) or equivalent, obtained after a minimum of three and a maximum of four years of full time study. Precondition for entering is a certificate of completion of secondary education, and it is evaluated with at least 180 or 240 ECTS credits. Exceptions are integrated studies of medical science group in the first cycle, which are valued up to 360 ECTS credits. The second cycle leads to the academic title of Master or equivalent, obtained after completing undergraduate studies, lasting one or two years, and it is evaluated with 60 or 120 ECTS credits, and the sum with the first cycle is 300 ECTS credits. The third cycle leading to a doctoral degree or equivalent lasts three years and it is valued at 180 ECTS credits ('Key Features of the Education System in Bosnia and Herzegovina', 2017).

Aside from these facts, it is very important to discuss the fact that Bosnian education system has been subject to severe criticism and pleas for reform and change. There are many irregularities in the system that have not been solved even with the various reform efforts made over the course of time. In 2008 the Council of Ministers adopted the Strategic Guidelines for the Development of Education in Bosnia and Herzegovina with the Implementation Plan for 2008 to 2015. The document confirmed that the jurisdiction over education in Bosnia and Herzegovina is decentralized. This way every government has its own right to create education laws, determine curricula and the textbook policy, plan and execute budgets (and thus determine the volume and irregularity/ regularity of the teacher's salary). The consequence of such fragmentation are unequal educational policies, emphasis on the national group of subjects and the absence of any valid external evaluation of student' achievements after completing elementary and secondary school ('Osnovnoškolsko i srednjoškolsko obrazovanje u BiH (trenutno stanje i preporuke za reforme)', 2015).

Another segment of Bosnian education system subject to criticism is teacher education. According to a study prepared under the USAID's Bosnia and Herzegovina Monitoring and Evaluation Support Activity (2016) teacher education is one of the weakest points of the

entire education system. Teacher education is provided by teaching faculties; mainly faculties of philosophy, science and mathematics, as well as art academies. The curricula are designed by the core department, and the faculty determines curricular content. Due to this reason the curriculum varies from faculty to faculty. Teaching methods employed in subjects of pedagogy, psychology, didactics and methodology are outdated. There is a severe lack of practical classes and internships and the majority of teacher trainees only get to experience peer teaching and a single independent class held at a primary or secondary school. The majority of students complain of not feeling competent or comfortable enough to teach after they finish their studies. The continuing professional education of teachers is also at an unsatisfactory level. A number of severe budget cuts have been implemented at various governmental levels over the past decade, with professional education being hit particularly hard (Branković, Husremović, Zečević, and Vukotić, 2016).

If we focus exclusively on the assessment segment, the Framework Curriculum (Model of the Framework Curriculum for BiH - Curriculum, November 2005, a project funded by the EU) states that the purpose of assessment should be a meaningful contribution to the progress of the student where the end goal cannot be a rating of knowledge and recommends that besides traditional assessment methods (oral examination, written work, tests) teachers also use alternative ways of assessing students. In 2014, the board of the Agency for Preschool, Primary and Secondary Education has adopted the common core curriculum for foreign languages defined on learning outcomes as one of the steps to reform the education system in Bosnia and Herzegovina ('Osnovno obrazovanje u Bosni i Hercegovini- kvalitet, kreativnost i inovativnost?', 2010). The common core curriculum for foreign languages defined on learning outcomes was designed according to the already established methodology of development of the common core curriculum; meaning, the document defines all areas that are comprised of components, learning outcomes are defined for each component, and descriptors with an emphasis on special features of learning foreign languages are defined for each learning outcome. The starting point for the development of the common core curriculum for foreign languages defined on learning outcomes are the Common European Framework of Reference for Languages, Identification of Key Competences and Life Skills in BiH (2011), and the results of the analysis of the curricula of Slovenia, Finland, Singapore, Canada and Australia.

Levels of language competence were designed in accordance with the Common European Framework of Reference for Languages which organizes learners into six levels, A1 to C2, which can be regrouped into three broad levels: Basic User, Independent User and Proficient User. The goal of the common core curriculum for foreign languages defined on learning outcomes is to improve the existing curricula for foreign languages in BiH, to improve the quality of educational practice in teaching foreign languages, and to ensure compatibility of documents with EU education systems and language standards specified in the Common European Framework of Reference for Languages ('Zajednička jezgra nastavnih planova i programa za strane jezike definisana na ishodima učenja', 2014).

Although there have been many attempts in reforming Bosnian education system and improving education practices, there is still much work to be done. Unfortunately, there is no evidence that European Language Portfolio or the Autobiography of Intercultural Encounters have been implemented in any Sarajevo schools, or any Bosnian schools for that matter. There are also very few opportunities for teachers to continue their professional development, and there are very few attempts from the government or private business to finance teachers' professional development. European practices can and should be taken as a great example and a guide for future reforms in Bosnian education system.

3. RESEARCH METHODOLOGY

The subject of the research is evaluation and assessment of language competences in Bosnian schools.

The problem of the research is effectiveness of methods and techniques used for assessment of language competencies in Bosnian primary schools.

The aim of the research is to determine which types and methods of evaluation and assessment are used by teachers in Bosnian primary schools, how effective these methods are, and to what extent the current education system in Bosnia and Herzegovina allows and supports the use of standard and alternative ways of evaluating and assessing the development of language competences.

Hypothesis:

Newly acquired knowledge of alternative ways of evaluating language competences suggests that traditional methods of evaluation are not an effective indicator of language competence.

3.1. RESEARCH METHODS

The descriptive method, the comparative method and the compilation method were used in the theoretical as well as practical part of the paper, while the method of proof was used in the discussion.

3.2. TECHNIQUES AND PROCEDURES

The technique of analyzing professional and scientific literature with the aim of defining key concepts relevant to understanding the problem was used in the theoretical part of the paper. The technique of anonymous online surveys was used in the research part of the paper with the intention of determining the standard and alternative ways of assessing and evaluating language competences at the level of Sarajevo primary schools.

3.3.RESEARCH INSTRUMENTS

Anonymous online questionnaires were used to gather data for research purposes. Questions comprising the questionnaires were adapted from a questionnaire used for diploma thesis research conducted in Masryk University in Czech Republic by Lenka Pribilova (2008), and modified for the needs of this research. Pribilova's questionnaire was written and conducted in Check language; however, it was translated to English language in the diploma thesis. It consisted of twelve questions, both open-ended and closed-ended. Upon analyzing the original questionnaire, it was decided to add questions regarding teacher's age, degree and employment status, as well as questions regarding alternative assessment in particular (as it had not been tackled at all in the original questionnaire). After modifying the original questionnaire, there were twenty-one questions in total: seventeen open-ended questions and four closed-ended questions. The questionnaire was written and conducted in English language. The questionnaire is included in the appendix.

3.4. RESEARCH SAMPLE

The research was conducted at the level of primary schools of canton Sarajevo. There are seventy-five primary schools in canton Sarajevo ('Osnovne škole u Kantonu Sarajevo', n.d.). The research did not cover all of the primary schools in canton Sarajevo. Moreover, it is impossible to state which schools in particular were a part of the research due to limitations encountered while conducting the research. Planned number of respondents was ten, including ten English language teachers from ten different primary schools in canton Sarajevo. By the end of the research the total number of respondents was nineteen. All respondents are English language teachers in primary schools in Sarajevo; however, some of them are currently unemployed (but were employed at a primary school in Sarajevo at one point). The respondents did not specify which primary school they are/were employed in. The age distribution of the research was 23 (twenty-three) to 37 (thirty-seven).

3.5. GATHERING INFORMATION

The process of gathering information for this research was challenging. The initial plan was to conduct an online survey among primary school English language teachers in canton Sarajevo. Since there are seventy-five primary schools in the canton, the main idea was to contact teachers currently employed at said schools and send them the questionnaire via email. This approach was too optimistic as it was impossible to get hold of teacher's emails. Schools that have been contacted in the attempt of obtaining teacher's emails declined to share that information.

Since the initial plan failed, the questionnaire was posted on various sites and Facebook groups used by teachers from Bosnia and Herzegovina exclusively. This resulted in only nine responses from teachers.

In an effort to obtain more information, the questionnaire was also sent to teachers that are not employed in primary schools in canton Sarajevo with a request to forward it to their colleagues who are employed there. The questionnaire was also sent to my colleagues with the same request. This effort resulted in ten more responses, which makes nineteen responses in total.

3.6. RESEARCH LIMITATIONS

Starting from the very creation of the questionnaire, it was challenging to modify the original questionnaire for the needs of this particular research. The majority of questions from the original questionnaire were kept; however, some additional questions regarding respondents' age and professional education, as well as respondents' knowledge and application of alternatives in assessment were added. It is important to add that in the process some questions were not properly worded which resulted in ambiguous and incomplete answers from respondents.

Moving on further in the research, the initial plan was to reach and question as many teachers from Sarajevo primary schools as possible. However, as the research progressed this plan turned out to be quite ambitious. The number of respondents was disappointingly low. Since there are seventy-five primary schools in the canton, this number cannot be seen as a true representative of all assessment practices in Sarajevo primary schools.

Lastly, since the questionnaire was anonymous and respondents were free to answer the question whichever way they like, some of them avoided answering questions regarding their employment status, information about the school they were currently employed in, as well as questions regarding their professional education. Lack of such information resulted in unclear respondents' profiles which do, in a way, affect the final analysis of the research results.

3.7.RESEARCH RESULTS

There were twenty-one questions in total: seventeen open-ended questions and four closed-ended questions.

OPEN ENDED QUESTIONS:

1. *What is your age?*

Although the questionnaire is anonymous, it is still important to gain better insight of the respondent's background and gather information that helps paint a picture of what respondents are like. The age distribution of this survey is twenty-three (23) to thirty-seven (37). Adults between the ages of twenty three (23) and thirty (30) are over-represented while adults over thirty-seven (37) are under-represented or not represented at all.

2. *What is the highest degree or level of school you have completed? If currently enrolled, highest degree received.*

Two respondents have completed Bachelor's degree, while the remaining seventeen respondents have completed Master's Degree. One respondent specified that they received their degree in English language and linguistics.

3. *What is your employment status? (employed for wages, self-employed, unemployed...) If possible, please name the school you are employed for.*

Thirteen respondents stated they are currently employed, while six respondents stated they are currently unemployed. None of the employed respondents specified which school they were employed in.

4. *How long have you been employed as an English language teacher?*

Respondents have been asked to answer this question whether or not they are currently employed which means that the numbers represented in the following passage might refer to their current employment or their past employment. One respondent has been employed for nine months. Five respondents have been employed for one year. Three respondents have been employed for two years. One respondent has been employed for three years. One respondent has been employed for four years. One respondent has been employed for six years. Two respondents have been employed for ten years. In summary, respondents' employment time ranges from nine months to ten years.

5. *Why did you decide to become an English language teacher?*

Although all answers to this question are different and very subjective, some patterns can be drawn. All of the answers can be divided into three groups: a) love of teaching, b) love of English language, and c) bad experience with past teachers.

The majority of respondents (seventeen) said they always wanted to be teachers and their love of languages and English language in particular drew them to this profession. Some respondents (eleven) also added that they always wanted to work with children and young learners and be able to transfer their knowledge to them.

Two respondents said they had bad teachers growing up and decided to become teachers themselves to help change that issue. Lastly, one respondent said they never wanted to become a teacher.

6. *What do you normally assess (participation, effort, knowledge, independence…)?*

Seven respondents said they assess knowledge and participation. Three respondents said they only assess knowledge. Eleven respondents said they assess everything listed in the question and more.

It should also be noted that some respondents added that they try to assess everything but they put the emphasis on knowledge.

7. How often do you assess your learners and why?

Respondents gave very subjective answers to this question which makes them difficult to sum up. Assessment incidence approximately varies from every day to every month.

Three respondents stated they assess their students every class. Two of them added that their reason for such frequent assessment is to keep track of students' progress and achievements.

Four respondents stated they assess their students every week or every two weeks. Their reasoning behind that decision is that they want to track their students' progress and give them feedback, but they do not see a point in assessing them more frequently.

Five respondents stated they assess their students once a month. Those who stated their reasoning behind that decision said they only assess their students when they need to grade them. One respondent also added that assessing once a month is ideal, anything beyond that would be too much.

The rest of the respondents (seven) stated they assess their students 'as often as they can', 'as often a needed', 'very often', or 'as often as possible'. The majority of them said they assess their students' knowledge with the intention of tracking their progress and keeping score of their achievements. On the other hand, some said they assess them as often as every day but they do not grade them that often.

8. When assessing learners via tests do you only assess a certain category or do you assess everything together?

Ten respondents stated they assess everything together. The majority of them explained their reasoning behind that decision in saying that it is easier to design tests that include all categories (vocabulary, grammar, comprehension, reading etc.). One respondent said that they see language as a whole so they assess it in the same manner.

One respondent said they assess exclusively by category. Their reasoning behind it is that such assessment is easier for both sides: teachers and students.

Eight respondents stated they use both approaches, depending on various factors. Some respondents said they assess by category if they were working on a certain category beforehand (vocabulary, grammar etc.) and they only want assess said category. Some

respondents said that they assess by category throughout the semester, and they assess all categories together at the very end of semester (to assess overall knowledge). One respondent said the prefer assessing by category because it is easier for students to prepare; however, they still assess all categories together because the majority of written tests are designed in that way.

9. *How do you test your learners most often: oral exam or written tests, and why?*

Six respondents stated they most often assess their students via written tests. According to them, written tests are much easier to organize and evaluate. Two respondents added that written tests require less time and are more enjoyable for the students. One respondent added that there is not always enough time to conduct oral tests.

Thirteen respondents said they employ both approaches. The majority of respondents said they prefer written tests because they require less time and are efficient, but they still conduct oral tests as well. Some respondents added that it is important to employ a variety of assessment techniques as to cater to everyone's differences and preferences, as well as to have insight into different types of student responses.

It is important to point out that none of respondents opted for oral tests as a main technique of assessment.

10. *How do you normally assess listening skills?*

All respondents stated they assess listening skills by letting students listen to audio or video material, take notes and try to understand as much of the audio as possible. After that, they assess their listening skills by asking them comprehension questions, or by asking them to retell the story from the audio material, or by asking them to have conversations among themselves about the story from the audio material. One respondent said they even ask students to reenact the story from the audio if possible, in order to make assessment more interesting and fun.

11. How do you normally assess speaking skills?

The majority of respondents stated they assess speaking skills through conversations, presentations, debates or oral tests. Two respondents mentioned that they prefer working in groups or pairs and initiating authentic conversations and use them as subject of assessment.

12. How do you normally assess writing skills?

The majority of respondents stated they assess writing skills through essays and written assignments including writing stories, poems, letters and emails. Some respondents also added that their reasoning behind choosing these methods is to get students to express themselves creatively.

13. How do you normally assess reading skills?

The majority of respondents stated they assess reading skills by assigning different texts to students (typically texts from the textbook) and having them read those texts out loud. A couple of respondents said they like to pick reading texts themselves, depending on the topic of interest. Some respondents added that they try to get more creative, especially with young learners, and do various reading activities they find online.

14. What is your opinion on the grading system?

Six respondents stated that they think the grading system is efficient as it is. They also added that it is easy to use and if you know how to use it there should not be any issues with it.

Eight respondents state that they think the grading system needs improvement. Their reasoning behind such opinion is that the current grading system lacks depth and meaningful explanations of the grades, grades very often do not reflect true knowledge, knowledge tends to be reduced solely to grades (and everything seems to be about getting the right grade rather than retaining the knowledge). The majority of them agree that the current grading system would be efficient if grades were accompanied with meaningful explanations and advice.

Five respondents said they either never thought about it or simply have no opinion on the grading system whatsoever.

15. What is the parents' and the learners' opinion on the grading system?

One respondent stated that they think both parents and learners are satisfied with the grading system.

Eleven respondents stated that they think both parents and learners are not satisfied with the grading system in a way that they often get frustrated with grades. Parents and learners sometimes feel like the grade is not reflective of the learner's knowledge. Some respondents added that a grade not being reflective of learner's knowledge is not system's fault and it is not something that can be changed just by changing the system. Some respondents suggested using less formal and less traditional ways of assessment in combating this issue.

Seven respondents said they never thought about this matter before, or they do not have an opinion on it.

16. Do you think it is necessary to combine various types of assessment?

All respondents overwhelmingly agreed that it is necessary to combine various types of assessment. However, some respondents added that very often it is not possible to do that in the classroom either because of lack of time or lack of financial support. This is the main reason why so many teachers opt for easier means of assessment.

17. Do you ever let the learners assess themselves or assess you?

Four respondents stated they let learners assess themselves and they let them assess them (the teacher). Three added that this is important because it helps improve the relationship between the teacher and learners as well as between learners themselves. One respondent added that it also helps learners develop their critical thinking skills.

Ten respondents stated they ask learners to assess them (the teacher) in form of feedback. They explained that they typically ask for feedback either at the end of every lesson, or whenever they have some time left. One respondent said that they do not ask learners to assess themselves because they think it is not appropriate to require that from young children.

Five respondents said they do not let learners assess themselves or assess them (the teacher).

CLOSED ENDED QUESTIONS:

1. *Which types of assessment do you recognize?*

Respondents were expected to choose from eight types of assessment: informal assessment, formal assessment, formative assessment, summative assessment, norm-referenced assessment, criterion-referenced assessment, performance-based assessment, and alternative assessment.

Figure1
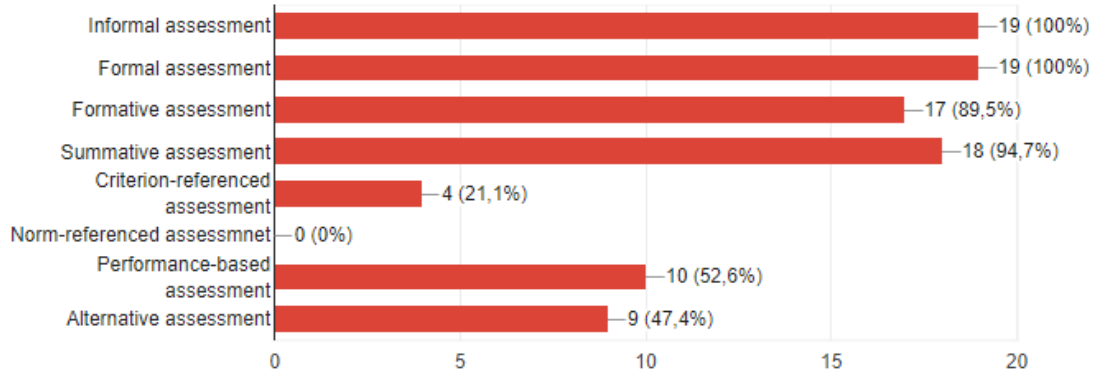Graphical representation of respondents' answers
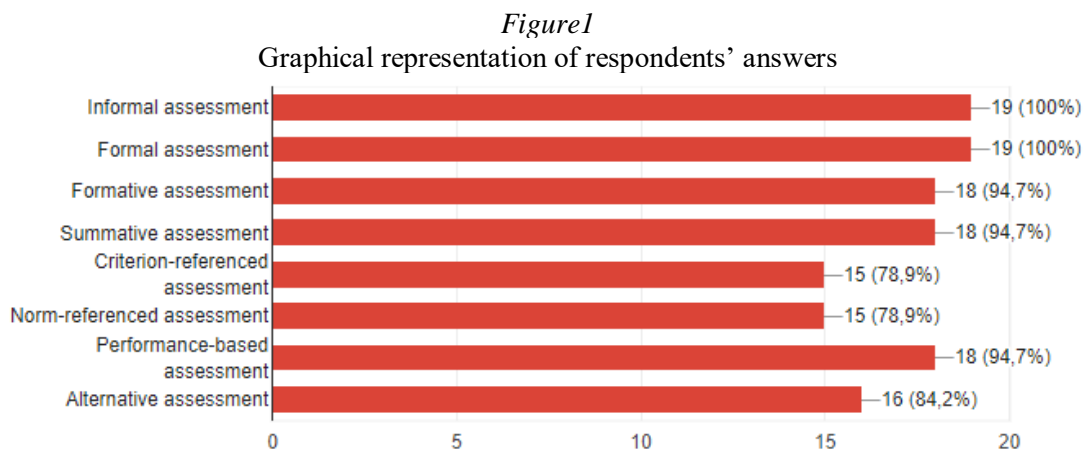
Nineteen respondents recognize informal and formal assessment. Eighteen respondents recognize formative, summative and performance-based assessment. Sixteen respondents recognize alternative assessment. Fifteen respondents recognize criterion-referenced and norm-referenced assessment.

2. *Which types of assessment do you use in your classroom?*

Respondents were expected to choose from eight types of assessment: informal assessment, formal assessment, formative assessment, summative assessment, norm-referenced assessment, criterion-referenced assessment, performance-based assessment, and alternative assessment.

Figure 2
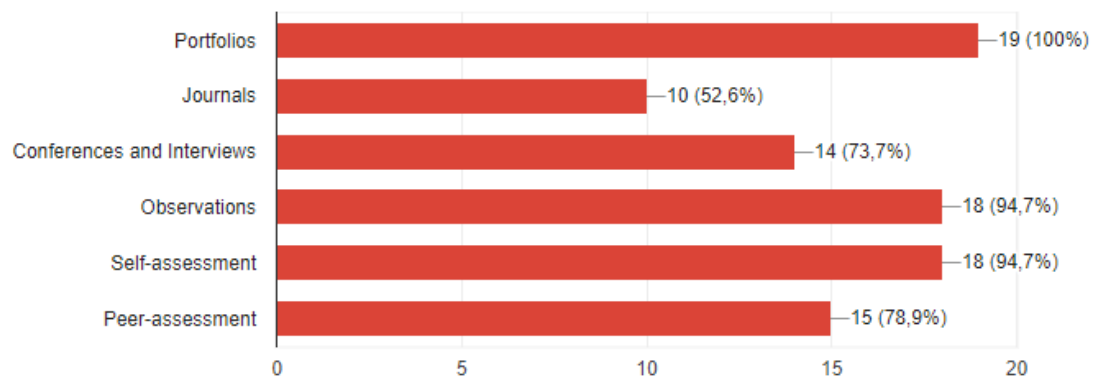
Graphical representation of respondents' answers

Nineteen respondents employ informal and formal assessment. Eighteen respondents employ summative assessment. Seventeen respondents employ formative assessment. Ten respondents employ performance-based assessment. Nine respondents employ alternative assessment. Four respondents employ Criterion-referenced assessment. Norm-referenced assessment is not employed by any of the respondents.

3. Which types of alternative assessment do you recognize?

Respondents were expected to choose from six different types of alternative assessment: portfolios, journals, conferences and interviews, observations, self-assessment, and peer-assessment.

Figure 3

Graphical representation of respondents' answers

Nineteen respondents recognized portfolios. Eighteen respondents recognized observations and self-assessment. Fifteen respondents recognized peer-assessment. Fourteen respondents recognized conferences and interviews. Ten respondents recognized journals.

4. Which types of alternative assessment do you use in your classroom?

Respondents were expected to choose from six different types of alternative assessment: portfolios, journals, conferences and interviews, observations, self-assessment, and peer-assessment.

Figure 4

Graphical representation of respondents' answers
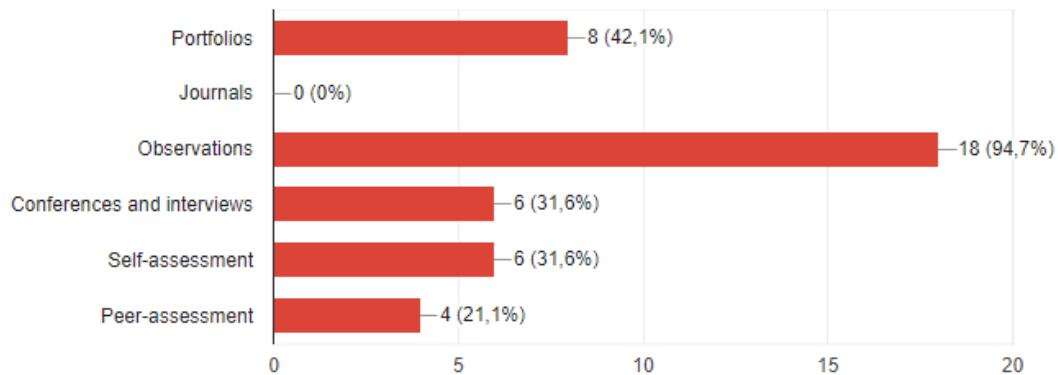
Eighteen respondents employ observations in their classrooms. Eight respondents employ portfolios. Six respondents employ conference and interviews and self-assessment. Four respondents employ peer-assessment. Journals are employed by none of the respondents.

3.8. DISCUSSION

The survey was conducted with the end goal of obtaining information on which methods and techniques of assessment are used in Sarajevo primary schools, and whether or not those methods and techniques are considered efficient by teachers in said schools.

The general conclusion of the survey is that teachers in Sarajevo primary schools are not employing all of the methods and techniques of assessment, whether that is because they are not financially able to support those methods or because they are not motivated enough to step out of their comfort zone and go beyond the traditional assessment route. It is important to note that this survey was conducted with limited resources and only for the purposes of this research. It cannot be taken as the ultimate indicator of assessment in Sarajevo schools. That being said, there are certain conclusions to be drawn as well as certain inconsistencies to be discussed.

 Starting from the very beginning of the questionnaire, it is apparent that the accessibility of the questionnaire could have potentially played an important role in how authentic the results ended up being. The survey was conducted online, with social media playing a major role in its accessibility. Although the popularity of social media has given rise in the last couple of years, and the majority of people do have access to internet, it cannot be claimed with certainty that every English language teacher in Sarajevo had access to this questionnaire.

When asked about their choice of becoming a language teacher, the majority of respondents said they have always wanted to become teachers and to have the opportunity to work with children. However, although a lot of teachers mentioned how they wanted to be able to change the system, and work on educating children in the best way possible, a considerable number of them still stated that they do not employ a variety of assessment techniques and only focus on traditional ones. Financial issues seem to be on top of the list of reasons why Sarajevo teachers choose traditional methods of assessment over alternative ones. According to the survey, lack of financial means prevents them from attending seminars and other types of classes aimed at teacher education. It also prevents them from employing methods and techniques which do require money. However, there are so many different

methods and techniques, both traditional and alternative, discussed in a variety of sources (McKay, 2006; Brown, 2004; Brown, 2005; Brown 2007; Harmer, 2005; Cameron, 2010) which do not require excessive amounts of money and could be successfully employed in any classroom. Moreover, the majority of teachers stated that some of the main reasons behind pursuing a career in education are their love of teaching, their love of children as well as their wish to improve and change the world one child at a time. This implies that their main source of motivation is their personal satisfaction gained through teaching. However, the majority of teachers complained about lacking motivation to employ alternatives in assessment. Teacher burnout[2] could potentially explain lack of motivation mentioned in the majority of teachers' responses. Many teachers over time experience lack of motivation due to stress and inability of coping with the pressures of their job.

Another inconsistency in their answers can also be seen when respondents were asked to clarify what they assess most in the classroom (knowledge, participation, independence, effort, etc…). The majority of respondents stated they assess everything that goes on in the classroom; however, the later part of the questionnaire dealing with particular methods of assessment indicated that respondents for the most part only use what we deem as traditional methods of assessment (written and oral testing). Motivation, independence, effort or cooperation would be very difficult/ impossible to assess through written or oral testing.

When asked how often they assess their students, respondents had quite mixed responses. The problem that might have occurred with this question is that it seemed as if a lot of teachers were confused by the terms 'grading' and 'assessing' and that might have influenced the final results of the question. Reasoning behind that conclusion is that a lot of teachers stated things such as 'they assess their students whenever they need to grade them' or 'they try not to assess their students too often'. Since assessment is a continuous process of tracking students' achievements and progress, and a grade is just a result of that process, it seems fair to assume that respondents might not have had proper understanding of those terms.

Questions relating to particular methods and techniques used to assess reading, speaking, listening or writing were meant to reveal what is commonly used in Sarajevo classrooms.

---

[2] According to Campbell (1983), teacher burnout could be defined as an overwhelming state of exhaustion (emotional or physical) caused by many issues teachers face such as oversized classes, inadequate salaries, excessive paperwork or even problems with parents, colleagues and administration.

Since the majority of teachers had very similar responses, it can be said that students in Sarajevo schools have a unified experience of assessment. Although all methods and techniques mentioned by the respondents are proven as efficient, lack of creativity is evident and worrying. Nowadays, access to different types of activities and ideas for authentic assessment is easy. Teachers can find inspiration online, whether that is on different sites targeted for teachers specifically, or by connecting with other teachers over social media and sharing ideas. Why is this the case, then? Assessment is usually not a priority in the classroom. A huge number of studies conducted in a variety of countries across Europe indicate that language teachers are either not skilled in language assessment, or do not see assessment as a valid focal point of language teaching (Edelenbos and Johnstone, 1996; Nikolov, 2000; Hill, 2000; Low, Brown, Johnstone and Pirrie, 1995; according to McKay, 2006). This implies that Sarajevo teachers might be putting an effort into delivering their lessons in a creative and authentic way, but not putting an effort into making their assessment practices just as authentic. Moreover, factors like teacher burnout, lack of motivation and lack of financial means mentioned in their responses could also be a reason for relying on more traditional methods of assessment.

Two questions relating to teachers', parents' and students' opinion on the grading system were quite poorly worded. These questions were originally written in Check language and translated to English language (by the author of the diploma thesis, Lenka Pribilova) which might have caused the issue with wording. The aim of the questions was to see if teachers agreed with the idea that grades should express and stand for much more than a single word explaining their meaning. However, it was very difficult to form a question without leading them on to an answer or an opinion. A lot of teachers did not even understand the question, or did not have an answer as they never thought about it before. Therefore, the aim of the question was not fulfilled.

It is also quite disheartening that when asked whether a variety of assessment methods and techniques should be employed in the classroom, the vast majority of teachers agreed but said that doing so is almost impossible in real life classrooms. Again, lack of time, money and energy is their main reasoning behind such claims. This is confirmed in the latter part of the questionnaire when teachers were asked which methods and techniques they employ in their

classrooms. A lower number of teachers use performance-based and alternative assessment, and the alternative assessment methods that they do use are those that naturally require less money and energy (observations).

Self-assessment and peer-assessment do not require money or energy but are still neglected as methods of assessment in Sarajevo schools. When it comes to self-assessment, the majority of teachers do not let students assess their own progress and their own work. The main reason why this is the case could be because there is a common misconception that young learners are not capable of assessing themselves. However, fostering learner autonomy has been proven to be immensely beneficial for the learner. It is very important to introduce young learners to the idea of evaluating their own work, and discussing their own strengths and weaknesses. Although they might not be able to truly assess their work at a young age, they would still be developing these skills through these discussions. Moreover, teachers very often dread not being in control of everything that goes on in the classroom (Campbell, 1983). When teachers let students assess themselves or assess their peers, they risk losing control as they cannot predict what these discussions might bring up. On the other hand, seeking feedback on their own work is quite common for the majority of respondents. They seem to value their students' opinions and have enough motivation and drive to incorporate their ideas into the classroom.

Lastly, the aim of the set of closed-ended questions was to obtain information on which methods of assessment teachers recognize versus which methods of assessment teachers use. These questions were formed as closed-ended questions in order to draw clear and visible comparisons between the two aspects. Since there is constant progress made in the field of assessment, it is possible to assume that some teachers might not be keeping up with the newest methods. It is important to clarify whether some teachers might not be using certain methods simply because they are not familiar with them. As it turns out, teachers do recognize the majority of assessment methods they were presented with, but they do not use all of them. Therefore, they are familiar with all of methods of assessment at the level of recognition but they are not familiar with them at the level of application. What that means is that they might be familiar with these methods because they learned about them years ago in college, or

because they simply heard about them, but they are not familiar enough with them to actually apply them in their own classrooms.

Lastly, the hypothesis stating that *'newly acquired knowledge of alternative ways of evaluating language competences suggests that traditional methods of evaluation are not an effective indicator of language competence'* can be confirmed. Both theoretical and practical part of the research suggest that traditional assessment methods, although efficient, are not an authentic indicator of language competence when used on their own.

4. CONCLUSION

Language assessment is something that seems to have been neglected seen as less important in the language learning process. Teachers are either not motivated enough to explore the vast field of language assessment, or simply do not see it as something that can have any effect on the language learning process. However, the effects of proper assessment practices are immense and one of the major claims that this paper has attempted to illustrate is that 'assessment has the power to change people's lives' (Shohamy, 2001; according to McKay, 2006; pg. 18).

Assessment can be immensely beneficial for everyone involved in the language learning/teaching process; however, it can also be harmful if done incorrectly. There are so many different methods and techniques of language assessment that have been used and proven efficient throughout history, as well as so many of those more experimental and alternative ways of assessment that have given rise in recent years. For a teacher in today's day and age, it would be inexplicable to only focus on using one or two most convenient methods of assessment. Unfortunately, such instances are still the reality of education systems in many countries, not only in Bosnia and Herzegovina.

That being said, the aim of the research was to determine which types and methods of evaluation and assessment are used by teachers in Bosnian schools, how effective these methods are, and to what extent the current education system in Bosnia and Herzegovina allows and supports the use of traditional and alternative ways of evaluating and assessing the development of language competences.

The results of the survey revealed that teachers in Sarajevo primary schools mainly use more traditional assessment methods consisting of written and oral testing. However, the results also revealed that the majority of teachers agree that these assessment methods should be combined with more authentic and alternative ones. There is a will among teachers to improve their assessment practices; however, the survey implies that the current education system does not truly allow that to happen. They are lacking knowledge, time, energy, finances and motivation.

Although there have been efforts made in the path towards reforming the education system in Bosnia, there is still room for improvement. Firstly, it is still very difficult to determine

with certainty whether teachers in Bosnia are following the common core curriculum for foreign languages defined on learning outcomes created by the board of the Agency for Preschool, Primary and Secondary Education. If not, that could be seen as the starting point for improvements in assessment.

Furthermore, teachers should be given more freedom when it comes to assessment because very often parents and the administration require tangible results of assessment (a grade). What this does is simply reduces learner's knowledge to a single number, as well as reduce teacher's possibilities in creating more authentic assessment practices. However, this issue could also be solved by teachers focusing their attention on taking proper steps in employing achievement standards suggested by APOSO into their classroom assessment. If teachers are using these standards as well as 'can-do' descriptors they will attach meaning to their grades which will help learners understand their level of knowledge better.

Moreover, teachers in training should be given more practice hours and possibilities for internships while attending college. There needs to be an increase in cooperation between universities and schools which are willing to offer internships for free. There needs to be a way in which teachers in training gain valuable experience which would also help them when looking for employment afterwards.

Lastly, teachers should be required to educate themselves as much as possible through seminars and projects in order to keep up with the newest trends in language teaching and education in general. One of ways in which this could be achieved would be an increase in interest from businesses and industries to sponsor seminars and educational programs for teachers according to the guidelines from APOSO (as it has been seen in examples from Finland and the Netherlands).

5. BIBLIOGRAPHY

Branković, N., Husremović, Dž., Zečevoić, I. and Čaršimamović Vukotić, N. (2016). *Brief Assessment of Basic Education in Bosnia and Herzegovina.* Bosnia and Herzegovina: USAID

Berger-Proßdorf, T. (n.d.). *Educational Systems in Europe*. Montana State: MSU-Billings.

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. White Plains: Pearson Education.

Brown, H. D. (2007). *Teaching by Principles. An Interactive Approach to Language Pedagogy*, Third Edition. Addison Wesley Longman.

Brown, J. D. (2005). *Testing in Language Programs: A compehensive guide to English language assessment*. New York: McGraw-Hill.

Brumfit, Christopher, Moon, Jayne and Ray Tongue. (1995). *Teaching English to Children: From Practice to Principle.* London: Longman Group UK Limited.

Byram, M, Barrett, MD, Ipgrave, J, Jackson, R and Mendez-Garcia, MC. (2009). *Autobiography of Intercultural Encounters.* France: Language Policy Division, Council of Europe Publishing.

Cameron, L. (2010). *Teaching Languages to Young Language Learners*. Cambridge University Press.

Campbell, L. (1983). Teacher Burnout: Description and Prescription. *The Clearing House* (pp. 111-113). Oxfordshire: Taylor & Francis, Ltd.

Dragunić, L. (2013). *Kvalitet i troškovi obrazovanja u BiH.* Sarajevo: Fond otvoreno društvo Bosne i Hercegovine.

Ellis, R. (1997). *Second Language Acquisition* (Oxford Introductions to Language Study). USA: Oxford University Press.

Furlan, I. (1996). *Upoznavanje, ispitivanje i ocjenjivanje učenika*. Zagreb: Pedagoško–književni zbor.

Halliwell, S. (1994). *Teaching English in the primary Classroom.* London: Longman Group UK Limited.

Harmer, J. (2005). *The Practice of English Language Teaching.* London: Longman   Group UK.

Ibrahimović, N. (2015). *Osnovnoškolsko i srednjoškolsko obrazovanje u BiH (trenutno stanje i preporuke za reforme).* Sarajevo: Inicijativa za monitoring evropskih integracija Bosne i Hercegovine.

Ioanou-Georgiou, Sophie, and Pavlou, Pavlos. (2003). *Assessing young learners, Resource books for Teachers.* Oxford: Oxford University Press.

Kyriacou, C. (1998). *Temeljna nastavna umijeća.* Zagreb: Educa.

McKay, P. (2006). *Assessing Young Language Learners*. Cambridge University Press.

Naletilić, M. (2014). *Zajednička jezgra nastavnih planova i programa za strane jezike definisana na ishodima učenja.* Mostar: Agencija za predškolsko, osnovno i srednje obrazovanje.

O'Malley, J.M., & Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. White Plains, NY: Addison-Wesley.

Osnovno obrazovanje u Bosni i Hercegovini- kvalitet, kreativnost i inovativnost?. (2003). Sarajevo: Centar za Politike i Upravljanje.

Owen-Jackson, Gwyneth. (2010). *Teacher education in Bosnia and Herzegovina: developments in the post-war, post-communist context.* In: ECER 2010 Education and Cultural Change, 25-27 Aug 2010, Helsinki, Finland.

Palomba, Catherine A., and Banta, Trudy W. (1999). *Assessment Essentials: Planning, Implementing, Improving*. California: Jossey-Bass Inc. Publishers.

Popham, J. (2003). *What Every Teacher Should Know About Educational Assessment*, Los Angeles: Pearson Education.

Pribilova, L (2008). *Assessing Young Learners*. Published diploma thesis. Brno: Department of English language and literature of the Faculty of Education.

Rixon, S. (1999). *Young Learners of English: Some Research Perspectives*. London: Longman.

Slatina, M. (1998). *Nastavni Metod*. Sarajevo: Filozofski fakultet.

Ur, P. (2005). *A Course in Language Teaching: Practice and theory*. Cambridge: Cambridge University Press.

Vizek Vidović, V., Rijavec,M., Vlahović-Štetić, V., Miljković, D. (2003). *Psihologija obrazovanja*. Zagreb: IEP – Vern.

WEBPAGES:

Teacher Education. (n.d.). Retreived from https://www.oph.fi/english/education_system/teacher_education

Illustrations of the European levels of language proficiency. (n.d.). Retrieved from https://www.coe.int/en/web/common-european-framework-reference-languages/illustrations-of-levels

European Langauge Portfolio. (n.d.) Retreived from https://www.coe.int/en/web/portfolio

Netherlands: Teacher and Principal Quality. (n.d.) Retreived from http://ncee.org/what-we-do/center-on-international-education-benchmarking/top-performing-countries/netherlands-overview/netherlands-teacher-and-principal-quality/

The CEFR Levels. (n.d.) Retreived from https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions

Mršić, D. (2017, May 20). *Key Features of the Education System in Bosnia and Herzegovina.* Retreived from https://ec.europa.eu/epale/

Osnovne informacije o obrazovanju u BiH. (2017, October 10). Retrieved from www.mcp.gov.ba/org_jedinice/sektor_obrazovanje/

APPENDIX

QUESTIONNAIRE

Set of open-ended questions:

1. What is your age?
2. What is the highest degree or level of school you have completed? *If currently enrolled, highest degree received.*
3. What is your employment status? *(employed for wages, self-employed, unemployed…)* *If possible, please name the school you are employed for.*
4. How long have you been employed as an English language teacher?
5. Why did you decide to become an English language teacher?
6. What do you normally assess (participation, effort, knowledge, independence…)?
7. How often do you assess your learners and why?
8. When assessing learners via tests do you only assess a certain category or do you assess everything together?
9. How do you test your learners most often: oral exam or written tests, and why?
10. How do you normally assess listening skills?
11. How do you normally assess speaking skills?
12. How do you normally assess writing skills?
13. How do you normally assess reading skills?
14. What is your opinion on the grading system?
15. What is the parents' and the learners' opinion on the grading system?
16. Do you think it is necessary to combine various types of assessment?
17. Do you let the learners assess themselves or assess you?

Set of closed-ended questions:

NOTE: Definitions of each type of assessment were mentioned in the questionnaire.

•Informal assessment involves observing the learners as they learn and evaluating them from the data gathered.

•Formal assessment involves evaluating a learner's level of language in a formal way, such as through an exam or structured continuous assessment.

•The goal of formative assessment is to monitor student learning to provide ongoing feedback that can be used by instructors to improve their teaching and by students to improve their learning.

•The goal of summative assessment is to evaluate student learning at the end of an instructional unit by comparing it against some standard or benchmark.

•Criterion-referenced assessment measures a student's performance based on mastery of a specific set of skills. It measures what the student knows and doesn't know at the time of assessment. The student's performance is NOT compared to other students' performance on the same assessment.

•Norm-referenced assessment measures a student's performance in comparison to the performance of same-age students on the same assessment. Normative scoring is based on a bell curve, meaning only half of those tested can score above the 50th percentile.

•The idea with performance-based testing is to require a demonstration of the scope of knowledge a student has on a subject rather than simply testing the accuracy of their response on a selection of questions.

•Alternative assessment is any classroom assessment practice that focuses on continuous individual student progress. Perhaps the best way to define alternative assessment is to say that it is the counter to traditional forms of standardized assessment

1. Which types of assessment do you recognize?

   a) Informal assessment

   b) Formal assessment

   c) Formative assessment

   d) Summative assessment

   e) Norm-referenced assessment

   f) Criterion-referenced assessment

   g) Performance-based assessment

   h) Alternative assessment

2. Which types of assessment do you use in your classroom?

   i) Informal assessment

   j) Formal assessment

   k) Formative assessment

   l) Summative assessment

   m) Norm-referenced assessment

   n) Criterion-referenced assessment

   o) Performance-based assessment

   p) Alternative assessment

3. Which types of alternative assessment do you recognize?

   a) Portfolios

   b) Journals

   c) Conferences and interviews

   d) Observations

   e) Self-assessment

   f) Peer-assessment

4. Which types of alternative assessment do you use in your classroom?

   g) Portfolios

   h) Journals

   i) Conferences and interviews

   j) Observations

k) Self-assessment

l) Peer-assessment

ORIGINAL QUESTIONNAIRE

Vážení kolegové a kolegyně, jmenuji se Lenka Přibilová a jsem studentkou pedagogické fakulty MU, obor učitelství Aj pro ZŠ a JŠ. V současné době se věnuji diplomové práci na téma „Hodnocení mladších žáků". Tímto bych Vás chtěla požádat o vyplnění přiloženého dotazníku. Dotazník je anonymní a jeho výsledky budou použity pouze pro potřeby výzkumu mé diplomové práce.

1. Proč jste se rozhodl/a stát učitelkou anglického jazyka?

2. Co hodnotíte u svých žáků?

 a. Aktivita

 b. Snaha

 c. Samostatnost

d. Zájem o výuku

e. Vědomosti

3. Proč a jak často testujete své žáky?

4. Jaké formy hodnocení znáte?

5. Co testujete u svých žáků?

 a. testuji vždy 1 kategorii (slovní zásobu, gramatické jevy atd.)

 b. mám rád/a komplexní testy zahrnující všechny oblasti

 c. testuji žákovy schopnosti poradit si s daným problémem

 6. Jakým způsobem žáky nejčastěji testujete?

 a. ústní zkoušení

b. písemný test

c. obě varianty

7. Zajímáte se o známky svých žáků?Snažíte se žákům individuálně pomoci? Jaký je váš postoj ke známkování žáků?

8. Znáte postoj rodičů vašich žáků a žáků samotných ke známkování?

9. Jaký je váš postoj ke slovnímu hodnocení?

a. Jsem jeho zastáncem

b. Nejsem jeho zastáncem

 c. Nemám vyhraněný názor

10. Slovní hodnocení se dle Vás zaměřuje na: Prosím, označte dle vaší důležitosti (1=hodně souhlasím-5=nesouhlasím)

a. Rozsah zvládnutí vědomostí

b. Postoje žáka k učivu

c. Osobností rysy žáka

d. Sociální dovednosti

e. Snaha žáka

 f. Celkové posouzení žákovi osobnosti

11. Kterou formu hodnocení byste volili pro své dítě? Proč?

a. Slovní hodnocení

b. Klasifikace

12. Myslíte si, že by bylo vhodné na 1.stupni kombinovat výše zmíněné formy?

 a. Ano

b. Ne

c. Nejsem si jistá

13. Mají žáci možnost se v hodinách sami hodnotit? Pokud ano, jak?

a. Ano

 b. Ne

ENGLISH TRANSLATION (contained within the thesis)

1. Why did you decide to become an English language teacher?

2. What do you assess most (participation, effort, knowledge, independence, interest in learning)?

3. Why and how often do you test your learners?

4. What forms of assessment do you know?

5. What areas do you assess, concerning the curriculum?

6. What is the most often used form of testing (oral testing, written testing, both forms)?

7. Do you know the attitude of your pupils´ parents towards assessment (yes, no)?

8. What is your attitude towards verbal assessment?

9. Verbal assessment is focused on (knowledge, attitude towards learning, individuality, independence, effort)?

10. Which form of assessment would you choose for your own child and why?

11. Do you think it would be appropriate to combine the above-mentioned forms in the lower primary school (yes, no)?

12. Do your learners have a chance to assess themselves in the lessons? If yes, how? (yes, no)

13. If your learners have a chance to assess themselves, how often do you let this (daily, once a week, once a month, as needed)?

REFERENCE:

Pribilova, L (2008). *Assessing Young Learners.* Published diploma thesis. Brno: Department of English language and literature of the Faculty of Education.