



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Dpto. de Lingüística Aplicada

Aplicación de las inteligencias artificiales en la enseñanza
de lenguas: una propuesta para aprendientes de lengua
inglesa de secundaria

Trabajo Fin de Máster

Máster Universitario en Lenguas y Tecnología

AUTOR/A: Sanchez Zugasti, Jaime

Tutor/a: Olmo Cazevielle, Françoise Therese

CURSO ACADÉMICO: 2021/2022

Agradecimientos

En primer lugar, gracias a mi tutora Françoise por orientarme, ayudarme y guiarme en la realización de este trabajo final de máster. Gracias también a la comunidad educativa de la Universidad Politécnica de Valencia por brindarme la oportunidad de aprender e investigar sobre el tema que desarrollamos en esta investigación. También debo agradecer al Colegio Teresiano del Pilar, al departamento de enseñanza de inglés y en especial al alumnado por permitirme realizar el estudio en tres de sus clases. Gracias a Kappa por los momentos de descanso. Muchísimas gracias a Natalia por ayudarme, corregirme y acompañarme en la realización de este trabajo y, por último, mencionar a mi familia, sobre todo a mis padres, sin los que nada de esta etapa de mi vida habría sido posible.

The application of Artificial Intelligence in language learning: a proposal for ESL students in secondary schools

This MA thesis consists of a study about Artificial Intelligence (AI) technologies' application in the field of language learning. After focusing on AI, this pilot project intends to analyse some of the existing apps, choose some of them and implement them for ESL students in secondary school. The usage of this kind of software is presented as a new tool and a complement for the classes and not as a substitute for the teachers' traditional figure. The results show, on one hand, a didactic proposal that can be extrapolated to other educational contexts of English teaching and, on the other, they evaluate the activities offered to the students and the teachers that have tried these apps.

Keywords: Artificial Intelligence, learning/acquisition of the English language, Communicative competence, digital technology, transversal competencies.

Aplicación de la inteligencia artificial en la enseñanza de lenguas: una propuesta para aprendientes de lengua inglesa de secundaria

Este trabajo fin de máster consiste en un estudio sobre la aplicación de las tecnologías de inteligencia artificial (IA) en el campo de la enseñanza de lenguas. Tras ahondar en las IA, este proyecto piloto pretende analizar las aplicaciones existentes, elegir algunas de ellas e implementarlas para el aprendizaje de la lengua inglesa en secundaria. El uso de *software* se plantea como una herramienta más, un complemento a las clases y no, como un sustituto de la figura tradicional del profesor. Los resultados exponen, por un lado, una propuesta didáctica extrapolable a otros contextos educativos de aprendizaje del inglés y por otro, evalúan las actividades ofrecidas por estas aplicaciones por parte del alumnado y del profesorado que las han experimentado.

Palabras clave: inteligencia artificial; enseñanza-aprendizaje de la lengua inglesa; competencia comunicativa; tecnología digital; competencias transversales.

Aplicació de la intel·ligència artificial a l'ensenyança de llengües: una proposta per a aprenents de llengua anglesa de secundària

Aquest treball de fi de màster consisteix en un estudi sobre la aplicació de les tecnologies d'intel·ligència artificial (IA) al camp de l'ensenyança de llengües. Després d'aprofundir en les IA, aquest projecte pilot pretén analitzar les aplicacions existents, elegir-ne algunes i implementar-les per a l'aprenentatge de la llengua anglesa a secundària. L'ús d'aquests software es planteja com una ferramenta més, un complement a les classes i no, com un substitut de la figura tradicional del professor. Els resultats exposen, d'una banda, una proposta didàctica extrapolable a altres contextos educatius d'aprenentatge de l'anglès i, d'altra, avaluen les activitats oferides per aquestes aplicacions per part de l'alumnat i del professorat que les han experimentat.

Paraules clau: intel·ligència artificial, ensenyança-aprenentatge de la llengua anglesa; competència comunicativa; tecnologia digital; competències transversals.

Table of Contents

| | |
|--|-----------|
| 1. INTRODUCTION..... | 7 |
| 2. THEORETICAL FRAMEWORK..... | 11 |
| 2.1 History and Evolution of Artificial Intelligence..... | 11 |
| 2.2 Modern AI and Its Applications..... | 21 |
| 2.3 The Application of AI into Education: Intelligent Tutoring Systems | 31 |
| 2.4 The Effects of Computer Assisted Language Learning (CALL)..... | 36 |
| 2.5 From CALL to MALL: The Use of Mobile Devices in Education..... | 41 |
| 2.6 English Standards, Common Goals and Transversal Learning..... | 44 |
| 2.7. Artificial conversations: the use of chatbots in language learning . | 51 |
| 3. METHODOLOGY OF THE RESEARCH..... | 54 |
| 3.1 Chatbot App: Andy English Speaking Bot..... | 54 |
| 3.2 Control Group and Previous Work | 57 |
| 3.3 Experiment: Previous Knowledge, Use of Andy and Satisfaction Questionnaire..... | 61 |
| 4. DIDACTIC PROJECT | 66 |
| 4.1 Test Group and Level of The Experiment | 66 |
| 4.2 Evaluation Items and Objectives of The School..... | 67 |
| 4.3 The Experiments Goals and Objectives..... | 68 |
| 4.4 Andy as a Conversational Auxiliary Teacher | 69 |
| 5. RESULTS..... | 71 |
| 5.1 Students' Previous Knowledge Questionnaire Results..... | 71 |
| 5.2 Students' Exercises Results in the App..... | 72 |
| 5.3 Students' Opinions Towards the Andy App | 74 |
| 5.4 Teacher's Assessment of the App and Its Use in the Classroom..... | 79 |
| 6. CONCLUSIONS..... | 81 |
| 7. REFERENCES..... | 85 |
| 8. ANNEX..... | 95 |

Figures

| | |
|--|----|
| Figure 1. How a knowledge-based approach works | 16 |
| Figure 2. Artificial Intelligence in Education | 58 |

Graphics

| | |
|---|----|
| Graphic 1. The use of educational apps in the classroom | 65 |
| Graphic 2. Available language learning apps | 65 |
| Graphic 3. Number of correct answers | 77 |
| Graphic 4. Correct answers in 'much, many' exercises | 78 |
| Graphic 5. Correct Answers in 'some, any' exercises | 79 |
| Graphic 6. Frequency and uses of English..... | 81 |
| Graphic 7. Use of language learning apps of the students..... | 82 |
| Graphic 8. Given punctuation to different features of the app | 83 |
| Graphic 9. Natural conversation with ANDY | 84 |

Images

| | |
|--|----|
| Image 1. Conceptual language analysis | 18 |
| Image 2. Image recognition software..... | 30 |
| Image 3. Chat interface in ANDY and speaker function | 61 |
| Image 4. Exercises in the app | 67 |
| Image 5. Example of QR slide to access app market. | 68 |
| Image 6. Types of exercise in the previous knowledge test..... | 69 |
| Image 7. Quantifiers exercises through the chat | 71 |

Tables

| | |
|---|--------|
| Table 1. Intelligences of AI systems | 24 |
| Table 2. Evolution of ITS definition | 35 |
| Table 3. Evolution of Computer-Assisted Language Learning | 43 |
| Table 4. Possible apps and their main features..... | 60 |
| Table 5. Features of ANDY, English Speaking bot..... | 62, 74 |
| Table 6. School's objectives and experiment's objectives. | 82 |
| Table 7. Most popular language learning apps..... | 85 |
| Table 8. Comments and opinions of ANDY..... | 81 |

1. INTRODUCTION

The learning and teaching world have recently suffered numerous drastic changes. The changes have permeated all levels and components of the educative environment, and all its members must adapt constantly. One of the agents that have influenced the most development of education is technology. Technology is in a never-ending development state; consequently, every other sector influenced by it also changes, develops and applies new technologies. Cloete (2019) refers to technology as “it is the accelerated pace at which very sophisticated technology is developing that makes it a game changer around the globe today” (2019, p. 1-2).

Our world is now driven by technological development. Our economy is technology-based, and many kinds of technology, like mobile devices, that are being developed focus on making our life easier and more accessible. In our era of modern technology, education suffers constant changes caused by adapting and incorporating new technologies that create better materials, new techniques or innovative ways of transferring knowledge. “The reliance and dependence of such an innovation, that simply makes life an easy, smooth journey is completely unavoidable these days even in schools, universities and colleges” (Raja and Nagasubramani, 2018, p. 33). The technologisation of our world has been shifting the trends in education completely for the last years, to the point that many learning and teaching models rely almost solely on technology, as we can see in online teaching, learning apps or academia’s platforms for. information sharing Inside the vast range of educational subjects and specialisations, we have chosen English language learning as the field of our investigation. Second Language Acquisition (SLA) has been hugely benefited by technological development as the communication, transfer of materials and interaction with native users of the objective language can now be immediate. Along with technology development, new fields of investigation and application have appeared inside SLA. Relatively new areas like Computer Assisted Language Learning (CALL) or Mobile Assisted Language Learning (MALL) are some of the applications that researchers and teachers have found for technological advancements in language learning.

It is inside CALL where we found the source of our investigation. CALL, and its actual evolution, MALL, are in continuous progression, as the computer field is one of the fastest developing technologies. Computers and personal devices are improving every day, and their possibilities widen steadily. One of the most significant developments in the world of technology is the creation of Artificial Intelligence (AI). With its origins in the middle of the last century, Artificial Intelligent systems are now present in every piece of technology we use and have some trajectory in the language learning field. AI's first applications to the learning environment can be traced back to the decade of the 60s. Within the early steps of AI systems, some researchers saw the possibility of a computer system acting as a teacher. Intelligent Tutoring Systems appeared as the first opportunity to apply Artificial Intelligence for teaching purposes. Intelligent Tutoring Systems (from now on ITSs) have remained one of the main options when applying AI to an education system, as we can see in modern language learning apps (Woolf, 2015, p. 38). If we analyse the actual state of AI in education, we can see that ITSs are the predominant systems in the market. ITS, now embedded in mobile device apps, presents a learning method where the student independently completes the exercises and follows the materials proposed by the AI tutor. These exercises and materials are customisable according to the student's past results and preferences. Instead of the AI tutor as a replacement for the human tutor, we focus our research on an AI tutor as an aid for the human teacher. An AI-powered tool that allows the teacher to practice some aspects of the language in a more profound and customisable way. We have chosen to test an AI-powered chatbot called ANDY for our investigation. ANDY is an AI chatbot that specialises in English learning and uses traditional exercises and simulated conversations to teach English to students. Our experiment consists of a trial of the Andy app in an English classroom in a Spanish high school. The use of technology in classrooms is still not very extended, and our research can help search for new methodologies, techniques and materials for learning a second language. Nowadays, schools have still not transferred their educational method to more technologically advanced platforms and methods, and the use of technology in the learning process is still very restricted. The schools still rely on traditional methods to teach the subjects, and technology remains in the background. "The

current use of digital resources in K12¹ and higher education can be described as dysfunctional: many school stakeholders can't find sufficient effective digital resources, while large collections of resources exist and sit online, waiting to be discovered" (Woolf, 2015, p. 40). The use of AI technology as a classroom tool can bring many benefits to second language acquisition, and with our research, we try to define these benefits and test the suitability of the app we chose to achieve them. The application of AI apps in language learning is still in its early steps of development, as most of the recent investigations are based on comparing the existing AI chatbots and the features that can be applied to education, like Haristiani (2019), Huang et al. (2022) and Pham et al. (2018) or investigations where new apps and systems are developed by the researcher, like Tono (2019), Tomlinson (2012) and Lee and Park (2020).

Our main objective is to test the effectiveness of an AI chatbot as a language learning tool for learners of English as a foreign language. This will allow us to see how the use of modern technology can be applied to traditional classrooms and techniques with the help of mobile devices and how it works as a support tool for the teacher. To achieve this, we set out the following specific objectives:

- Define the various benefits that can arise from using an AI application in a real classroom.
- Demonstrate that an AI-powered tool can be used to improve the English language skills of secondary school students.
- To show students the learning possibilities that their mobile device can offer and how technology can be applied to traditional learning methods.

In order to achieve these objectives, we pose the following research questions to which we will try to seek answers.

- Are AI apps easily applicable in a language classroom?
- In which aspects does a language-oriented chatbot benefit learning? Does it bring improvements for the teacher, students or both?
- Are the students familiarised enough with educative technology? Does it boost the student's motivation?

¹ K12: Term used to describe the years financed by the Government in the United States. From Kindergarten to high school

- Does an AI chatbot work as a conversational support tool in an ESL classroom?

By answering these questions, we will be able to prove whether the use of an AI chatbot is viable in an English learning class and how its use can be adapted to the school's educational goals. Our investigation is mixed, with both quantitative and qualitative analysis, to consider all the features of the app and the student's progression. First, in the second chapter, we revise the history and evolution of Artificial Intelligence and its application for educational purposes. We also analyse the effect of technology on language learning and the most recent systems, theories and approaches. The next step in our research, in Chapter 3, is to design the experiment where the teacher is asked about the learning objectives and methodology of the class, we check the previous level of the students, the students use the app in a class session and finally, the students are asked about their opinions and impressions towards the activities performed in the app. Then, in Chapter 4, we suggest a teaching model that includes the use of this app in the teaching process. After analysing, in the Chapter 5, the results of both the questionnaires and the app's exercises, we can conclude by answering the research questions and assess the completion of our objectives in Chapter 6.

2. THEORETICAL FRAMEWORK

2.1 History and Evolution of Artificial Intelligence

Nowadays, technology is a force that shapes the world we live in. Everything, from society and its culture to scientific research and knowledge preservation is influenced by technological advancements. All around us has suffered a technological advancement in recent years, an advancement that is unstoppable and faster as the time passes. Technological goals are becoming more and more ambitious to the point of creating human intelligence inside a device. This objective, one of the most ambitious of history, is known as Artificial Intelligence. If we try to find the origin of the idea of artificial intelligence, the first concepts and theories of an autonomous mechanical artificial being date back to ancient Greek mythology, where the god Hephaestus created Talos, “the giant bronze automaton” (Mayor, 2018, p. 1). The idea of artificial thinking stayed in the minds of philosophers, writers and mathematicians for many centuries. In the early 16th century, Jonathan Swift depicted in *Gulliver's Travels*, his most famous novel, a device known as the engine. This device was a human-like mind that allowed exceptional writing performance:

Every one knew how laborious the usual Method is of attaining to Arts and Sciences; whereas by his Contrivance, the most ignorant Person* at a reasonable Charge, and with a little bodily Labour, may write Books in Philosophy, Poetry, Politicks, Law, Mathematicks and Theology, without the least Assistance from Genius or Study. (Swift, 2005, p. 171)

Although fictional, we can consider this device as one of the first examples of a human-assisted computer. Furthermore, this is one of the first examples of a human language-based computer. With the appearance of the first modern computers, the idea of a “digital human mind” started to strengthen. Some examples began to appear in popular culture, like the Machine Man in Fritz Lang’s *Metropolis* (1927). Shortly after, the first complex computers stirred up the world of science by resolving complex equations in mere seconds. Claude Shannon published *Programming a Computer for Playing Chess* (1949), the first article that presented the idea of an autonomous computer program capable of challenging

a human to a chess game. Although simple, the possibilities, combinations, and movements represented the first example of the computer imitating the human way of thinking. Some years after, John McCarthy first described Artificial Intelligence (AI) as “the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it” (McCarthy et al., 2006, p. 1). In 1950, Alan Turing presented the basis of one of his most representative works, a test capable of testing the “humanness” of a computer program. With the Imitation Game (Later known as the Turing Test), the AI will be developed in a way that, according to Turing, will be hard to differentiate from a human mind:

I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9 , to make them play the imitation game so well that an average interrogator will not have more than 70 percent chance of making the right identification after five minutes of questioning. (Turing, 1950, p. 442)

This test developed over the years, and it became the standard evaluation when putting Artificial Intelligence to the test. The main aim of the Turing Test is to find whether a human investigator can differentiate between an AI's and a human's responses. Several years later, in 1955, the first Artificial Intelligence was developed at Carnegie Mellon University. “The AI system, called Logic Theorist, was designed by Allen Newell, Herbert A. Simon, and implemented by J. Clifford Shaw” (Benko and Sik, 2009, p. 4). This system contained 40 mathematical theorems, and the system was used to create a paper applying these theorems. Still, the Journal of Symbolic Logic later rejected the report as the editors did not notice that the AI system was one of the co-authors. The summer next year, in 1956, a group of researchers formed by John McCarthy from Dartmouth, M.L. Minsky from Harvard, N. Rochester from IBM and C.E. Shannon from Bell Telephone created a summer research group at Dartmouth university that finally coined and established the term Artificial Intelligence. Nowadays, 1956 is considered to be the year when Artificial Intelligence started as a field of study and from that moment onwards, research and progress started to develop (Dartmouth, 2006).

The same investigators that created the Logic Theorist shifted the focus of their study and changed from mathematical application to a more universal problem focus. In 1959, Simon, Newell and Shaw created the General Problem Solver (GPS). This AI program was the first to focus on the means and the ends, not only the latter, as the previous ones did.

The General Problem Solver (GPS) was the first useful AI program, written by Simon, Shaw, and Newell in 1959. As the name implies, it was intended to solve nearly any problem. This is important to note. Obviously, GPS was not the first software ever written; but software was previously written with very specific goals. The software solved one problem. G.P. S., however, was designed to solve lots of kinds of problems, using the same "reasoning" mechanism (i.e., algorithm) for every problem. (Eckroth, 2021, p. 1)

This program was instructions and variables that allowed it to solve general problems in a way that tried to resemble the human chain of thought. Although rudimentary, the programming allowed this AI to resolve problems like the Seven Bridges of Königsberg or the puzzle of the Hanoi Towers. This first program set the basis of the cognitive simulation, "AI systems general schemes of human ways of problem solving" (Benko and Sik, 2009, p. 4). John McCarthy (1959) joined MIT's AI research and presented the paper titled Programs with Common Sense at The First International Symposium on Artificial Intelligence. McCarthy defines the main characteristics of this program in the paper:

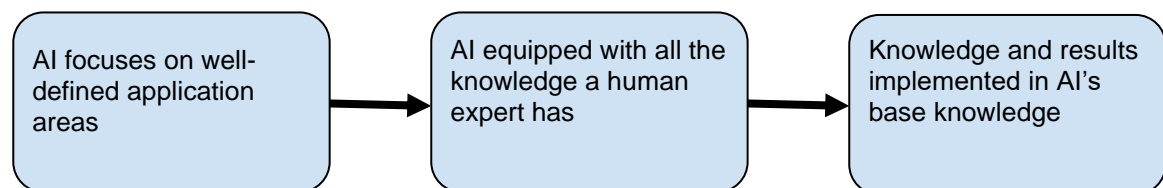
Programs to manipulate in a suitable formal language (most likely a part of the predicate calculus) common instrumental statements. The basic program will draw immediate conclusions from a list of premises. These conclusions will be either declarative or imperative sentences. When an imperative sentence is deduced the program takes a corresponding action. These actions may include printing sentences, moving sentences on lists, and reinitiating the basic deduction process on these lists. (1959, p. 1)

After several unsuccessful attempts at publishing his investigation, McCarthy designed a programming language specialised in the design of an Artificial intelligence. For two years (1956-1958), he created a programming language whose characteristics eased the creation of AI systems. Some of the main features were: “computing with symbolic expressions rather than numbers, representation of symbolic expressions and other information by list structure in the memory of a computer” (Liskov, 1974, p. 173). LISP is still used today and remains the second oldest computing language. This programming language cleared the way for many other LISP-based dialects and other languages to be developed, languages that favoured the creation of new AI systems as their characteristics and method of programming were AI-oriented.

In the decade of the 60s, a new approach to AI appeared. Until then, the two main perspectives towards the design of an AI were the logic-based approach (with a given set of rules, the system creates an inference) (Calegari et al., 2020) and the cognitive simulation approach (with narrow hand-given input, the system simulates human thinking and gives results that had to be reviewed (Forbus, 2010). In 1965, the knowledge-based approach was first implemented in the chemistry field. As can be seen in Figure 1, Benko and Sik indicate that this approach has three main characteristics:

Figure 1.

How a knowledge-based approach works:



Note. Adapted from *History of Artificial Intelligence* (p. 5), by Benko and Sik, 2009, Encyclopaedia of Information Science.

With these three approaches extended in the field of research on AI, John R. Searle, an American philosopher, divides the AIs into “weak” and “strong”. The main difference between these two categories is that “weak” AIs “enable us to formulate and test hypotheses in a more rigorous and precise fashion” while in ‘strong AIs’, ‘the computer is not merely a tool in the study of the mind; rather, the

appropriately programmed computer really is a mind" (Searle, 1980, p. 2). This division firmly separated the AIs according to their capabilities, and Searle, in the same text, proposes a new test for an AI as an alternative to Turing's test called The Chinese Room. In this test, Searle compares a human trapped in a room to a computer. Searle pictures himself inside the room and follows the instructions and symbols given to him in Chinese under the door, to the point where he can pronounce correct Chinese words but without understanding them. Chinese-speaking people outside that room will understand and think that Searle knows how to speak Chinese, but in reality, he doesn't; he is just following instructions. Searle compares this false-speaking situation to the functioning of an AI. Many AIs can succeed when performing a task, but only because they are given specific functions and instructions. In the same way, the human in the room lacks the understanding of the Chinese language, weak AIs do not have any kind of understanding of the data they are given. Also, the concept of Artificial Intelligence started to take its first steps in cultural media. As the investigation in this field progressed, the possibilities that the AI offered widened, and the presence of AI in mainstream culture increased, as can be seen in HAL from *2001: A Space Odyssey* (1968) or in C-3PO in *Star Wars* (1977).

As we have seen, AI development quickly shifted from mathematical models and applications to a linguistic approach. Because humanlike-minded machines were the main objective, AI engineers focused on giving these systems the ability to understand and communicate with human natural language. Therefore, in the early years of AI development, many significant linguists got involved in language theories and systems that greatly affected AIs. American linguist Noah Chomsky published in 1966 his work *Topics in the Theory of Generative Grammar*. In his work, he presented a new linguistic theory regarding language learning and human understanding of the language, which also affected AIs. With this theory, Chomsky presented the idea that a finite number of language rules can create an infinite number of language combinations. "The rewriting rules of the base component and the rules governing ordering and arrangement of transformations generate an infinite class of T-markers" (Chomsky, 1966, p. 55). This linguistic model presented by Chomsky has stayed at the centre of the development of computer science and Artificial intelligence, and still today, many systems are created following this model. Following Chomsky's research, the field of

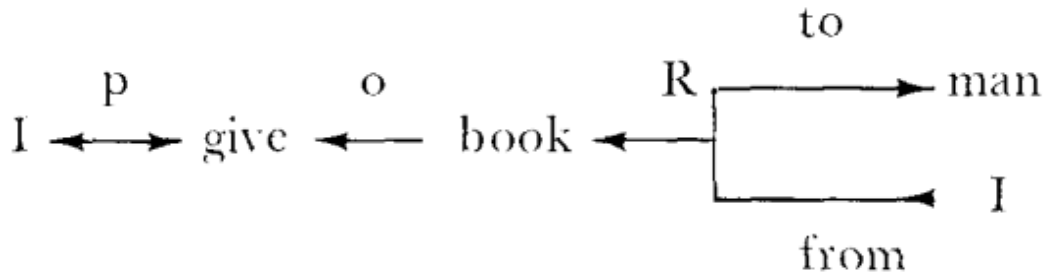
mathematical linguistics gained relevance. “Mathematical linguistics refers to the use of rigorous mathematical frameworks to model aspects of languages, either natural or formal” (Fulop, 2012, p. 2110). This new field of research allowed the researchers to apply mathematical rules to both formal and natural language, which proved to be very useful in the development of Artificial intelligence and computer science. The second main linguistic theory that greatly affected early artificial intelligence is the conceptual dependency theory by Roger Schank, an American Artificial intelligence theorist. In his work *Conceptual dependency: A theory of natural language understanding* (1972), Schank defends the importance of concepts against the previously mentioned theory that placed syntax as the most important component of language.

A conceptual base exists; that its elements are concepts and not words; that the natural language system is stratified with the actual language output being merely an indicator of what conceptual content lies beneath it; and that the conceptual apparatus that we tend to call thinking functions in terms of this conceptual base, with concepts and the relations between these concepts as the operands. (Schank, 1972, p. 554)

This theory organised the concepts that form language in a series of different relation types and hierarchies, similarly to, as Schank defends, the human mind. These relations and hierarchies are presented in different kinds of graphs that help us understand this model. In Image 1, we can see an example of a graph representing the phrase ‘I gave the man a book’.

Image 1.

Conceptual language analysis



Note. Taken from *Conceptual dependency: A theory of natural language understanding* (p. 567), by R. Schank, 1972, *Cognitive Psychology*.

This representation helps us visualise the kind of relationship each word has with the adjacent ones. This theory quickly gained recognition and promoted the creation of new theories and models that tried to represent this concept-based linguistic model. Two of the most significant new models proposed are frames by Martin Minsky and scripts by Robert P. Abelson and Schank. These models helped and served as the basis for the development of Natural Language Processing (NLP), one of the most relevant AI-related technologies. "Natural language processing (NLP) is a collection of computational techniques for automatic analysis and representation of human languages, motivated by theory" (Chowdhary, 2020, p. 604). Thanks to these concept-based models, many systems of NLP systems were developed, and the application of NLP has widened and extended to many other fields since. One of the first examples of this technology is ELIZA, a natural language processing program developed by Joseph Weizenbaum of MIT that "was incorporated in various programmes and given a human interface" (Guan et al., 2020).

The third linguistic theory that strongly shaped the development of AI is cognitive linguistics by George Lakoff. He published his book *Women, fire and dangerous things: What categories reveal about the mind* (1987) and in it, he challenged and criticised what was known as strong AIs. His main attack was against the "boxing" of the concepts of previous models. For Lakoff, our concepts and their representations in our minds are not only formed by the concept itself, but it is a combination of the "crude" concept, our senses and what that concept has produced in our imagination. He introduces the term "fuzzy", categories outside

the binary categorisation of 0 or 1 and can be conceived as an “in-between” (Lakoff,1987, p. 21). According to Lakoff, many things can affect our categorisation of concepts, and it is impossible to create a universal representation of these concepts. We are strongly influenced by our personal experience, our body and our imagination, and according to modern philosophy and psychology, a machine that tries to imitate our thinking should have more similar behaviour to the one our brains have.

As a consequence, human reason crucially depends on the same factors, and therefore cannot be characterised merely in terms of the manipulation of abstract symbols. Of course, certain aspects of human reason can be isolated artificially and modeled by abstract symbol manipulation, just as some part of human categorisation does fit the classical theory. But we are interested not merely in some artificially isolatable subpart of the human capacity to categorise and reason, but in the full range of that capacity. (Lakoff, 1987, p. 25)

Modern AIs' objective should be to mimic the categorisation of the human mind, and as Lakoff defends, that is not possible under the traditional conception of reality and categorisation. A more modern approach refuses the absolutes (or the binary division) and has to take into account the influences our reason can be affected by. This theory presented the possibility of an AI imitating our brain in a more literal way. Instead of following a programmed chain or thought process, an AI should be composed of “neurons”, and its ramified connections would confer the program more human-like reasoning. This theory focuses on the importance of the connections between concepts, not the concepts themselves. This priority given to the connections resulted in the creation of connectionist models.

Connectionist models take inspiration from the manner in which information processing occurs in the brain. Processing involves the propagation of activation among simple units (artificial neurons) organised in networks, that is, linked to each other through weighted connections representing synapses or groups thereof. (McClelland and Cleeremans, 2009, p. 2)

These models took neurons as the smallest processing unit and tried to imitate them. Neural networks had been presented previously, but due to their basic design and few applications, researchers in this field did not receive enough support and financing. Despite the lack of financing and significant research projects, researchers kept investigating the possibilities these new models could offer. In 1986, the investigation results by David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams were finally published. In their work *Learning representations by back-propagating errors*, they presented a model where “neuron-like units” form a network that is able to autonomously “adjust the weight of the connections in the network to as to minimise a measure of the difference between the actual output vector of the net and the desired output vector” (Rumelhart et al., 1986, p. 533). This system was able to analyse its own process and result and compare both with the desired result/output. By comparing these elements, the system could adjust the parameters of its connections in order to achieve a better result, all this without little to no human intervention. This autonomous working system was one of the first examples of an AI capable of mimicking the human feature of “learning”. This research revived the interest in multi-layered neural networks, so the number of academics interested in this field rose. This interest peaked at the 1987 First International Conference on Neural Networks (Caudill, 1987). Shortly after this first conference, an extended trend in this field was to imitate the brain's biological functioning as the model for designing AIs. This approach tries to copy our brain's structural and functional design. Then, inside this segment of AI research, a specialisation known as evolutionary computing started to gain strength. This field of research tries to apply the natural equations that we have observed in the process of evolution into a digital and technological platform.

The simplest evolutionary algorithm can be viewed as a search procedure that generates potential solutions to a problem, tests each for suitability, and then generates new solutions. It's important to understand how this process differs from exhaustive search or blind random search. (Fogel et al., 2016, p. 216)

When applied to an AI, Evolutionary computing tries to achieve the optimal solution to a problem. Several potential solutions are calculated simultaneously, and the system chooses the optimal one. This selection process is similar to natural selection and combined with learning capabilities; we can compare it to mutation processes. Following the evolutionary approach, the academic proposed other models that could apply natural engineering to technological systems. The main four were: “genetic algorithms, evolution strategies, evolutionary programming and genetic programming” (Miettinen, 1999). AI’s related communities took the biological-imitating approach as one of the most proper ways of designing Artificial Intelligence, so it has remained one of the principal methods and models of AI design.

Another significant AI model design is pattern recognition. Pattern recognition dates several years back before the term Artificial Intelligence was coined and extended, but its application to AI technology made it gain recognition in the computing field. Jain, Duin and Mao define “pattern recognition” as “Automatic (machine) recognition, description, classification, and grouping of patterns” (2000, p. 4), understanding patterns as several common features that can be recognised in unknown objects (objects represent any unit that can be subject of study). Pattern recognition can benefit strongly from all AI’s features that we have previously mentioned. Nowadays, pattern recognition is used in the vast majority of services and computational processes thanks to technological development.

The rapidly growing and available computing power, while enabling faster processing of huge data sets, has also facilitated the use of elaborate and diverse methods for data analysis and classification. At the same time, demands on automatic pattern recognition systems are rising enormously due to the availability of large databases and stringent performance requirements (speed, accuracy, and cost). (Jain, Duin and Mao, 2000, p. 5)

2.2 Modern AI and Its Applications

Since the start of the new century, AI implementation has extended to every aspect of technology. Since AI's first creations and development, its application was reserved for very few fields, like scientific research or military simulations. The expansion and availability of technology were incredibly fast, and we have moved from specialised, slow computers to personal, hand-sized devices. All the devices we use today, and the companies and systems needed for them to work according to our needs, make use of AI systems that combine many of the models and features previously described. "Our lives are already saturated with AI-evidenced in the rise of chatbots, Google Maps, Uber, Amazon recommendations, email spam filters, robo-readers, and AI-powered personal assistants such as Siri, Alexa and Echo" (Elliot, 2019, p. 20). In the last years of the 20th century, AI as a concept had extended widely, and some of the feats that had been set as objectives in previous years had been achieved. In 1996, Russian chess grandmaster Garry Kasparov started a series of matches against IBM's supercomputer *Deep Blue*. After several mediatic games and despite Kasparov's initial victories, one of the best human chess players had been defeated by a computer (Campbell, Hoane and Hsu, 2002, p. 57). This highly mediatic set of matches was one of the first examples that demonstrated to the public the capabilities of a modern AI

As we previously mentioned, most recent AIs are designed by combining several systems and models in order to achieve better performance. The combination of previous technologies and advancements in both software and hardware has caused a mutation in how we classify AIs. Kaplan and Haenlein (2019) propose a new classification in their paper about the implication of Artificial intelligence. Instead of focussing on the systems and models that make the AI work, they classify the programs by whether or not they possess several kinds of "intelligence". These intelligences are all possessed by a human being, and AI's main objective of AIs is to fulfil all these. The four intelligences that humans have are: "Cognitive intelligence, Emotional intelligence, Social Intelligence and artistic creativity" (Kaplan and Haenlein, 2019, pp. 17-18). Until now, AIs have only been able to conquer three of the four intelligences proposed by Kaplan and Heinlein. The intelligences and the evolution of AIs are shown in Table 1.

Table 1.

Intelligences of AI systems

| Intelligence | Expert systems | Analytical AI | Human Inspired AI | Humanized AI | Human Beings |
|------------------------|----------------|---------------|-------------------|--------------|--------------|
| Cognitive Intelligence | × | ✓ | ✓ | ✓ | ✓ |
| Emotional Intelligence | × | × | ✓ | ✓ | ✓ |
| Social Intelligence | × | × | × | ✓ | ✓ |
| Artistic creativity | × | × | × | × | ✓ |

Note. Adapted from *Siri, Siri, in my hand: Who's the fairest in the land?* (p. 18), by A. Kaplan and M. Haenlein, 2019, *Business Horizons*, 62 (1).

As we can see in the chart, artistic creativity is the only one that AIs have still not achieved. Actual AIs, on account of being designed following human example, can recognise human emotions and answer accordingly (emotional intelligence) and interact by being conscious of the other interlocutor (social intelligence). Both abilities are considered complex human features. Artistic creativity is the only one that the AIs are still unable to perform, but due to recent advances like DALL-E2 and story generators, there is an ongoing debate about whether to classify these advances as artistic creativity (Marcus et al., 2022).

As we have previously mentioned, the devices capable of running highly demanding programs like AI's were only accessible to powerful companies, well-funded universities or governments. With the quick shift in the availability of technology for consumers, many companies appeared intending to provide the huge mass of "technologized" consumers with all the services they may need. The vast majority of these services are now online or at least digital, so companies have been forced to shift their strategies regarding employees, physical establishment and customer service. In order to perform correctly, these companies need immense quantities of all kinds of data, not only from their

customers but from almost every other component of the technological world we live in. That's why we are currently living in the world of big data. "Big data is a term for massive data sets having large, more varied and complex structure with the difficulties of storing, analysing and visualising for further processes or results" (Sagiroglu and Sinanc, 2013, p. 42). Today, consumers are exposed to a continuous and immense flow of information throughout several channels. Despite having easy access to all the available information, this over-exposition to unfiltered information can cause negative effects. Elena Benito Ruiz, a researcher at Valencia Polytechnic University, defines two of the main negative effects that information can cause us today. First, she refers to our inability to keep up with the information flow. "Information Fatigue Syndrome (IFS): The cognitive inability to keep up with the ever-increasing amounts of available information" (Benito-Ruiz, 2009, p. 60). And then defines and updates the term Infoxication, a term that combines information and intoxication to give a name to the current state of the information system.

Infoxication 2.0: Infoxication 2.0 is a viral process, a ripped, mixed and burned virus coming from our most essential needs (information and communication), exponentially worsened by the myriad of Web 2.0 communication and networking possibilities. It refers to an intoxication of excessive informational and communicational demands. (Benito-Ruiz, 2009, p. 60)

For this reason, many of the consumer-oriented AIs developed today try to customise our experience and funnel the information we receive according to our likes. In order to do this, companies design these systems in a way that uses customers' information and automatically designs a custom feed of information and recommendations. "Besides dynamic pricing, firms use consumer profile data and make product recommendations" (Taylor, 2004 p. 632). Examples of these AIs are Amazon Shopping recommendations, Twitter recommended tweets or, the most famous of these technologies, Google's custom ads and news feed. All these technologies gather bits of our information, process it with an autonomous AI and offer us custom information according to our recent searches, conversations and much more. This technology has raised controversy in the last

years as many companies make gigantic profits just by trading with customer information, and because a great number of people are against it, the number of more-private browsers, platforms and devices is rising. AI application has permeated all levels of consumer technology, so the custom feeds and recommendations are not the only technologies that base their results on our past inputs. Written communication is made almost exclusively through digital means, so many communication platforms have implemented systems that help us write faster and more correctly.

Predictive systems are designed to offer suggestions that reduce writers' typing effort by offering shortcuts to enter one of a small number of words that the system predicts are most likely to be typed next. As such, the suggestions are, by construction, the words that are the most predictable in their context. (Arnold et al., 2020, p. 128)

As the user writes, the system can register and process the language and add it to its database. The system will later suggest similar words and structures that have been previously used by the user that fit in the context that the system automatically detects, and even place personal vocabulary of non-existent words registered by the user. Predictive text systems allow us to write faster and more efficiently.

As we have previously mentioned, one of the main objectives of early AI development was successful communication via human natural language. If we successfully combine the use of natural language with an autonomous system, we can get an artificial process of communication that mimics a real conversation. These systems receive the name of chatbot, and Adamopoulou and Moussiades define it as “a computer program, which responds like a smart entity when conversed with through text or voice and understands one or more human languages by Natural Language Processing (NLP)” (2020, p. 373). Text-based chatbots can be traced back to 1994 when Joseph Weizenbaum of MIT developed ELIZA. This system analysed the input sentence, recognised keywords and gave a preprogrammed response from its database (Weizenbaum, 1966, p. 33). Although very restricted and preprogrammed, this interaction gave the impression of real reasoning and conversation. Shortly after, in 2001, a trio of

investigators created a chatbot named Eugene Goostman. This chatbot, created to act as a 13-year-old boy, is considered to be the first chatbot that passed Turing's test. "The tests involved in which the machine, Eugene Goostman, achieved a 33% success rate when compared with a hidden human correspondent in each case" (Warwick and Shah, 2016, p. 990). This chatbot was an early teenager to justify its lack of knowledge in some fields and possible grammatical errors. Modern chatbots are designed to work in several different contexts and with various features. Depending on the intended application of the chatbot, it will have certain features. Lokman and Amedeen (2019) classify these features into four main classes:

Modern chatbot design can generally be categorised into several base elements which are: (1) knowledge (open or close domain), (2) response generation (retrieval or generative), (3) text processing (vector embedding or latin alphabet), and (4) machine learning (ML) model (usually using neural network). (p. 1012)

For our investigation, the two main classes that affect the app we use are knowledge and machine learning. When a chatbot has a general culture database and can converse with the user about any topic, it is considered to be an open domain chatbot. Many modern chatbots have models that allow the chatbot to access the Internet and retrieve information about the proposed topics in order to create a more natural and rich conversation. On the other hand, if a chatbot does not have the ability to treat every topic and can only use the topics stored in its database, it is a closed domain chatbot. Many companies use closed domain chatbots in the customer service processes to solve problems that can be easily solved without human intervention. 'Closed-domain chatbots are generally goal – oriented – the user is likely attempting to accomplish a task, such as asking a question, setting an alarm, or making a reservation" (Suta et al., 2020, p. 506). These closed domain chatbots have now extended to many fields, and now we can find one of the big four chatbots in every modern personal device. These big four chatbots, Apple's Siri, Amazon's Alexa, Google's Assistant and Microsoft's Cortana, can be used in many ways, and users are able to control their device and its features and many other devices like T.V.s, air conditioning, door locks

and many other, even cars. These personal virtual assistants are constantly updated and improved, and now they make use of voice recognition technology to “learn the words and phrases of the user’s voice in order to interact with users in a personalised manner” (Suta et al., 2020, p. 502). Voice recognition technology has gained importance in recent years as many AI’s rely on direct vocal input from the users. This technology is crucial for the development of chatbots and education software, and we will discuss it further on.

Although these assistants are classified as closed domain, users can ask the system for any kind of information, and if it is not within the assistant’s field of knowledge, it has relative freedom to make a search on the Internet for the requested information. Closed domain chatbots function following two main ways of generating a response. The “retrieval-based system” chooses the most appropriate response from its database according to the human input. This selection process can be fixed (following a set of rules), or machine learning can be applied to improve the relation of the responses. ‘Generative-based systems’, the other main model of development, can generate completely new responses not stored in its database. “After analysing the input, the chatbot will create an appropriate response, so it can give an answer to both known and unknown inputs” (Qian et al. 2021, p. 1467). These later systems are much more difficult to create, as the context and content of the conversations cannot be predicted, although if successful, conversation with the chatbot will be much more natural and human-like. This is one of the reasons why many chatbots have been implemented with a self-learning component.

As we have previously mentioned, actual AI models allow them to “fix” their own errors and gradually improve without human intervention. For example, the chatbot Replika is able to learn from the millions of simultaneous conversations with the users, and after analysing and comparing users’ input through several steps and models, it gives a response according to past users’ conversations, the topic and the context (Replika, 2022). This “self-improving” process is known as machine learning. As Bini defines it, “ML learns from experience and improves its performance as it learns” (2018, p. 6). The AI system is given by the developers a set of significant features that the system must recognise. The AI will learn from this ‘learning data set’, and it will be able to recognise these features in new input that has never been analysed before. Then, this input is incorporated into the

data set to be used as a reference in the future. The bigger the “learning data set” (also known as ‘ground truth’), the faster and more precise the system will be (Bini, 2018, p. 6). Developers try to use massive datasets to have an initial functioning system, like ImageNet for image recognition systems or GLUE dataset for English text analysis (Paullada et al., 2021, p. 1). Machine learning technology is now implemented in many AI systems. Not only in general, customer-oriented AIs but in more specific and private AIs, like medical or scientific AIs. “Big data” technology also brings a huge benefit to machine learning, as it makes possible the quick analysis of massive quantities of information from where the AI can learn.

Big data refers to datasets that are too awkward to work with using traditional, hands-on database management tools. This might be the case for a number of reasons – including (but not limited to) limitations in traditional methods’ ability to capture, store, manage or analyse data on the scale in question. (Yiu, 2012, p. 10)

Datasets can be automatically produced with big data analysis, and this hugely benefits the creation of new machine learning AIs whose initial datasets contain a massive number of examples and allows the system to learn and recognise features at an incredibly fast pace. These enormous datasets serve as the basis for AI’s intelligence, but a way to apply it is needed. Nowadays, Artificial Intelligence programs are developed with one or several types of inputs and outputs in mind. For this reason, different types of datasets are adopted depending on the kind of input and the output the AI will receive. The three most extended types of input in AIs today are images, text and voice. For these three types of inputs and databases, specific technologies must be implemented in the systems for the correct recognition of features and functioning of the AIs.

Many of the actual AIs are specialised in image recognition, and classification and their application are already present in many fields. Image recognition AIs analyse the image input, compare its features with a given model and then, basing their decision on a “confidence score”, gives the user the output (Fritz, 2021) as can be seen in Image 2.

Image 2.

Image recognition software.



Nota. Adapted from Image Recognition Guide [Image], by FritzLabs, 2021, Fritz.ai (<https://www.fritz.ai/image-recognition/>). © All rights reserved

Image recognition software can be programmed according to specialisation, and the classification of the images can be broadened to produce more than one tag from a single image input. This technology is known as “multi-classification”, and although “binary classification” can still be used, “multi-class classification is a more practical and significant problem in real-world applications” (Gu et al., 2014, p. 400). Image recognition software is applied in many everyday applications like Instagram’s automatic censoring algorithm (Cobbe, 2021, p. 741) or Google Maps orientation image-based system, whereby activating the user’s device camera, the app can precisely set the location by just analysing the immediate surroundings (Zamir and Shah, 2010, p. 256). Now, certain AIs can receive text-based inputs, analyse them, and produce an image-based output by creating a completely new image, a result of a combination of the main features stated in the text input. The main example of this type of technology is the AI DALL-E 2. (OpenAI, 2022). Image-fed and image-producing AIs are being developed at a very fast pace, and it allows users from every field to process and use a variety of inputs and optimise and handle image databases.

As we previously mentioned, Artificial intelligence, shortly after the first developments, tried to process human natural language. Natural Language Processing (NLP) research first focussed on written text recognition, but not much

time passed when researchers started to develop software that was able to process the other main human form of communication, the voice. In 1962, IBM presented a rudimentary voice recognition system considered to be the forefather of today's system.

The "IBM shoebox" was a rudimentary computer that could perform speech recognition of 16 spoken words and the digits 0 through 9, producing either a printed output of simple arithmetic computations or serial activation of light bulbs. The basic elements of that prototype still serve modern-day VR systems. (Fox et al., 2013, p. 191)

Modern voice recognition systems (VR) suppose a lot of benefits for both developers and users. For developers, it allows the systems to function properly without the need for additional input or output peripherals. For users, it drastically simplifies communication with the device and allows the user to communicate with the device with simple voice commands without even being close to it (Kim, 2020, p. 1). One of the best examples of this technology is Google's home assistant. The device captures words or phrases produced by the user, converts the acoustic waves into an electric current through a microphone, and these words are transformed into text that is processed by an NLP engine. (Arriany and Musbah, 2016, p. 3) These voice inputs can be completely similar and have the same function as text-based inputs.

Voice-based AI transforms the user inputs into the text to process them with a text NLP system. Text processing AI was one of the first focuses of the AI's research community, first with numbers and then with Natural language, as we previously mentioned. For this investigation, we have done an experiment with an English-learning-oriented chatbot, so text NLP systems are crucial in order for this app to work. NLP and text processing have many applications that widen and increase every day. From simple vocabulary translation to complex conversational chatbots, the systems need to be developed and updated constantly to be able to keep up with the information flow and the needs of users and researchers. After many years of development and different approaches, current NLP systems have a common course of action when analysing text fragments.

First, the intelligent system can detect the language of the text input, then the text needs to be fragmented into smaller units (clauses or single words) for the system to process it easier. Next, in order for the program to classify the words and understand their possible connections, language units get marked up grammatically (whether the word is a verb, noun, or adjective). Then, the program can automatically clean the data and delete all the bits of information that do not provide useful information (non-textual data, prepositions, articles, etc.). Lastly, the NLP system will define the relations between the language units and process them to achieve the goal of its programming (Quasthoff et al., 2014, pp. 3-24). As in any other AI, the logarithm can be varied, and in NLP systems, we mainly find two: ruled-based algorithm or machine-learning algorithm.

Machine learning algorithms usually consisted of intelligent modules which have the capability to learn from historical data. Before ML approaches, NL tasks are commonly carried out using rule-based approaches. In rule-based approaches rules were constructed manually by linguistic experts or grammarians for particular tasks. (Khan et al., 2016, p. 98)

Unless very specific, NLP systems have a combination of many of the aforementioned features. The function of the modern AIs can strongly vary depending on its main purpose, and developers choose the most suitable kind of algorithm in relation to the goal of the AI Today's main AI goals that can be found are text classification, text extraction, machine translation and natural language generation (NLG) among many others (Chowdary, 2020, pp. 608-609). In our case, we can see three of these four applied in the app used for our investigation. Being a chatbot, text classification, text extraction and natural language generation are needed in order to create a successful conversational system. Artificial Intelligence has become a crucial feature in today's technology. It is present in many apps and devices not only for general public-oriented but also for scientific and specific purposes. Its utilisation and applications have reached every aspect of society, and one of the most affected ones is education. In the next section, we will discuss the usage of AI-powered tools in education, how AI technology shifted the approach in education and how language learning and

teaching make use of AI, NLP and NLG for an optimal and more approachable way of language acquisition.

2.3 The Application of AI into Education: Intelligent Tutoring Systems

The development in technology greatly affected education and all its processes. The new platforms, exercises and availability of materials supposed one of the most significant changes in the history of education. As soon as Artificial intelligence started to be developed, its application in education was seen as a more than viable option by experts in the field. Oxford dictionary defines Artificial Intelligence as “the study and development of computer systems that can copy intelligent human behaviour” (Oxford, 2022). So, this imitation of “intelligent human behaviour” could be applied to the educational field. Teachers and students quickly saw this technology as a new tool to improve the educational process by applying it to all aspects.

The shift from paper and analogical equipment to digital devices allowed tutors to work much more efficiently and effectively. Technology kept developing, and it infiltrated all levels of education, with constant growth in the number of users and the quality of technology. If we limit the application of AI to the education environment, we can track the first examples of its use back to the decade of the sixties. Jaime Carbonell created an intelligent program to study South American geography, which gave students their results in natural language (Carbonell, 1970, p. 191). Today, this is considered the first example of what we know today as an Intelligent Tutoring System (ITS). As the technology of AI developed, the models and trends in education also changed. The term ITS was redefined several times, and its definition and functions broadened as its capabilities grew, as can be seen in Table 2. Several experts gave their own definitions, reflecting the faculty and the role the Intelligent Tutoring System was given at that moment in history.

Table 2.

Evolution of ITS definition

| Year | Definition |
|------|---|
| 1987 | AI techniques can permit the intelligent tutoring systems itself to solve the problems which it sets for the user, in a human-like and appropriate way, and then reason about the solution process and make comments on it. (Ross, 1987) |
| 1997 | The goal of intelligent tutoring systems (ITS's) would be to engage the students in sustained reasoning activity and to interact with the student based on a deep understanding of the student's behaviour. (Corbett et al., 1997, p. 849) |
| 2009 | Artificially intelligent tutors respond dynamically to the individual learning needs of each student. That is, an AI tutor does not employ a set of "canned" instructions, guides, or problems that are preprogrammed to anticipate particular student responses. Instead, an intelligent tutor constructs responses in real-time using its own ability to understand the problem and assess student analyses. (Johnson et al., 2009) |
| 2018 | An intelligent tutoring system is a computer system aimed at providing customised instructions and feedback to learners. Human tutors have the capability to observe most changes in affective state, but an intelligent tutoring system has the capability to recognise, respond and react to affect. (Odo, 2018, p. 521) |
| 2021 | Intelligent tutoring systems (ITS's) are a promising integrated educational tool for customising formal education using intelligent instruction or feedback. (Guo et al., 2021, p. 441) |

As we can see, AI capabilities and ITS's competencies grew concurrently. The first ITS system is defined as a tool capable of solving the challenges for the user and analysing the results in the same way a human tutor would. In the most recent definitions, the system is designed not only to function as similar as possible to a

human tutor but also to customise the materials, the exercises and the experience according to the users and their past results. With the advance in this field of research, ITS's goals and possible applications grew, so its components and features also had to grow. Sottolare (2016) defines in his work the four main components of an ITS.

First, "the domain model" consists of all "skills, knowledge, and strategies/tactics", usually topic-specific, prepared for the students. It also contains a similar level of knowledge that an experienced human tutor would have and all the usual errors of the students. Second, "the learner model" is more psychology-oriented, as it involves all the "cognitive, affective, motivational, and other psychological states" related to the learning process. This is one of the aspects that would be more affected by the customisable features of the ITS. Third, "the tutor model" combines the two previous models and chooses the best "tutoring strategies, steps, and actions" for the human tutor to do next. Lastly, the fourth model is the "user interface model". This model receives and understands the users' input, which can be varied in form (spoken, text, clicks, etc.) and creates an output that can also vary in form. "In addition to the conventional human-computer interface features, some recent systems have incorporated natural language interaction" (Sottolare, 2016, p. 3). These four components are commonly found in ITS and provide the system with the functionality of working in a similar way to a human tutor.

Intelligent Tutoring Systems are applied in a wide variety of fields, but unlike general AI systems like open-domain chatbots, their field of knowledge and application is very restricted. ITS are designed to tutor a certain subject, so all the previous components that form an ITS are single-goal-oriented and are more successful.

Computer science was the main ITS function background. While the theoretical basis for ITS's was education and psychology, the integration needed to be realised through technology. Therefore, the related computer science studies were focused on specialised computer science skills and the incorporation of educational or psychological frameworks. (Guo et al., 2021, p. 448)

As we can see, many educational fields make use of ITS's but, in the end, all fields need to adopt some computer science integration. Actual ITS's consist of a mix of computer technology, most of the time in relation to Artificial Intelligence, deep learning and user inputs, and any of the fields chosen by the researchers or developers. This is why Computer Assisted Language Learning (CALL), and Intelligent Tutoring Systems are fundamentally related. From the beginning of the developments of what were considered "intelligent" computer systems, Language Learning software was seen as one of the first and best options to progress.

First, individualised instruction by a competent tutor is far superior to the classroom style because both the content and the style of the instruction can be continuously adapted to best meet the needs of the situation. Secondly, students learn better in situations which more closely approximate the situations in which they will use their knowledge, i.e., they learn by doing, by making mistakes, and by constructing knowledge in a very individualised way. (Ferreira, Atkinson, 2009, p. 277)

Once again, customisation of the content and learning procedures is one of the most important features when choosing an ITS as a language learning tool. Technologies like Artificial Intelligence, machine learning and NLP, when applied to an ITS, have the capacity to customise users' experience by analysing and processing their inputs, results and other variables. Modern ITS can take into account many variables to further customise the learning experience. Several options of accessibility like "adaptive image visualisation" (Ahn et al., 2018, p. 413) and psychological variables like "positive and negative affect" (Odo, 2018, p. 521) suppose the next step in ITS development, as the systems are able to take into consideration some of the users' characteristics that only a human tutor could. The deepest adaptation the system achieves, the better the learning process will be. Contents will be better and more deeply learned by the user, and the learning process will be more attractive. This is similar if done with a human tutor, so ITS tries to imitate that teacher-student relation. "98% of the students with private tutors performed better than the average classroom student, even though all students spent the same amount of time learning the topics" (Nwana, 1990, p. 254). For this reason, many of the actual attempts on creating ITS's

focus on creating a personal, individual tutor for each user, and personal smartphones suppose an immense step towards just that. According to Corbett (1997), despite extensive research, ITS's have not achieved the expected popularity. He states two main reasons for it. "There are several reasons for this lack of penetration. Intelligent tutoring systems are expensive to develop, and until relatively recently, the necessary computing power was expensive to deploy" (Corbett et al., 1997, p. 850). Until now, processing power was too much for the regular user's devices, but with the development of household devices and the appearance of small personal devices, this last inconvenience has been slowly disappearing.

The massive technologization of culture and population has favoured the appearance of a large number of language learning apps. Nowadays, the expansion of smartphones as personal learning devices has allowed many companies to develop language learning apps that do not need the presence of a teacher figure. Without this figure, the developers had to find a way to modify and adapt the learning process without human intervention. With the collaboration of many other computer science fields, these apps use Artificial Intelligence to achieve these user adaptations.

A growing incorporation of intelligent tutoring systems (ITS) into self-developed courseware has been due to technological breakthroughs from the fields of computational linguistics, information retrieval, cognitive science, artificial intelligence, and discourse processing. ITS's are designed to simulate a human tutor's behavior and guidance. A well-designed ITS system provides users with convenient access to individualised practice, models good problem-solving approaches, and helps learners based on their learning rates and specific needs. (Liaw and English, 2017, p. 66)

These new apps and systems are the next step in Computer-assisted language learning. The whole learning experience can be customised by the AI autonomous "thinking" with the student's input. The exercises, results, and contents can be adapted to the students' needs without human intervention. The

number of tools, apps and systems that make use of AI assistance increases every day, and the computational capabilities of these AIs expand rapidly. The AIs can manage much more data in much less time, thanks to new technology that no longer relies solely on hardware. Several factors have benefited the technology for it to not solely rely on hardware and physical systems. During the first steps of AI technology, the totality of the computing process was based on computers that could take up a whole room. ENIAC, one of the first computers, was formed by “thirty separate units, power supply and forced-air cooling, that weighed over thirty tons” (Weik, 1961, p. 571). As the years passed, engineers managed to reduce computers' size and weight until now, when we can have a computer with 1300 times the processing power (AntiqueTech, 2013) but with the size of our hand, a wristwatch or even glasses. The other crucial development in AI technology is the Internet. After years of development since the 80s, Internet has become the new platform for AI computing. The new cloud technology allows users to let powerful servers do the computing work and receive the results of the “AI-thinking” in their devices. “Cloud computing has been greatly developed and applied owing to its high cost-efficiency and flexibility achieved through consolidation, in which computing, storage, and network management functions work in a centralised manner” (Ai et al., 2018). This centralisation also has experienced changes; the central, individual servers in the headquarters of some companies have been replaced with “cloud” servers, managed by bigger companies with better structures and with several servers in several locations for safety reasons. As I previously stated, this de-centralisation has helped AIs be more both accessible and affordable economically and technologically speaking. New cloud-based AIs are now able to function in small or old devices thanks to a simple internet connection, as all the processing is made remotely.

2.4 The Effects of Computer Assisted Language Learning (CALL)

The appearance of language oriented ITS was not immediate as the development of language learning was progressive. The academic community found a variety of applications to the technology that reached the general public. Since the arrival of computers and digital technology, the language learning community clearly saw their application and potential and swiftly developed tools to use technology for language learning purposes. In the first two decades of Computer-assisted

language Learning (CALL), the tools were developed following the traditional language learning techniques. “The computer learner was treated in a stimulus-response relationship due to behaviourism, and drill-and-practice remained the main method” (Li and Lan, 2021). Digital tools were used as a new environment but followed the same learning methods used with previous analogic systems. Many types of exercises that we still use today (fill-the-gaps, flashcards, repetition, etc.) were transferred into digital tools that offered a new and attractive way of learning (Otto, 2017, p. 10). Every time a new technology was implemented in the language learning environment, it was evaluated to decide whether it could mean a drastic change and improvement for the students. In the end, many of these technologies supposed a benefit to the presentation of traditional materials. As Sue E.K. Otto defends, tutors and specialists always searched for the same features in all the materials, “quality of materials, teacher engagement and training, suitability of technology to specific instructional goals, and sound pedagogical principles” (2017, p. 10). In the early 20th century, we can find the first attempts to incorporate innovative technology into the learning process. It was first focused on the oral comprehension exercises of the class as “vinyl records, film, and audio and video tapes brought native speaker voices, visages, and culture into the foreign language classroom” (Hubbard, 2017, p. 94). These exercises tried to improve students’ auditive comprehension of the target language and, by accommodating it, the improvement of the pronunciation. This sound-focussed model gained importance throughout the years, and in mid-20th century, Skinner presented his behaviourist model, which defended “behaviour as a result of its consequences.” A good or a bad consequence will determine the repetition probability of that behaviour (Catania, 1988, p. 5). This education theory made the linguistic community focus on formula repetition to learn knowledge. Most universities built language-specialised laboratories where students could listen to native recordings and practice the language (Allen, 1960, p. 355).

These were the first steps towards CALL as we know it today. Shortly after, computers started to be accessible to the general public and the educative community, so researchers saw the opportunity to use this technology in their favour. These first researchers saw in computers many of the CALL benefits we still have today.

Beneficial aspects of CALL: 1) multimodal practice with feedback, 2) individualisation in a large class, 3) pair or small group work on projects, 4) the fun factor, 5) variety in the resources available and learning styles used, 6) exploratory learning with large amounts of language data, and 7) real-life skill building in computer use. (Tafazoli and Golshan, 2014, p. 34)

Pioneers of CALL development first focused their efforts on vocabulary and grammar exercises, as these computers only had text as possible input (Otto, 2017, p. 12). Because of computers' limited processing power, early CALL programs presented the same tests and questions that were made on paper. A significant portion of the teaching community did not see in CALL the benefits that we previously mentioned, so CALL was not considered a successful model until several years later. In 1977, Apple released the Apple II, a microcomputer that allowed people to have a personal computer in their own homes due to its small size. The small CALL community of that time saw microcomputers as an opportunity thanks to the processing power of these machines, especially for graphics and images (Davies et al., 2012, p. 25). The 60s gave way to a 20-year period when many new linguistic approaches, models and theories were presented. Many of these new approaches focused on communication and interaction as one of the most important parts of language learning. (Richards and Rodgers, 2001). Many specialists saw the interaction and real communication as the best way of learning a language, refusing the previous approaches that focused on repetition and grammar practice (Otto, 2017, p. 14). It was not until the 90s that the learning paradigm changed and technology advanced enough. With it, the focus and model of exercises shifted.

The cognitive approach rejected behaviourism for language learning in the 1980s and the 1990s, although the actual paradigm shift from behaviourism to cognitivism occurred two decades earlier. During this period communicative exercises were emphasised, and fluency, rather than language analyses and grammar, was the major focus of language teaching. CALL software and language games also began to flourish during this period. (Li and Lan, 2021, p. 362)

This new approach backed successful communication as the optimal way of learning language and prioritised fluency over the analysis of the language and correct grammar. Until that moment, CALL's development was heavily restricted by technology, so its goals and functioning had to grow concurrently with technology. At the moment that the widely available technology had developed enough, the change of approach could also be done in CALL's environment. As we can see in the Table 3, CALL's types, main objectives and the use of the different platforms have changed at the same time:

Table 3.

Evolution of Computer-Assisted Language Learning

| | | | |
|------------------------------|---|---------------------------------------|---|
| Stage | 1970s–1980s: Structural CALL | 1980s–1990s: Communicative CALL | 21st Century: Integrative CALL |
| Technology | Mainframe, first microcomputers | PCs | Multimedia, Internet and mobile devices |
| English-teaching paradigm | Grammar Translation and audio-lingual | Communicate language teaching | Content-Based |
| Use of Computers | Drill and practice | Communicative exercises | Authentic discourse |
| Principal Objective | Accuracy | Fluency | Agency |

Note. Adapted from *CALL, past, present and future* (p. 15), by S. Bax, 2003 System, 31 (1).

With the new century, digital development increased, and the possibilities in the educational field widened. One of the most critical changes in the area of CALL was the substitution of “tutorial CALL” for “tool CALL”. In Tutorial CALL, “the computer must be programmed by “experts” in programming and in that subject.

The student is then tutored by the computer executing the program(s)” (Taylor, 2003, p. 243). On the other hand, in tool CALL, the digital resources are created or presented by the program, and it is the teacher who presents them to the students (Levy, 1997, p. 101). The difference in the role of the computer in both models of CALL is significant. The later way of using the digital resources was maintained by CALL for many years, a teaching model where the teacher supported the learning process with digital resources (exercises, images, audio, etc.) that exemplified or applied the content that was being learned. Later on, these two models of CALL usage were adopted by more prominent learning publishers (Tutorial CALL) and smaller institutions and teachers (Tool CALL). As Sue E.K. Otto (2017) states in her chapter about the evolution of CALL tools:

Tutorial CALL production has largely become the domain of publishers of language textbooks and of instructional language software for business, education, and home use. All language textbook publishers offer multimedia components for their books, some more elaborate than others (...) With massive resources at their disposal these companies are able to develop and sustain their products, whereas major projects funded locally by educational institutions or by federal grants are hard-pressed to do so over time. (p. 20)

Apart from tutorial CALL and tool CALL, new types of CALL have recently appeared as researchers made new improvements and discovered new applications for computers in the field of language learning. Shortly after the first developments in Artificial Intelligence and Natural Language Processing, CALL researchers applied some features of this technology to language learning software, creating “intelligent” programs that had more freedom if we compare them with other rule-based CALL systems. This “understanding” of the computer toward students’ inputs allowed the system to correct unforeseen errors that were not inside its rules or database (Heift and Vyatkina, 2017, pp. 28-29). This model of CALL has developed strongly in recent years due to AI and NLP evolution. While CALL represents an evolution in the functioning of CALL, Integrative CALL represents a shift in its goals and objectives of it. “The aim of the last phase of the CALL was to overcome the obstacles of language learning and teaching and

therefore to optimise the opportunities for integrating new technologies in the language classrooms” (Tafazoli and Golshan, 2014, p. 34). Integrative CALL's main platform relies on the Internet, where tutors and teachers can find authentic materials created by native speakers and apply them to the classes. Integrative CALL assimilates the communicative features of the previous chapter of CALL and takes a step further by emphasising its social and context-based aspects (Yamazaki, 2014, p. 3). The majority of the language learning community is now at this phase, as social interaction and gathering authentic materials through the Internet are very present in the education process. CALL has suffered many changes since its first application in the middle of the last century. CALL has been developed to the point that we can find every kind of previously mentioned CALL on our devices. While Tool CALL has seen its area of effect being restricted to schools and language centres, tutorial CALL and Integrative CALL are present in many devices and can fulfil a wide variety of users' necessities.

2.5 From CALL to MALL: The Use of Mobile Devices in Education

Language learning has always been influenced by technological advancements. The first technology-led changes were language classes recorded in storage devices that were played in personal, portable devices like Walkmans that later evolved into iPods and MP3 Players. These devices kept evolving until what we have today, small computers with global communication possibilities with high storage and processing power. Although very important, hardware is just half of language learning development. “However, it is not just—or even primarily—hardware enhancements of the iPhone generation devices that hold the most promise for use in language learning. Equally important is the software and the new opportunities that arise from mobile application development” (Godwin-Jones, 2011, p. 3). The use of mobile devices for learning a language supposes an evolution of CALL, to the point that some researchers make use of the term MALL (Mobile Assisted Language Learning) to name this new trend of language learning apps (Gangaiamaran and Pasupathi, 2017, p. 11243). The evolution from CALL to MALL has brought several benefits. According to Kukulska and Humle (2009), the four main advantages of mobile devices against traditional computers are “portability and mobility, social connectivity, context sensitivity and individuality”. Mobile devices suppose the next step in technology-assisted

language learning due to their flexibility and possibilities for both the users and the developers. Thanks to the quick and innovative changes, the educational model and way of learning with these apps have widened the possible exercises. Some of the exercises and practices that we can find in these apps are adapted from the previous analogical platforms, while others, thanks to technology, are now doable by every user in every situation. “English learning apps have the capability to integrate different media, for example, text, picture, animation, audio and video can be integrated in order to create a multimedia instructional material, as well as prompt students’ interest in studying” (Gangaiamaran and Pasupathi, 2017, p. 11247). Before CALL, some exercises that required special devices (like oral comprehension and oral expression) could only be done in the education centre where the devices needed for the exercises could be found. When personal computers became the standard, the devices needed to complete these exercises could be used in public places, schools and even at home. These days, we have seen how these computers, big, bulky machines, have been incredibly reduced in size, and their processing power has multiplied. Another significant improvement that MALL has brought is the easiness of use. As many of the users have been raised with continuous use of these kinds of devices, they are much more familiarised with technology and tutorials, instructions and preparations are much less needed (Gangaiamaran & Pasupathi, 2017, p. 11243).

As we previously stated, computers, smartphones, and similar devices have allowed the language learning process to be made in an individual, autonomous and portable way. Now, all the processes can be made remotely, and users are no longer tied to some restrictions as they were before. Although some activities can now be done remotely and without special equipment, the content of these activities is the same. Despite the substance being the same, the rest of the learning process can drastically change, which strongly affects the study's effectiveness

Even when 'old' content is served up on mobile devices, the ways in which learners use it and hence how they learn will be different, since context and location of use are associated with environmental factors such as partial attention, shifting motivation, opportunistic scheduling of study, availability of physical space, real or perceived costs to the user, social conventions of device use, and so on. (Demouy and Kukulska, 2010, p. 218)

Mobile devices are now an essential tool for every person living in the modern world. Due to this fact, much of the practice that a student can make may go unnoticed. Apart from the apps, which have exercises, clear objectives, and a clear purpose, the use of a mobile device can produce a total immersion in the goal language (Mavropoulou and Arvanitis, 2020, p. 122). Something as simple as a web search or using social media can be transformed into language learning by immersing the users in the goal language and "forcing" them into using it. Also, the use of language learning apps can cause a boost in the motivation of the students, as the apps usually present the contents in a more attractive way. Gamlo (2019), a researcher at King Abdul-Aziz University, concludes, "students had a strong motivation to learn English and were looking for every possible method to acquire high grades in the English subject. The survey results indicate that students had instrumental motivation to learn English" (p. 53).

Nonetheless, MALL also has some inconveniences that tutors have to face in order to provide successful language teaching. One of MALL's main benefits is the user's autonomy from a traditional tutor. One of the most common issues is when, with the lack of motivation, autonomy loses its function (Anggriyashati, 2020, p. 613). Another problem with MALL can be found in the physical aspect of the devices. Some researchers think that in most cases, the devices are too expensive for all the users to own or for the school to provide them (Stockwell, 2008, p. 267). Another problem with mobile devices is that they can be too small, which can cause problems for some users when writing or reading (Thornton and Houser, 2005, p. 219). Some of these problems are present in the implementation of CALL and MALL, but as in many other fields where technology is developing, benefits strongly overcome inconveniences, to the point of almost making them go unnoticed.

2.6 English Standards, Common Goals and Transversal Learning

There is no denying that mobile devices are the next step in language learning. Millions of users are learning new languages thanks to apps that bring them the possibility of learning without depending on external factors like tutors, price, place or time availability. Despite this, MALL's usage is constantly growing in the classrooms as teachers and tutors notice the benefits in opposition to more traditional education. Although the way of teaching can vary, the linguistic content is the same, as it is produced according to the most commonly accepted standards of language evaluation. In order to make these apps and materials both recognised and successful, they need to be created following the standards of the most common language frameworks of reference for each language, as these standards are the model that most of the users and institutions follow. In this case, the most common standards for English language evaluation are CEFR (Common European Framework of Reference), Cambridge, IELTS (International English Language Testing System), TOEFL (Test of English as a Foreign Language) and PTE (The Pearson Test of English). These English Assessment standards have gained international recognition, and similarities in their levels can be found among them. In our environment, the Common European Frame of Reference is taken as the common ground for the other recognitions. "The most influential parts of the CEFR are the many illustrative scales" because "without the scales, the CEFR would have been largely ignored in the European language education" (Alderson, 2007, p. 661). The European Union organisation first proposed the CEFR standard in 1996, and since then, its impact has undoubtedly affected the way languages are taught in Europe.

Most accounts of its use have emphasised its potentially positive contribution to enhancing the transparency of curricula and examinations in the different nation-states of Europe, and through that enhanced transparency, to facilitating the mobility of students and labour across the continent. (Alderson, 2007, p. 660)

Although strongly influenced by the Anglo-American world, the evaluation model that the CEFR follows covers the four main aspects that a learner tries to master

in any second language: speaking, writing, reading and speaking. The results obtained by the students in the several exams that test these four abilities give the student a position in one of the six sections of the scale (A1, A2, B1, B2, C1, C2). The popularity of this framework and the EU's effort to popularise it has made this scale a "common currency in language education, and curricula, syllabuses, textbooks, teacher training courses and not only examinations" (Alderson, 2007, p. 660) and the scales of other standards like IELTS "while not based on CEFR levels, has been linked to them" (Bezy and Settles, 2016). Thanks to these efforts made by the European language organisations, the CEFR has become one of the essential standards in language learning and the reference for many other models of learning. Thanks to CEFR's recognition, many materials are created following the scale proposed by this standard. "The use of CEFR in teaching practice and teachers' Implementing CEFR in secondary education development was addressed at numerous conferences, at which good practice and examples of teaching materials were presented" (Moonen et al., 2013, pp. 229-230). CEFR's influence has surpassed Europe's frontiers in the last years, so many other countries that do not form Europe take this recognition as a valid way of scaling the knowledge levels of a language. As Yukio Tono (2019) explains in his work about CEFR implementation in Japan:

With the growing influence of the CEFR beyond Europe, people working in foreign language teaching and learning, notably in a number of Asian countries (Japan, Vietnam etc.), have started to explore the potential of the CEFR in their fields. The most important impact of which has been made in the area of language testing. Many foreign language proficiency tests are aligned to the respective CEFR levels and claim to be mutually comparable. (p. 5)

In his project, he transferred the CEFR evaluation scale and items to the Japanese language and analysed the effectiveness of several exercises on Japanese students. He also develops several apps following CEFR's recommended items that help students learn vocabulary and grammar (Tono, 2019, pp. 13-15). This can be taken as an example of the progressive expansion of CEFR and how it is being included not only in the creation of books, text and

other traditional materials but also in the creation of modern language learning technology like apps or digital resources.

In 2001, the European Union council presented CEFR, a model that contained a list of the main competencies and abilities that any learner should have according to his level on the scale. The four abilities that the CEFR define as needed in every second language learner are “oral reception and production, written reception and production, interpretation and translation” (Martin-Monje et al., 2014. p. 2). Each of these skills has several descriptors that describe them specifically with smaller aptitudes. In a conventional class or course of a second language, all of these abilities are treated equally, and the time devoted to each one is, in most cases, equally divided. As we previously mentioned, the actual trend in language teaching is to focus on the communicative aspect of the language and on using authentic native materials as it is beneficial to learn grammatical, cultural and contextual information simultaneously (Birajdar, 2020, p. 10). On the other hand, most of CALL and MALL’s programs tend to focus solely on one of CEFR’s four main abilities. There is a wide range of language learning apps in the app market, but the majority of them just encompass one or two of the previously mentioned talents. A high number of these apps concentrate on grammar, vocabulary and memorisation tasks, despite the actual trend on communicative tasks. “They were necessarily eager to try out themselves due to various reasons, especially the lack of interaction and personal contact when learning through an app” (Puebla et al., 2022, p. 175). As we can see, there is a dichotomy between the available language learning apps and the most accepted English recognitions, specially CEFR. While apps and CALL systems tend to focalise on grammar, memorisation and vocabulary, the extended standards invite learners and teachers to consider communication and social in-context interaction as the cornerstone of language learning. Collaboration is needed between language learning app developers and the Council of Europe to simultaneously modernise CEFR certification and its applications, and to widen the approach of the apps and include social interaction and authentic material. Language learning resources, whether they are digital or traditional, should have a higher goal, an objective that states its benefit for society and the world in general. For this reason, The United Nations and its member nations adopted a series of objectives and goals that try to ensure a better future for society.

The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015, provides a shared blueprint for peace and prosperity for people and the planet, now and into the future. At its heart are the 17 Sustainable Development Goals (SDGs), which are an urgent call for action by all countries - developed and developing - in a global partnership. They recognise that ending poverty and other deprivations must go hand-in-hand with strategies that improve health and education, reduce inequality, and spur economic growth – all while tackling climate change and working to preserve our oceans and forests. (The United Nations, 2015)

All projects, no matter if made by a company, a government or an individual, should obey as many SDGs as possible. Although the responsibility of achieving the majority of these goals falls into the political powers of a country (Pradhan et al., 2017, p. 1169), private initiatives should also have some of these SDGs as objectives. Our research is not an exception, and some of the SDGs can be applied to our investigation project, so we have adopted them as goals to achieve. First, our project is destined to prove technology efficiency in language learning classrooms, so educational systems and their process and goals are a very important part of our research. According to the United Nations (2022), 50% of the world's primary schools lack computers and Internet connections. This hampers the schools from obtaining actualised materials, new teaching techniques and modern activities for students to learn. If a school manages to have all the basic necessities covered, computers or devices for students should be the next priority. This project can be included inside SDG number 4, as it can “ensure inclusive and equitable quality education and promote lifelong learning opportunities for all “(The United Nations, 2022). The use of technologies like the one proposed in our work brings the possibility of remote learning and teaching, which has been the norm during last year's pandemic (Hodges et al., 2020). The use of technology brings teachers and students the opportunity to benefit from quality education systems, even in difficult situations or remote locations. Apart from the accessibility benefits, schools can retrieve new and modern resources, and actualised techniques and boost students' motivation with the use of technologies and educational games. Our proposal of the use of technology can

also side with SDG 9. This SDG encourages countries to “build resilient infrastructure, promote inclusive and sustainable industrialisation and foster innovation“ (The United Nations, 2015). We support the use of innovative technologies like Artificial Intelligence. The development of these kinds of technologies can make the developed countries invest in less developed countries and robustly strengthen these countries’ research development and innovation (RDI) (Cooper, 1972, p. 3). Innovation can substantially affect every aspect of society, and technology should be a priority for governments once basic needs and industry are established. Lastly, the use of electronic devices can significantly reduce the use of traditional materials. SDG number 12 states, “Ensure sustainable consumption and production patterns “(The United Nations, 2015). Many schools still support their education in non-digital platforms like textbooks and sheets of paper, a valuable resource due to the decreasing quantity of raw material needed. According to David Wees, a regular school in the United States can use around 2000 sheets of paper a day. With an average price of 5 cents per photocopied paper, it sums up to 2 billion dollars a year per schooled child.

For comparison, providing each child in the US with a \$100 laptop (recently available on the market) would cost about 8 billion dollars assuming even the little kindergarten children get one. In other words, we could pay for a laptop per child in the US in 4 years by stopping using paper in schools. (Wees, 2009, p. 5)

Although difficult due to the availability of materials and cost of the devices, a progressive transition from paper to digital platforms could suppose a significant reduction in the use and waste of paper. Digital resources can be updated each year without the need to print another edition and are more accessible for communities in remote locations. Computers and mobile devices can be difficult to achieve or buy for some teaching communities, but this can be easily solved, as only 22.8% of electronic waste is correctly collected and recycled (United Nations, 2022). The use of technology in education can contribute to an improvement of the ecological situation in our countries, benefiting from the digital media and recycling processes (The United Nations, 2022).

Apart from the global objectives that the United Nations propose, any project, research or initiative related to education should commit to certain competencies that students can learn within the education process.

General or transversal competences, which can be defined as a set of competences that can be applied in any professional situation or task, regardless of where they were attained. Thus, general competences are required for all types of jobs and are the basis for the attainment of more specific or technical competences. These competences are transversal and transferable to different contexts. (Sá and Serpa, 2018, p.13)

Transversal competencies (TVCs) were first defined in the Tuning Educational structure project. This educational project was presented in the year 2000, and the European countries that were members of the Bologna process also signed the creation of this project. The Tuning's members defined the TVCs as "a dynamic combination of cognitive and meta-cognitive skills, knowledge and understanding, interpersonal, intellectual and practical skills and ethical values" (Gonzalez and Wagenaar, 2008, p. 9). The TVC are specific skills that are highly valued in the job market and that every person should learn throughout their education life. This meeting of around 30 European countries pursued the unification of the Higher Education systems in Europe and promoted common grounds for mobility, an equal system of credits and transparency (van de Wende, 2000, p. 5). Although based on United States' model, the representatives of the attending countries established a number of objectives which every Higher Education institution should try to achieve. For this reason, every Higher Education Institution in Europe, especially universities, has defined several TVCs to be applied in all their projects and initiatives. Depending on the project, some of them would not be fulfilled, but the author should plan the project thoroughly to achieve as many as possible. In this case, the Polytechnic University of Valencia defined and implemented the last version of its TVCs in 2014. Initially, the UPV defined 13 competences that ought to be included in all its courses and projects. During this last academic year, the UPV's competences have been revised and reduced to 5. These 5 new Transversal competences have been revised and approved by the University's governing board and are the following: "Social and

Environmental commitment, innovation and creativity, teamwork and leadership and responsibility and decision making” (Secretaria General UPV, 2022, pp. 3-6). Although reduced in number, this new set of TVCs contain all the skills described in the previous versions and are adapted to the current year and society. This set of skills should be attainable by any student, no matter their level, specialisation or duration of the course. These objectives cover a wide range of effects and consider important aptitudes that are crucial for the student, not only during his/her educative stage but also when entering work life.

The first Transversal Competencies is social and environmental commitment. All students should learn how to act ethically towards social issues and environmental concerns and to be aware of the diversity and differences that exist in our world. Also, students should have the knowledge to create, develop and collaborate on solutions that try to fix these problems, always taking into account the previously mentioned SDGs proposed by the United Nations. The second TVC makes reference to the need for innovation and creativity. All students should be able to “propose creative and innovative solutions to complex situations or problems, specific to the field of knowledge, in order to respond to diverse professional and social needs” (BOUPV, 2022, p.3). This experiment supposes an exploration of innovative techniques in the field of language learning, and the students make use of an example of technological advance in the field of language learning as the language-oriented chatbot Andy. The third competence establishes that all students should be able to work in a team and to adopt leadership positions to achieve common goals via collaboration. This section also reinforces the importance of multidisciplinary teams and the possible enrichment of different kinds of knowledge in the same project. Students must act with empathy towards the other team members and comply with all the opinions, points of view and ideas in the collaborative work. The next Transversal competence makes reference to the student's ability to communicate efficiently and always adapt to the context and the necessities of the interlocutor or the audience. This successful communication must be achieved both in verbal and non-verbal communication, must be well structured for easier comprehension, and all the produced texts should be created according to the current scientific and socially accepted standards. Also, students should have the capacity to communicate and use digital platforms accordingly. In our teaching proposal, a

mobile device, a platform that students are highly familiarised with, is given a new purpose. Students learn to get on with a digital platform and to use it according to the goal of learning the English language. Furthermore, students can change their perception of technology and specifically mobile devices, as the app, they are using presents a different way of learning a language that does not relies on traditional exercises but on casual conversation practice. The use of technology in our research can motivate students to find new and different applications for technology in the education field and to develop new techniques that can be applied to everyday technology. Lastly, students should act with responsibility and improve their decision-making inside and outside the learning environment. Acting autonomously is very important in many aspects of life, and the use of this app can teach students to self-carry their learning process without the constant supervision of a teacher or tutor. Knowledge extracted from solving problems can be applied to other situations, and students should also learn to manage this capacity and adaptability. This knowledge extraction should be accompanied by critical thinking, effective bibliographic research and reliable source usage.

All these Transversal Competencies are expected to be learned in a higher or lower grade by every student as they are useful abilities considered both valuable and useful for work life and life in general. Analysing our investigation, we can see that the three TVCs applied are successful digital communication, development and use of innovative solutions to existing problems and the improvement of an autonomous work ethic.

2.7. Artificial conversations: the use of chatbots in language learning

The teaching environment has suffered many changes in the last years, as many new technologies are tested constantly in search of better systems, techniques and results. Previously, we have mentioned many types of these new technologies and the application that the language learning community has found for them. Innovative techniques like Artificial Intelligence, Intelligent Tutoring Systems, Image and text recognition, Machine learning and Chatbots are some of the technologies that we can find today in computer software, apps and online platforms. In our case, we research the use of chatbots in language learning education. Chatbots are commonly extended systems in today's technological world. Combining Artificial Intelligence, machine learning, and text recognition

serve the purpose of establishing a human-like conversation in order to solve a problem or achieve a goal (Misischia et al., 2022, p. 422). Present applications for this technology may vary, from customer service to medical support. Each day the usage of chatbots in many fields extends, and their application in education is starting to be seen as a viable option. One example is the use of chatbots for custom support in large, multitudinous, online lectures. Without the possibility of having one teacher per student or a single lecturer helping more than 100 students, chatbots can solve this situation by helping students with a customised solution to their lecture-related issues. (Winkler and Söllner, 2018, p. 2). If we narrow down chatbot application to language learning, its definition changes:

Chatbot-supported language learning refers to the use of a chatbot to interact with students using natural language for daily language practice (e.g., conversation practice), answering language learning questions (e.g., storybook reading) and conducting assessment and providing feedback (e.g., vocabulary test). (Huang et al., 2022, p. 238)

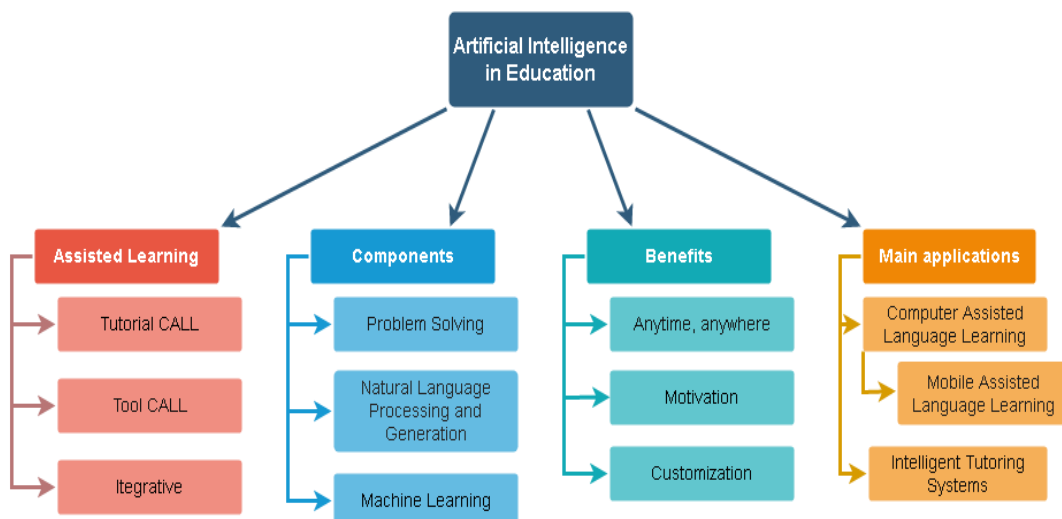
All these functions can be comparable to the ones performed by a human tutor, so this is one of the reasons why chatbots are forging their own way into language learning. Chatbots also bring several benefits that we have previously mentioned in relation to ITS and learning apps, like remote learning or 24hr availability. Fryer and Carpenter (2006) present new positive aspects of the use of this technology, like the reduction of “lack of confidence, shyness, and unchecked mistakes in grammar and pronunciation” (2006, p. 8). Fryer (2019) finds even more benefits like “a broad range of expressions/questions and vocabulary” and the ability to “keep on talking and enable repetitive practice” (2019, p. 29), which a human interlocutor can get tired of. Even with all these benefits, language-oriented chatbots still have some issues that need to be addressed. In order for a chatbot to perform correctly, it needs to be advanced enough because if it is not the case, a chatbot can answer incorrectly if the student’s input is wrong or if the topic is not accepted (Huang et al., 2022, p. 238). Chatbots are becoming more and more used in learning contexts, and the development of these systems is allowing both teachers and students to benefit from their features and capabilities.

For these reasons, we have chosen the app Andy, created by developer Andrey Pyankov in 2016. This app is available both for Android and iOS with a trial version free of charge and a premium option. Andy is an English teaching chatbot with Artificial Intelligence and customisation processes. This app was designed to teach English through the use of a chatbot that simulates “human-to-human” casual conversation. Apart from the chatbot conversation, users can also complete traditional exercises made through the same chat interface. The utilisation of this type of technology in a language learning class is still very unexplored, and we thought that more research is needed in order to extend its usage and benefits in more language learning institutions. For this investigation, both teachers and students were questioned about the use of technology in high school English classes. Then, students completed a general English questionnaire and used the app to complete several exercises. After the practice, the teacher and the students completed a satisfaction questionnaire regarding the results of the exercises and the use of the app. We will analyse all the materials, questionnaires and results further on to discuss whether the use of chatbots in an English class is viable.

Artificial Intelligence has been present in the educational field for some time now. Despite this, many new systems, apps and approaches are developed continuously to keep up with the technological advances of our time. In Figure 2 we can see a scheme representing the use of AI in education and its advances until the moment.

Figure 2.

Artificial Intelligence in Education



3. METHODOLOGY OF THE RESEARCH

In this section we will present the methodology used for this project. This section includes all the previous work needed before the experiment, the different options considered and the variety of tools that were used in the previous work, the experiment and the analysis of the results.

3.1 Chatbot App: Andy English Speaking Bot

The language learning method is constantly suffering an evolution that change all its features. Materials, processes, approaches and platforms are updated according to the community's needs and following technological development. We have chosen to research the functionality of AI — powered chatbots in a class of English as a second language (ESL) learners in a school. Currently, there are several AI — powered language learning tools in the market, and we had to analyse several of them in order to fulfil certain requirements. First, the program must be widely available in our country and under a free — use license. Many options had to be paid for from the start, or the free trial version was too limited for us to use in the research. Next, another necessary requirement was that the app had to be relatively modern and recently updated. According to Pixalate's study (2022), "over 1.5 million apps between the Google Play Store and Apple App Store appear to have been 'abandoned,' meaning that they haven't been updated in over two years" (Pixalate, 2022, p. 1). If the app was abandoned, the AI might work incorrectly, or the topics may be outdated. The use of a mobile device in the experiment was not mandatory, but the availability of this kind of app is far more significant in the personal devices' market". Mobile technology provides various resources and tools for language learning that encourage learners to be more motivated, autonomous, situated (site-specific), and socially interactive" (Kim and Kwon, 2012, p. 35). Another needed feature was the possibility to complete exercises about topics and materials treated by the research group. These exercises had a result system that could be individually registered and metered. Lastly, the AI app had to be attractive enough for the students to use it and to present a work method viable for the teachers to reuse in the future or for the students to operate independently. We had several possible options available in the market. We analysed the features of apps like Read My World, an image recognition app focused on vocabulary, ELSA Speak,

a pronunciation tutor driven by AI and Duolingo, the most popular multipurpose language learning app in the market. In the Table 4 we can see a comparison of the possible apps and their main features.

Table 4.

Possible apps and their main features

| App | Main purpose | Platform | Price |
|---------------|-----------------------|--------------|---------------|
| Read My World | Vocabulary | Android, iOS | Not available |
| ELSA Speak | Pronunciation | Android, iOS | Free version |
| Duolingo | Multi-purpose | Android, iOS | Free |
| ANDY | Chatbot and exercises | Android, iOS | Free version |

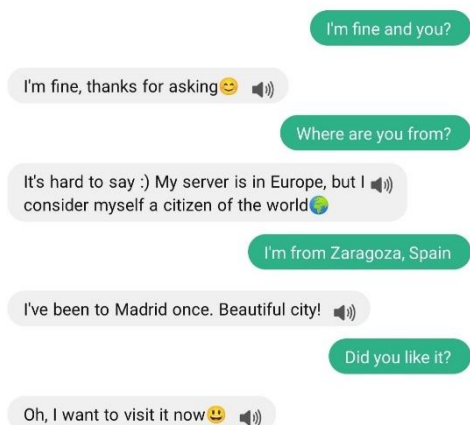
After long research on the usable options, we decided to use the Andy app. Andy is a language — oriented AI chatbot that allows users to practice English through traditional exercises and casual written conversation, similar to texting apps.

Andy English Bot was created in 2016 by Russian developer Andrey Piankov. This app is available for all mobile devices, both Android and Apple, in the app market. The app allows the user to use the chatbot and some exercises with the free version and offers more options like more grammar lessons, more vocabulary or more chatbot topics with monthly, yearly or permanent payment options. As we will see later on, the free version restricts the level of English that can be practiced. Once we start the app, it presents the user with its possibilities without the need to log in or create an account. Another reason we chose it is the opportunity to use it without making any kind of account. The app can be used the moment it is installed, which significantly eases its use in the classroom. ANDY app presents four sections, chat, exercise, grammar and stats, with a straightforward design that facilitates its use. It shows four main areas, three are destined for different types of exercise, and the last one contains statistics about the user and how the app is used. All the user — app interaction is made through the app. When the user wants to do an exercise, the app will automatically change to the chat section, and the practice is presented by the chat interlocutor as if a human interlocutor made it via chat. Additionally, as we can see in Image

3, each message sent by the bot has a speaker button, and when pressed, the message will be read through the speakers. This extra feature can help students check the pronunciation of a specific word or phrase or help people with reading difficulties understand the conversation.

Image 3.

Chat interface in ANDY and speaker function



Once the exercise has been selected, the chatbot will ask questions and present exercises that the student must answer by tapping one of the options that the chatbot presents. The option is sent as a message for the system to recognise, and the chatbot gives the consequent answer in a new message. The total integration of the exercises in the chat interface can provide the students with a more profound conversation sensation while screening the activities and the learning process. As mentioned in section 2.3, ITSs and Machine learning — capable systems allow the program to customise its functioning to the user's needs and results. In this app, the system will record the user's results and used words and create a list of “known words” to use in future exercises. When a user asks about a word or makes a vocabulary exercise, the chatbot will send a message that says, “Cool, I've added it to your KNOWN words list 😊” (ANDY, 2021). This customisation allows the system and the user to follow the learning progress and motivate the use of the app with motivational messages after several successfully answered questions (See Annex 3). Although the paid version of the app gave many more options, the free version was enough for us to try it on our experiment, and the capabilities of the free version were similar to

the ones of the paid version. Hence, the quality of the results is similar. As we can see in Table 5, this app presents many benefits and improvements over against more traditional language learning methods. These improvements help students enhance their learning process and impulse a better knowledge acquisition.

Table 5.

Features of ANDY, English Speaking bot

| Accessibility features | Linguistic features | Customization |
|--|---|---|
| Free version, no account needed | Variety of lessons and exercises | Custom vocabulary lessons and lists |
| Voice-over messages | Simulated conversation | Custom chatbot topic |
| Visual support with emojis and motivational messages | Improvement of written comprehension and production | Statistics about the use of app for each user |

3.2 Control Group and Previous Work

For our investigation, we want to test whether an AI chatbot is a suitable support tool for an English teacher in a Spanish high school. For that reason, our research is mixed. The quantitative aspect of the investigation analyses the effectiveness and the data regarding the use of language in a research group made up of three high school classes. This type of research will give a numeral value to the results similar to the ones used in schools to classify student's performance. Also, quantitative research allows us to find trends and singularities in the results we obtain from the different research sources. On the other hand, the qualitative aspect analyses the results and opinions of the app given by the teachers and students. Qualitative results are equally important due to the public — oriented nature of the app and its necessity of users' experiences to improve certain features. "Qualitative research contributes to an understanding of the human condition in different contexts and of a perceived situation" (Bengston, 2016, p. 8). The interpretation of the evidence from these two types of research would provide us with the suitability of a chatbot app as a support tool for an English teacher in the class.

In order to have reliable data about the app's usage, we needed to find a group of students that were familiarised with the use of technology and were learning English as a second language. I contacted the English teaching department in the high school Teresiano del Pilar in Zaragoza, Spain, where I attended. After presenting the experiment, explaining its steps and goals and expected results, the department teachers agreed to conduct the research in three classrooms of 4º ESO (15 — 16 years old). The reasons to choose this level were the availability of the students, as higher levels were no longer attending class, schedule compatibility and these classes' English level was similar to the app. The study was conducted on June 16th, 2022, and 37 students participated. The number of students was not very high. Still, it allowed us to analyse the individual results and contrast them with the variable English levels in the group.

The first step in our research was to talk with the teacher, that in this case is the same for the three classes and question her about the average level of the classes according to CEFR level, what was her opinion about the use of technology in English class and how she implemented technological tools in the class. As we expect the app to work as a complementary tool for the English teacher, she was also questioned about the English content learned by the students during the year, so the exercises on the app treated this content and acted as a support, not as a teaching platform for new content. All this information was gathered through a questionnaire sent to the teacher days before the experiment. Some of the answers to the questionnaire (see Annexe 1 for complete questionnaire) gave us the information necessary for the experiment's development. First, the teacher set the average CEFR level of the three classes at a B1, which confirmed the use of the app in this group, as the level of the app is similar to B1. We will discuss the English level and competencies of the app and the students further on. The teacher was asked about the types of exercises most used in class, and the answer was: "Textbook exercises (fill the gaps, rephrasing), exercise sheets with similar grammar practice and true or false / short answer listening exercises" (Sánchez, personal communication, June 15, 2022). Thanks to this previous practice, students did not have any problem doing the test, as the app's exercises were similar to the ones used by the teacher in class. Also, the teacher thinks that the text writing exercises are the least practised in class, so the app can also reinforce this aspect by making students

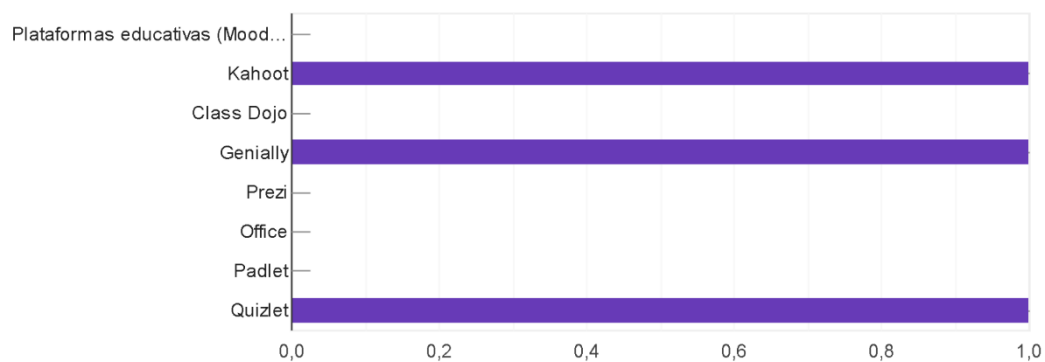
write short texts and conversation fragments through the app and pay attention to grammar and orthography. It is also significant that the teacher valued the use of technology in class as a five out of ten. Related to this, the questionnaire presented a list of the most common non — language — specific education applications, and the teacher marked those that have been used in the class. Those used in class are shown in Graphic 1.

Graphic 1.

The use of educational apps in the classroom

12. ¿Que aplicaciones y herramientas de la siguiente lista usa o ha usado alguna vez en la clase?

1 respuesta



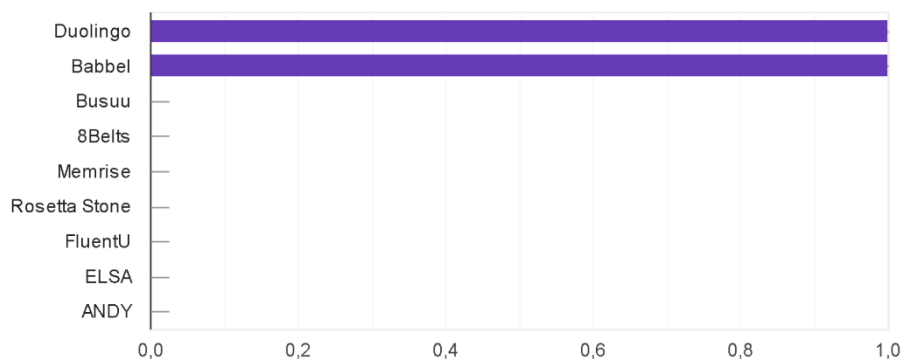
As shown in Graphic 1, only three of the eight presented apps have been used in the language classes. When comparing and analysing these three apps, they all have multimedia usage and interactivity as the primary focus. These apps allow the students to collaborate or compete in materials that the teacher previously created. The next question, also related to the use of apps, presented the most popular language learning apps and the teacher had to mark those that she knew or used in the past. As we can see in Graphic 2, the teacher knows only two of all the presented apps.

Graphic 2.

Available language learning apps.

13. ¿Cuales de las siguientes aplicaciones de aprendizaje de idiomas conoce?

1 respuesta



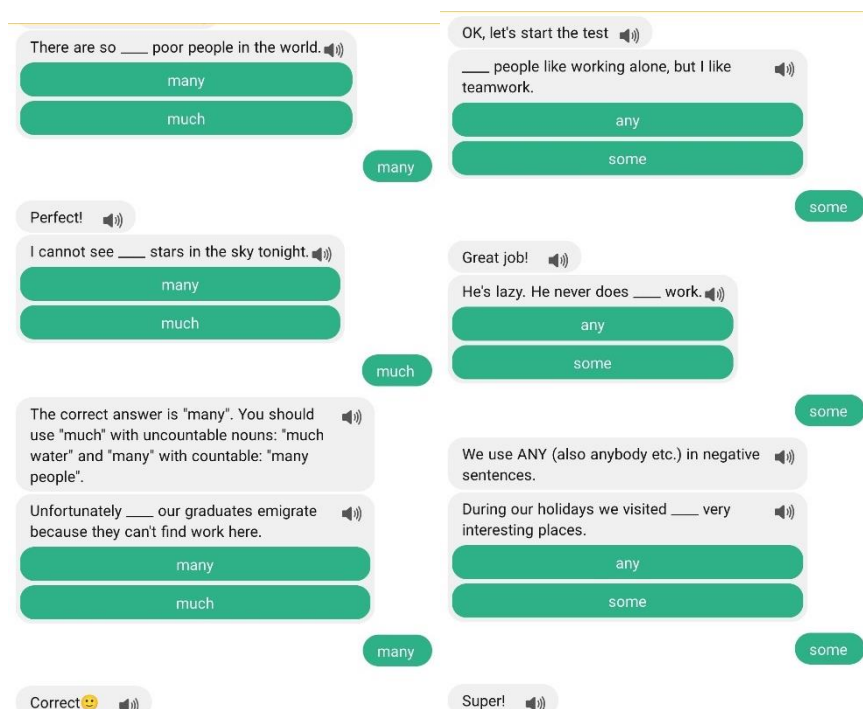
This can indicate that this class still strongly relies on physical materials and traditional teaching methods. Also, we can suppose that this situation is general to the Spanish English classrooms, as all schools must follow the current education law, the *Ley Organica de modificación de la LOE Dec 3, 2020*.

After analysing the results of this initial questionnaire, we had to search for a series of exercises available on the app that coincided with content that had been taught to the student early that year. In our case, we propose using the ANDY App as a support tool for the teacher. The chatbot app complements the teacher's classes following the Tool CALL model. Unlike Tutorial CALL, the teacher is not substituted by the application but used as a tool that reinforces the students' learning process. For this reason, we looked into the materials in the textbook used by the class that year to find a grammatical lesson that could be treated in the app. Moreover, the lesson had to be difficult enough to test the students' knowledge and for the variation of results in the test to be noticeable. After discarding some of the lessons for various reasons, we finally chose the lesson about the quantifiers *some* and *many* and *much* and *many*. Through its exercises, Andy allowed us to test the students' knowledge regarding these quantifiers that they were familiar with. The chosen exercises presented a familiar topic in an also familiar style of exercise. The main difference was the platform the students

completed them through. As we can see in Image 4 the exercises proposed by the app are similar to the ones practiced in class via traditional methods and the students will be completely familiar with these and will need no explanation to complete them.

Image 4.

Exercises in the app



As can be observed, students complete the exercises, and the app gives the feedback of the sentence immediately. If the students give the wrong answer, the chat shows a message that repeats the explanation given at the start of the test.

3.3 Experiment: Previous Knowledge, Use of Andy and Satisfaction Questionnaire

One of the main goals of this investigation is to adapt the use of AI technology to an English class. For this reason, the whole experiment should not last longer than 1 hour. In Spanish schools, classes last 1 hour, so our experiment had to adapt to this timetable not to damage students' motivation and to prove that this type of technology can be integrated into regular classes without altering their duration or functioning. After choosing the exercises in the app, we needed to

test the previous level of the students regarding the grammatical content we decided. In order to check the effectiveness of the app, students' knowledge of the topic had to be checked before experimenting with the app, so after the exercises in the app, we can prove whether it helped to learn the contents. After checking the school's availability of mobile devices, the best option is for the students to use their own mobile devices. As previously mentioned, mobile device ownership is now widely extended, and most students are familiar with its use. As Sun and Gao (2020) propose in their work “. The increasing use of mobile devices in daily lives and the fast development of mobile technologies decreases the traditional technical difficulties in MALL” (2020, p. 1192) Apart from this, using personal mobile devices in the classroom can boost students' motivation and teach them alternative smartphone uses (Elaish et al., 2019, p. 13327). As every student carries their own device, we created a PowerPoint presentation with QR codes, created with the tool QR.io, for the students to access different sections. This presentation is projected on the classroom screen, and when scanning the QR code, the mobile device automatically opens the questionnaires or the app market on their device, as we can see in Image 5. This solves the problem of accessing the different sections for the students and manually installing the app on the school's devices. (See Annex 5 for full presentation)

Image 5.

Example of QR slide to access app market.

Android & iOS



After a short introduction to the project, students' first step was to complete a previous knowledge questionnaire. This questionnaire, created and completed with Google Forms tool, is similar to an exam whose main objective is to test students' knowledge about the quantifiers some, any, much and many. This questionnaire presents a traditional — style exercise of a sentence with a gap in the position of the quantifier. The student is given two options and must choose the right one. In Image 6 we can see two of the questions of the questionnaire. (See Annex 1 for full questionnaire)

Image 6.

Types of exercise in the previous knowledge test

| | |
|--|---|
| My family owns ___ properties in Italy | Please go outside, you can't stay here ___ longer |
| <input type="radio"/> Many | <input type="radio"/> Any |
| <input type="radio"/> Much | <input type="radio"/> Some |

The result is given to the student immediately after every response, and after finishing a total of eight questions, the total amount of correctly answered questions is shown. We will discuss the results of this questionnaire in the Results section and how the outcome of these exercises can affect the use of the app or how it can vary after using the app.

After finishing this previous questionnaire, a brief explanation of the app's functioning and features was given to the students. Then, students were asked to scan the QR that allowed them to install the app, Andy. As we previously mentioned, the app does not need an account to work, so the students were able to use the chatbot as soon as the app got installed on their devices. Each student started chatting with the chatbot, simulating a real conversation. Although all conversations with the chatbot started similarly, students introduced themselves, stated the country they lived in and similar topics, many of the topics changed quickly to the ones chosen by the students. Some students asked the bot about its origin and ability to know English, others chose sports as a conversation topic, and another group of students decided to talk about tourism and travel experiences. All the students used typing as the primary form of input for the app,

even though they were informed that the app supported voice input to write the message. Only ten students used the voice feature to hear the messages the chatbot produced. Although the voice feature was explained and present from the start of the test, only a reduced number of students made use of it. One of the reasons for these maybe is that earphones or similar devices were needed to listen the voice carefully and not disturb other students with the device's high volume. The students that used the voice feature mainly used it for the instructions for the exercise and some of the questions. Students were allowed to chat with the app for 10 minutes after completing the exercises. After 10 minutes of individual, autonomous conversation, the group was asked to change into the "Grammar" tab inside the application and complete the two available exercises. The reason behind the time destined for autonomous conversation was to mix the timing of an oral conversation exercise and a writing exercise. Oral conversation exercises done in class tend to last from 3 to 5 minutes, while writing exercises of more complex texts can last up to 20 minutes. After discussing with the teacher, we agreed on giving the students 10 minutes to chat with the bot, as it is a middle ground between the timing of the two types of exercises and it allows students to understand how to use the app and talk about several topics. Then, the students had 15 min to complete the exercises in the app. These exercises are similar to the ones conducted in the previous questionnaire. The tests presented by the app are fill - in – the - gap sentences where the student chooses the right option between the two options given by the app. As we can see in Image 7, all these exercises are completed through the chat. The chatbot gives a short explanation of the two possibilities and gives instant feedback depending on the student's response.

Image 7.

Quantifiers exercises through the chat

The screenshot displays a chat interface with the following content:

- Lesson Header:** "Some - Any" lesson please
- Instruction:** In general we use SOME (also somebody/someone/something) in positive sentences and ANY (also anybody etc.) in negative sentences and questions.
- Example:** We sold SOME books. We didn't sell ANY books. Did we sell ANY books?
- Exercise:** There are so ___ poor people in the world. (Answers: many, much)
- Lesson Header:** "Much - Many" lesson please
- Instruction:** OK. The rule is very simple: you should use "many" with countable nouns: "many cats". And "much" with uncountable nouns: "much water". You can't count water, but can count cats: 1 cat, 2 cats, 3 cats, etc. Do you understand?
- Example:** I cannot see ___ stars in the sky tonight. (Answers: many, much)
- Correction:** The correct answer is "many". You should use "much" with uncountable nouns: "much water" and "many" with countable: "many people".
- Exercise:** Unfortunately ___ our graduates emigrate because they can't find work here. (Answers: many, much)
- Feedback:** Perfect!
- Final Exercise:** There are so ___ poor people in the world. (Answers: many, much)
- Feedback:** Correct 😊

Students had approximately 15 minutes to complete all the exercises. Once they finished, they could keep chatting and interacting with the app until all their class colleagues finished. During the whole test, students were allowed to ask or to look for any unknown vocabulary that could appear in the exercises. When finished, we asked some of the students about the results of the app's exercises. Once all the students had finished, we asked the students to fill out one last questionnaire. This last phase of the experiment focuses on the impressions and opinions of the students toward the app. With the results of this questionnaire, we can determine whether the students are comfortable using an app with these characteristics or if this app and similar ones, are ready to be applied in an English language learning context. The results of this last questionnaire will be discussed further, in the section Results. At the same time, we also asked the teacher to answer some questions about the experiment and the results of the students. These questions aimed to get the teacher's perspective on the use of AI powered apps in the classroom environment.

4. DIDACTIC PROJECT

This investigation tries to apply a language learning tool into an ESL classroom, and, as such we need to create an educational project that follows a set of guidelines, goals and objectives. This section presents the goals and educational structure behind the use of the app in an English classroom.

4.1 Test Group and Level of The Experiment

The main objective of this research project is to analyse the viability of applying an AI-powered chatbot app in an English language learning class. This mixed experiment, both quantitative and qualitative, tries to analyse the number of students that successfully used this technology and their opinions on the use of AI apps, and innovative technology in a class context. To achieve this, we conducted an experiment that simulated the use of the app in a regular lesson as a supportive tool for the teacher. In this section, we will compare and discuss the evaluation process and learning goals of our experiment in relation to the learning guide of the school, which is created following the valid education law.

The experiment was conducted in three classrooms of 4th ESO in the school Teresiano del Pilar, in Zaragoza. The number of students who participated was 37. In the previous questionnaire to the teacher, we asked her to set an average level for the three classrooms according to the Common European Framework of Reference (CEFR) and the answer was an average of B1. Although the difference and variation in level among the students are inevitable, the materials, exercises, and evaluation models are generally learned and achieved by all the students in order to pass the subject. According to CEFR descriptors, students with a level equivalent to B1 should be able to:

Understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans. (Council of Europe, 2001, p. 5)

As we can see, the competencies expected of a B1 student are focused on the communicative aspect of the language and the topics to be treated are familiar themes. The student experiences these situations in daily life so the knowledge about them is deep enough to understand them in English. These competencies described by the Council of Europe, serve as a reference for the schools, institutions, and governments to create the learning policies that set the objectives of the students at each level.

4.2 Evaluation Items and Objectives of The School

Schools design the subjects' evaluation items and competencies according to the current education law. In Spain, each autonomous region is responsible for the creation of its own laws regarding education. Educative laws are created by each region following the general law proposed by the state, which sets the basis and guidelines (Digón, 2003, p. 5). Henceforth, the school where we conducted our investigation bases its learning objectives on the *order ECD/489/2016*. In this law, we can find four main sections that encompass all the competencies previously mentioned in the B1 descriptors. These four main sections are "oral texts comprehension, oral text productions: expression and interaction, written text comprehension and written text production" (Order ECD/489/2016). All the exercises, materials and tests made for the students must be created according to one or more of the descriptors inside these categories. Each category is further divided into more sections that describe the individual competencies of linguistic proficiency, like the use of "subordinate clauses", "adverbs" or "collocations" (Order ECD/489/2016) and the previously described transversal competencies that can be acquired by following this evaluation model. The main TVCs that this document describes are "comprehension ability and strategy, sociocultural and sociolinguistic aspects and communicative functions" (Order ECD/489/2016). In the case of English, the law establishes a series of objectives that are classified into four main sections, which at the same time are divided into two more sections each (TVCs and linguistic competencies) that contain specific descriptors for abilities and evaluation items that should be obtained by the student. If we read these specific descriptors, we find that many of these can be applied to the use of the app we chose in class.

4.3 The Experiments Goals and Objectives

When developing this investigation, we had to set a series of objectives for the students to achieve when using the app. We previously described (see 2.6 section) the Transversal Competencies that students can master through the use of this app. Apart from TVCs, we set linguistic objectives that proved the English-learning capabilities of the app. The main objectives are learning the use of the quantifiers (much, many, some and any), improving the written production and comprehension of short texts and messages, reducing grammatical errors, discovering new vocabulary and enhancing the use of technological advancements and previous technological knowledge, like an AI app. As we can see in Table 6, we can compare the school's objectives with ours and see the similarities.

Table 6.

School's objectives and experiment's objectives.

| School's objective (Order ECD/489/2016) | App's objective |
|---|--|
| "Quantifiers: (a) few, (a) little, a lot of, lots of, plenty of, several, etc." | Use of quantifiers: much, many, some, any |
| "Comprehension of general and specific information in authentic or adapted texts, written in a language of habitual use." "Composition of creative written texts on current topics or topics of personal interest" | Improving the written production and comprehension of short texts and messages. |
| "Use of all punctuation marks and conventions of use" | Reducing grammatical errors |
| "Use words of similar meaning; define or paraphrase a term or expression." "Locate and use linguistic or thematic resources appropriately." | Discovering new vocabulary |
| "Building on and making the most of previous knowledge." "Understanding basic digital language" Locate and use linguistic or thematic resources appropriately." | Enhancing the use of technological advancements and previous technological knowledge |

Looking at this table, we can see that the app's objectives and the ones proposed by the educational law, which the school must follow, are aligned. By comparing the goals, we suggest for our experiment with the goals that the teacher tries to achieve with the students, we prove that implementing this technology in an English class is not purely technological but also follows the approved models of education and teaching.

4.4 Andy as a Conversational Auxiliary Teacher

We suggest the usage of the app in an English classroom as a tool for the teacher. The teacher can use the app to assist the traditional teaching methods and support the theory materials and exercises found in textbooks and similar

media. Also, teachers can benefit from using this application by “allowing” the app to reach “places” the teacher cannot. The app can provide the teacher with the ability to give each student a personal, conversational companion, which is highly customisable and available at all times. Currently, many Spanish schools have an auxiliary English teacher whose primary purpose is to serve as “corrective models in L2 at the phonemic, grammatical, lexical, semantic, interactional or register level” (Sánchez, 2014, p. 78). Usually, one or two auxiliary teachers are not enough for one or more English classes of different levels, and their contracts are typically temporary, so this app can supply the teacher with conversation exercises and reinforce the past lessons through conversation. Simultaneously, the use of the app can support all the aspects we previously mentioned, like vocabulary, correct grammar and written comprehension and production, all this through the use of casual chatting. As we have described, both the chatbot and the exercises combine conversation, grammar and writing knowledge. For the grammar exercises, it is easy for us to register the results, as we do in the Results section, but the interaction with the chatbot is harder to register. Chatbot’s interaction and results cannot be automatically registered and the only way for us to obtain the messages is to ask the students to provide images of the interaction. For this reason, we value student’s opinions and suggestions as the evaluation for the chatbot interaction between AI and user. Our investigation’s main aim is to test the suitability of use of the Andy app, in an English learning classroom. We established a series of objectives that follow the law and the school’s objectives and competencies so it can be applied legally and in line with the rest of the lessons and materials of the course. Also, we consider the use of this app as a tool and complement for the teacher, not as a substitution. Andy can take the role of conversational companion giving the students huge customization possibilities and immediate-constant availability.

5. RESULTS

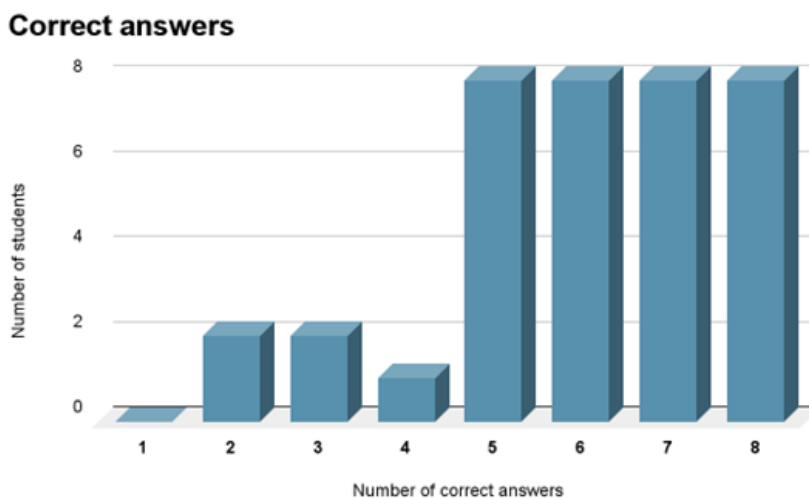
Our research tries to find the features and benefits of a chatbot app, powered by Artificial Intelligence, in order to apply it in a learning environment. We have subjected the research group to 3 questionnaires and tests and in this section, we analyse the results to discuss them further on. By analysing the quantitative and qualitative results we will be able to analyse them and confirm our objectives and answer our research questions.

5.1 Students' Previous Knowledge Questionnaire Results

Before asking the students to use the app, we need to register the students' level and, after analysing the results, check if the students' knowledge has been affected by the use of the app and if the application of technology improved the results. In this previous questionnaire, students had to complete exercises similar to the ones presented in the app. The lesson, topic and style of the exercise are familiar to the students as they have been treated previously during the scholar year. The results of this previous test show us a variety in the class level concerning this kind of exercise. From a total of 37 replies, the average score is six out of eight. As we can see in Graphic 3, the number below and above the average score is similar.

Graphic 3.

Number of correct answers.



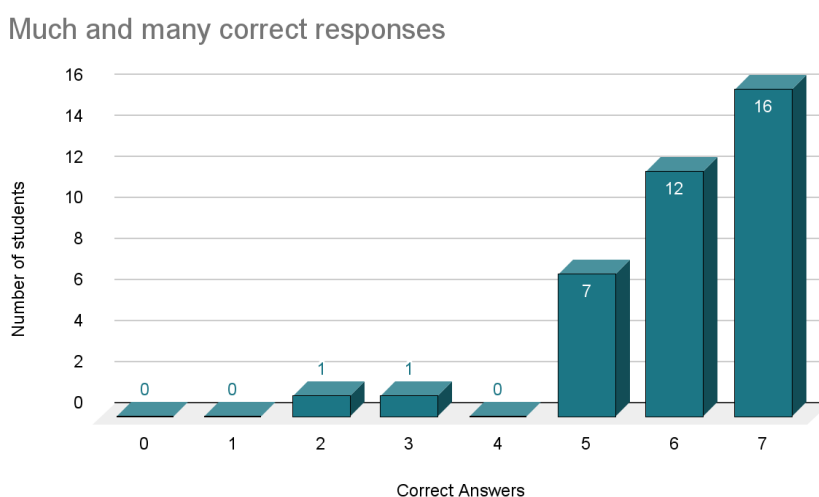
Although the number of students with no errors was eight, the majority of the group had committed at least one error, which gave room for improvement through the app. The result of the previous questionnaire proves that a variable level can be found in the group. This level variation may also be found in the app's results and in the final questionnaire.

5.2 Students' Exercises Results in the App

The students, after completing the previous knowledge questionnaire and the time of autonomous conversation with the chatbot had to do the exercises in the app. These exercises of fill-in-the-gap sentences treated the lesson of the quantifiers much, many, some and any, the same as the first questionnaire (see 5.1). These two lessons, one for much and many and the other for some, had seven exercises each, making a total of fourteen exercises. Students had 20 minutes to complete the exercises, which gave them approximately a minute and a half for each sentence. When all students finished, we gathered their results manually by asking each of the students for the number of correct answers in each lesson. In Graphics 4 and 5 we can see the number of correct answers for each lesson.

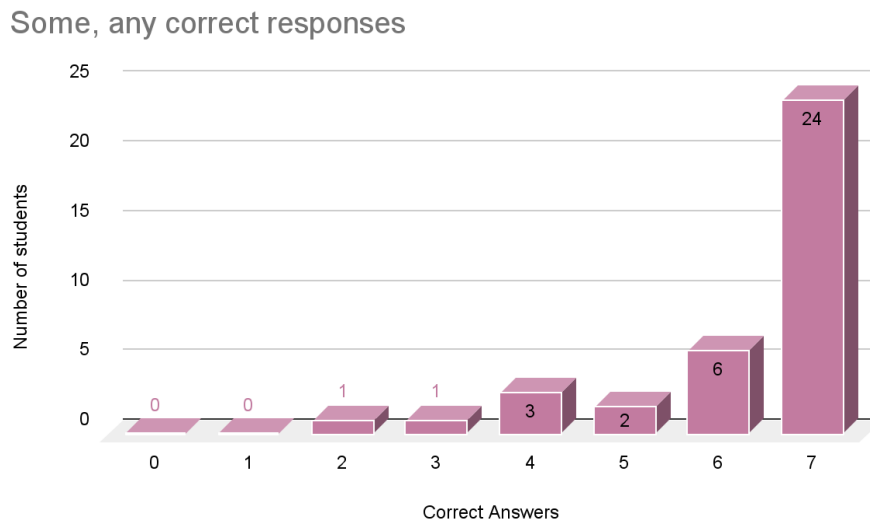
Graphic 4.

Correct answers in 'much, many' exercises



Graphic 5.

Correct Answers in 'some, any' exercises



If we compare these results with the ones obtained by the students in the previous knowledge questionnaire (see Graphic 3 in the methodology section), we can see an improvement. In the first set of exercises, the average score is 6,27 correct answers (an equivalent of 8,96 out of 10), while in the second, it is 6.24 average points (8.91 out of 10). The average score in the initial questionnaire was six correct answers out of a total of eight questions (7.5 out of 10), demonstrating a significant improvement in the class' average score in this second set of exercises. The average score of the app's activities is 6,25 (8.93/10) against the previous questionnaire average of 6 (7.5/10). Similar exercises have obtained different results in a short period of time. One of the reasons for this is the initial explanation of the possible options that the app gives before starting the exercises. Unlike in the previous questionnaire, the explanation is given to the student in the chat and can always be looked up in case of doubt. Also, we can see a shift in the variation of the class. While in the previous questionnaire, the majority were in the middle-upper scale of the results (32 students were in the 5-8 interval), in the app's exercises, the majority of the students are in the higher number of results, especially in the 'some, any' lesson where the majority of students (24 students) obtained a perfect score. Overall, the group's exercise results improved in the app's environment and, despite the differences in level, the result's variation of the class was reduced due to the general rise in the number of correct answers.

5.3 Students' Opinions Towards the Andy App

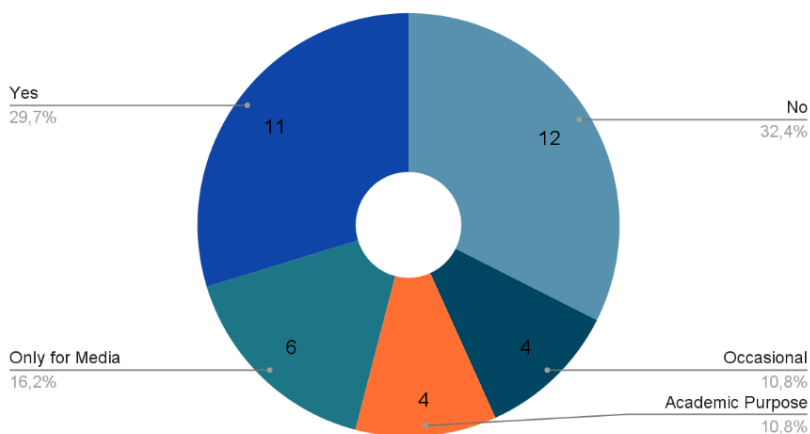
Apart from the results' data, our mixed investigation needs qualitative research on the learning method we propose. The combined analysis of quantitative and qualitative research will determine the appropriateness of the teaching and learning method which makes use of an AI-powered app. For this reason, students were asked to complete one last questionnaire after finishing the app's usage. This final questionnaire's primary purpose is to register the opinions and impressions that the app-driven learning method caused in the students. This section of our research evaluates some of the most critical features of AI apps, like motivation aspects, ease of use, the authenticity of the materials and usage of language learning apps.

In the first question, students were asked to rate their enjoyment of English on a scale from 1 to 5. The results are one student values it as 1, five students as 2, nine students as 3, eleven students as 4 and twelve students as 5. A considerable percentage of the group (32 students) positioned English in the upper half of the scale, which indicates that the majority of the group positively values English, and their predisposition towards using the app could be positive. In the following question, students were asked to indicate how frequently and with which purposes they used English in their daily lives. As we can see in Graphic 6, the results were very varied.

Graphic 6.

Frequency and uses of English

Graphic 6. Frequency and uses of English

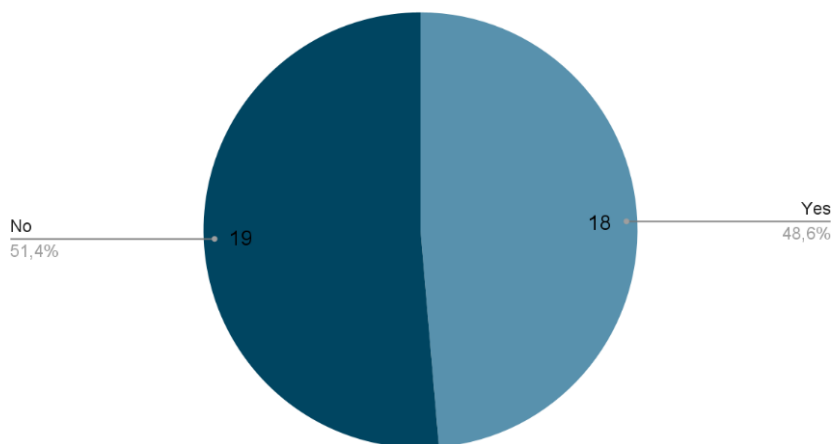


Again, a noticeable variation in the group can be found. While a big portion of the group does not consider that they use English daily, an almost similar quantity of students uses English daily and for a wide range of purposes and environments. Also, two other smaller groups only use English for a specific purpose, like academic goals (to learn English or another subject) or to understand and consume social media, films, shows and video games. A smaller group only uses English when needed for special situations. This wide range of language uses can strongly affect opinions towards the app, as not every student has the same concept of English. In the same way that we asked the teacher about the language learning apps, students were also asked about their use of language learning apps in their free time. As we can see in Graphic 7, the number of students using and not using some kind of language learning app is very similar.

Graphic 7.

Use of language learning apps of the students

Do you use, or have used, language learning apps?



The answers to this question show us that almost half of the research group has previous experience using language learning apps, which can influence the opinion of the students regarding several of the app's features, like the ease of use or the quality of the exercises. With this data, we asked the students that answered yes in the previous question to state the apps that they used. In Table 7 we can see each mentioned app with the number of students that use them.

Table 7.

Most popular language learning apps.

| Name of the app | Number of students |
|-----------------|--------------------|
| Duolingo | 15 |
| SmartUp | 1 |
| DeepL | 1 |
| Andy | 1 |

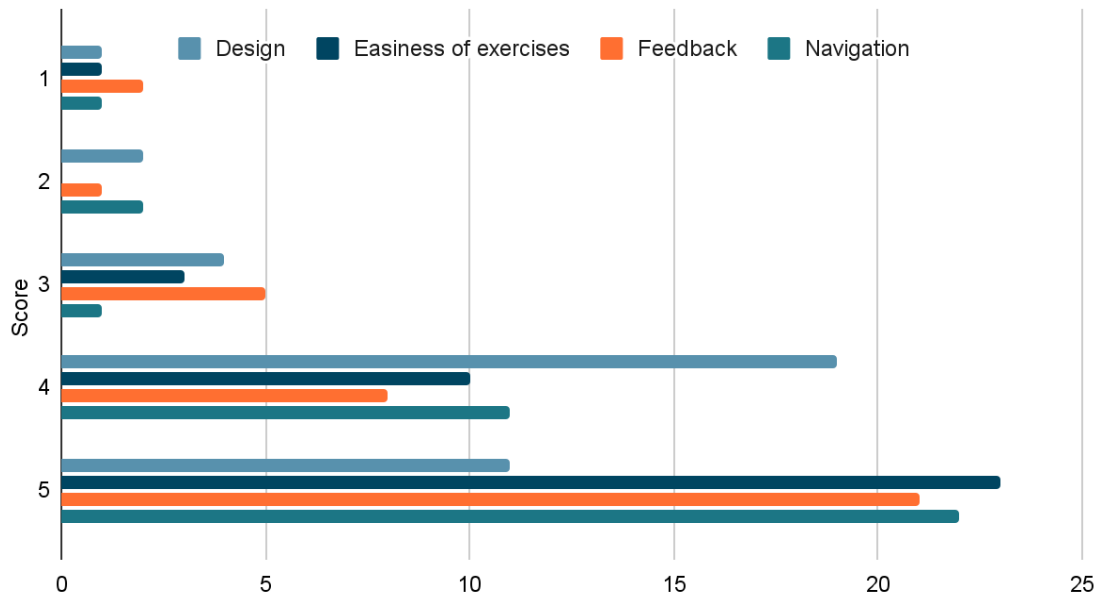
As we can see, Duolingo is the most popular language-learning app in the market. Surprisingly, two students mentioned two AI-powered apps like Andy, the chatbot we use for our experiment, and the AI-powered translation tool DeepL. The popularity of Duolingo can prevent students from discovering new applications and different approaches to language learning, like the chatbot method we used in this research.

For the following three questions, students had to mark four different aspects of the app on a scale from one to five. The design and style of the app, the easiness of the proposed exercises, the app's evaluations, and corrections after completing practice and how easy it was for them to navigate through the app and find the lessons section. In Graphic 8, we can see the values given to these features by the students.

Graphic 8 .

Given punctuation to different features of the app

Design, Easiness of exercises, Corrections y Navigation



A high number of students in the research group value positively these four characteristics of the app. Among these four, the design of the app is the one feature that received the lowest average score, meaning that the students saw it as the worst aspect of the app. Many times, the style and design of the app are decisive for the public to use it, and in this case, this app may need improvement in this aspect. On the other hand, the highest average score was given to the easiness of the exercises. As we previously mentioned, the initial explanation given by the app can help students complete the exercises correctly. We should also point out the score given to the feedback feature. One crucial aspect of these apps is a good feedback system that gives the user appropriate corrections in case of a right or wrong response, and in this case, students valued it positively with an average score of 4.2 out of 5. As we can see in the graphic, students found it easy to navigate through the app, this could be due to a combination of an intuitive design of the app and the student's familiarity with mobile devices and similar apps.

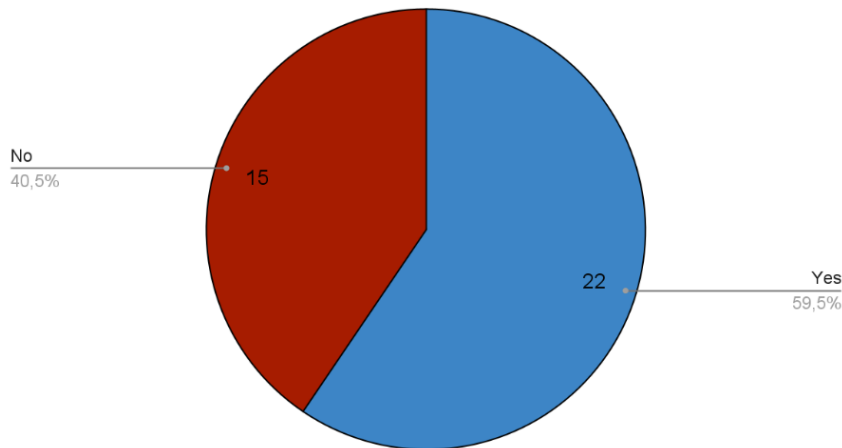
As we previously described, chatbots simulate human-to-human conversation. The better the chatbot is developed, the better and more realistic the conversation would be with the user. In the next question, students were asked whether the

conversation with the chatbot seemed real and authentic. In Graphic 9, we can see that more than half of the group saw the conversation as natural.

Graphic 9.

Natural conversation with ANDY

Graphic 9. Was Andy's conversation natural?



More than half of the research group classified the conversation with Andy's app as authentic. The way the app reacted to the students' messages made sense, was coherent and followed the topic introduced by the student. Despite this, some of the students saw this conversation as one of the app's features that needed more improvement. In the final question, students had to write any kind of comment, question, or advice about the app. We have classified these comments into three categories, negative comments/ something to improve, neutral comments/ no comments, and positive comments. In the Table 8, we can see the number of comments for each type and the most significant comments.

Table 8.

Comments and opinions of ANDY

| Negative comments/ Improvable features | Neutral/no comment | Positive comment |
|--|--|---|
| Number: 6 | Number: 18 | Number: 13 |
| Examples: “It should give more natural responses” “Improve responses to look more human and enhance loading time” “More free exercises should be available” | Examples: “Nothing to add” “Nothing” | Examples: “Very good app 😊” “It is very good” “Very interesting app” |

As in the other questions, the responses are varied. Some students think that the chatbot capability of the app can be improved, as the responses of the system were inaccurate or off-topic. Additionally, there were two negative comments regarding the limited features of the app in the free version. On the other hand, and if we consider neutral comments as a kind of positive, the majority of the group valued the app positively and saw no problems or difficulties in its use. Apart from, the students' opinions we also gathered the teacher's impression of the app, as she is responsible for creating the exercises and designing the teaching plan where the app could be used.

5.4 Teacher's Assessment of the App and Its Use in the Classroom

In addition to the students' assessments of the app, the teacher was also asked about several of the app's features. All these questions were made orally, and answers were recorded after finishing the experiment with the students. First, the teacher was asked to value the design and style of the app. She described it as

“simple, with a good choice of colours and the fact that is basic can help the app work in school’s devices” (A. Sánchez, Personal communication, 16 June 2022). She also considered this simplicity as a negative aspect of the app, as “images, videos and media in general can be more appealing to students” (A. Sánchez, Personal communication, 16 June 2022). The next question asked about the chatbot capabilities, and the exercises students completed in it. The teacher valued the chatbot positively because “the responses were good enough to produce a nice interaction with the students and the topics were familiar enough for students to know the topic and how to apply some vocabulary” (A. Sánchez, Personal communication, 16 June 2022). She also thought that the known-words feature was useful and that the use of emojis in both the messages and the app’s corrections made the conversation seem more real. In the exercise section, the teacher believed that the input method could be improvable. “It would be better if the app allowed the students to write the answer instead of tapping the response. It can help students learn orthography and vocabulary, maybe the student can choose the method of input, tapping or writing” (A. Sánchez, Personal communication, 16 June 2022). She also described the app’s voice as “very robotic and with improvable pronunciation” (A. Sánchez, Personal communication, 16 June 2022). Lastly, when asked about the usability of this app in the class, the teacher stated the following:

“On many occasions, schools lack the resources to buy apps like this one. The number of licenses needed is too high and group licenses are not available in most apps. This app can easily be applied to class objectives, and it can be used both in class and at home as homework. Also, the free version is limited and can only be used in 2º and 3º ESO.” (Sánchez, Personal communication, 16 June 2022)

As we can see teacher values the app positively and supports its use in the classroom, although some improvements can be made for it to be more successful in a learning environment. With all these results we can now decide whether we have managed to answer our research questions and completed our objectives with our experiment.

6. CONCLUSIONS

This dissertation has focused on successfully applying an AI chatbot app to a language learning class. With both quantitative and qualitative research, we can analyse the experiment's results and set an example of the fruitful use of AI technology in the language learning process. After the analysis of all the results of our investigations, we can give an answer to the research questions we presented in the introduction. In this section, we define the various benefits that the app brings to the language learning process and assess how these benefits can be achieved in other aspects of language learning.

Until now, this research has made use of several questionnaires to register the results and the app ANDY, where the research group has operated an AI-powered chatbot and completed grammatical exercises.

After analysing all the results from the various sources, we can notice significant outcomes. Some improvements and benefits have appeared while using the app. First, the use of the app boosted students' motivation within the learning process. The use of technology presents an attractive method of learning a second language as it diverts from traditional models that rely primarily on repetition and memorisation. During the experiment, students showed interest and improved motivation in both sections of the app's usage, using the chatbot and completing the exercises. The motivation boost is one of the first benefits that partly answers our second and third research questions. This enhancement improves the student's learning process and makes the teacher's tasks more manageable, as students are more willing to participate.

The second and most notable improvement is the rise in the number of correct answers in similar—style exercises. If we compare Graphic 3 with Graphics 4 and 5 (see Results Section), which portray the previous knowledge questionnaire and the app's results, respectively, an increment in the average score of the class is easily noticeable. The average score of the class rose from 7.5/10 on the previous level test to 8.93/10 in the app's exercises. The better performance was homogeneously in the research group, as every range of results was improved, and more students achieved better results that fall in the upper range of results (12-14 correct answers). One of the reasons for this is the explanation given by the chatbot before the exercises. The chatbot explains the possible responses to

the practices in a message that stays in the chat. This message can be consulted by students whenever doubt or difficulty appears while completing the exercises. The explanation of the activities also reinforces the idea of the app as a tool, as it shifts the approach of exercises as a test, where the student only learns by repetition of the theory or error revision, towards the approach of exercises as an accompanied practice, where the theory is always available, and the student can revise it. These results also partly answer our second and fourth research questions, as the app's usage built an improvement in the students' results in a specific grammatical lesson while working as a supportive tool.

Another significant result is that no students had a problem downloading or using the app. By analysing the results of the last questionnaire, we can observe that approximately half of the research group had previously used a language learning app (See Graphic 7). Despite this, all students successfully installed and used the app without issues and more importantly, both students and the teacher valued the navigation through the app very positively. This shows that students are completely acquainted with learning technology and answers our third research question. Learning technology application in a classroom is no longer a challenge because the students fully know how to use it and even more if they make use of their personal mobile devices.

Regarding the chatbot capabilities of the app, the teacher and most of the students acknowledged the interaction and conversation as natural (See Graphic 9). The chatbot had enough intelligence to address the familiar topics introduced by the students and was able to keep up a conversation. Although some students opined that the responses could be more natural, the chatbot software was successfully used in the intended way in an English classroom, thus, answering our last research question.

Additionally, we must take into account the teacher's assessment of the app in view of the app proposed as a support tool for the English teacher. Our experiment with the app successfully aligned with the class's objectives and evaluation methods. It also transferred traditional fill—in—the—gaps exercises to a digital platform and achieved better results. Despite the positive feedback, the teacher also saw a necessity for better free versions or collective license buying options, as on many occasions schools' funding is restricted.

If this kind of AI technology is to be applied in any other language learning process, the app or system used must have several requirements to successfully achieve its purpose. The app must adapt its functioning according to the students' needs, provide accessibility possibilities and present a new and attractive way for learners to study the new language. Most students are completely familiar with the use of technology in a daily basis so the app must present a good design and appealing features.

After the analysis of the results, we have noticed some aspects of our investigation that could have been better. First, a more extensive research group would have positively affected the quantitative aspect of our study. A higher number of students would have allowed us to analyse and compare more results and apply different lessons and exercises in different research groups. When we conducted the experiment, the scholar year was also finished, so several students did not attend class anymore. If we had conducted it in the middle of the year, the totality of the students could have participated in the experiment. Also, related to this, an investigation conducted in classes of different levels would allow us to prove the efficiency of the app in various language levels and contexts. Lastly, the results could have included an analysis of the students' conversational improvement if the app allowed the automatic registration of the chat's messages. With this feature, we could have evaluated students' messages and interpreted their grammar, vocabulary, structures and topics. These limitations have appeared mainly because of a lack of time. With broader and longer research, many other features and approaches can be analysed. With the resulting investigations, new applications of AI technology and optimised systems can be discovered and applied to language learning. The research on other AI features can throw light into their applications in ESL, features like Machine Learning, speech recognition or automated monitoring, which can strongly influence the learning of a second language.

In conclusion, according to the results obtained, we have successfully applied an Artificial chatbot app as a complementary tool in an English as a Second Language classroom. The use of the app has boosted students' motivation, improved the results of traditional exercises by transferring them into a digital platform, and enhanced students' writing and conversational abilities. Although additional finance or better economic accessibility may be needed, our

experiment has proved that the integration of innovative language learning technology can be easily achieved and that it brings numerous benefits for both the students and the teacher. The union between this kind of technology and language learning is constantly increasing. With this work, we hope to have helped prove the benefits and objectives this implementation can bring to the learning community. We also hope to have laid the foundations for artificial intelligence applications to find their place in language learning.

7. REFERENCES

- Adamopoulou, E., & Moussiades, L. (2020). *An Overview of Chatbot Technology* (pp. 373–383). https://doi.org/10.1007/978-3-030-49186-4_31
- Ahn, J., Chang, M., Watson, P., Tejwani, R., Sundararajan, S., Abuelsaad, T., & Prabhu, S. (2018). *Adaptive Visual Dialog for Intelligent Tutoring Systems* (pp. 413–418). https://doi.org/10.1007/978-3-319-93846-2_77
- Ai, Y., Peng, M., & Zhang, K. (2018). Edge Computing Technologies for internet of things: A Primer. *Digital Communications and Networks*, 4(2), 77–86. <https://doi.org/10.1016/j.dcan.2017.07.001>
- Alderson, J. C. (2007). The CEFR and the Need for More Research. *The Modern Language Journal*, 91(4), 659–663. <http://www.jstor.org/stable/4626093>
- Allen, E. D. (1960). The Effects of the Language Laboratory on the Development of Skill in a Foreign Language. *The Modern Language Journal*, 44(8), 355–358. <https://doi.org/10.2307/321068>
- Anggriyashati, R. (2020). Negative Effects of MALL on the Improvement of Learners Autonomy and Motivation. *The International Conference on Innovations in Social Sciences and Education*, 1(1) (pp. 613-622).
- AntiqueTech. (2013). *The ENIAC vs. the cell phone*. AntiqueTech. Retrieved May 26, 2022, from http://www.antiquetech.com/?page_id=1438
- Arnold, K. C., Chauncey, K., & Gajos, K. Z. (2020). Predictive text encourages predictable writing. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 128–138. <https://doi.org/10.1145/3377325.3377523>
- Arriany, A. A., & Musbah, M. S. (2016). Applying voice recognition technology for smart home networks. *2016 International Conference on Engineering & MIS (ICEMIS)*, 1–6. <https://doi.org/10.1109/ICEMIS.2016.7745292>
- Bax, S. (2003). CALL—past, present and future. *System*, 31(1), 13–28. [https://doi.org/10.1016/S0346-251X\(02\)00071-4](https://doi.org/10.1016/S0346-251X(02)00071-4)
- Bengtsson, M. (2016). How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, 2, 8–14. <https://doi.org/10.1016/j.npls.2016.01.001>
- Benito-Ruiz, E. (2009). Infocination 2.0. In *Handbook of Research on Web 2.0 and Second Language Learning* (pp. 60–79). IGI Global. <https://doi.org/10.4018/978-1-60566-190-2.ch004>

- Benko, A., & Sik Lányi, C. (2009). History of Artificial Intelligence. In *Encyclopedia of Information Science and Technology, Second Edition* (pp. 1759–1762). IGI Global. <https://doi.org/10.4018/978-1-60566-026-4.ch276>
- Bézy, M., & Settles, B. (2016). The Duolingo English Test and East Africa: Preliminary linking results with IELTS & CEFR. *Reading*, 5(1.0), 4–85.
- Bini, S. A. (2018). Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? *The Journal of Arthroplasty*, 33(8), 2358–2361. <https://doi.org/10.1016/j.arth.2018.02.067>
- Birajdar, L. M. (2020). Application of CEFR (Common European Framework of Reference) in Testing the English Language Proficiency Level of Polytechnic Students. *Journal of English Language Teaching*, 62(1).
- Calegari, R., Ciatto, G., Denti, E., & Omicini, A. (2020). Logic-based technologies for intelligent systems: State of the art and perspectives. *Information (Switzerland)*, 11(3). (pp. 1-29) <https://doi.org/10.3390/info11030167>
- Campbell, M., Hoane, A. J., & Hsu, F. (2002). Deep Blue. *Artificial Intelligence*, 134(1–2), 57–83. [https://doi.org/10.1016/S0004-3702\(01\)00129-1](https://doi.org/10.1016/S0004-3702(01)00129-1)
- Carbonell, J. R. (1970). AI in CAI: An Artificial-Intelligence Approach to Computer-Assisted Instruction. *IEEE Transactions on Man-Machine Systems*, 11(4), 190–202. <https://doi.org/10.1109/TMMS.1970.299942>
- Catania, A. (1988). The operant behaviorism of B.F. Skinner. In S. Harnad (Ed.), *The Selection of Behavior: The Operant Behaviorism of B. F. Skinner: Comments and Consequences*. CUP Archive.
- Caudill, M. (1987). *IEEE First International Conference on Neural Networks* (Vol. 3). SOS Print.
- Chatterjee, S., & Bhattacharjee, K. K. (2020). Adoption of artificial intelligence in higher education: a quantitative analysis using structural equation modelling. *Education and Information Technologies*, 25(5), 3443–3463. <https://doi.org/10.1007/s10639-020-10159-7>
- Chomsky, N. (1966). *Topics in the Theory of Generative Grammar*. DE GRUYTER. <https://doi.org/10.1515/9783110903843>
- Chowdhary, K. R. (2020). Natural Language Processing. In K. R. Chowdhary (Ed.), *Fundamentals of Artificial Intelligence* (pp. 603–649). Springer India. https://doi.org/10.1007/978-81-322-3972-7_19
- Cloete, A. L. (2019). Unstoppable: A critical reflection on the socio-economic embeddedness of technology and the implications for the human agenda. *HTS Teologiese Studies / Theological Studies*, 75(2). <https://doi.org/10.4102/hts.v75i2.5550>

- Cobbe, J. (2021). Algorithmic Censorship by Social Platforms: Power and Resistance. *Philosophy & Technology*, 34(4), 739–766. <https://doi.org/10.1007/s13347-020-00429-0>
- COE. (2001). Structured overview of all CEFR scales. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*.
- Commission, E., Centre, J. R., Samoli, S., López Cobo, M., Gómez, E., de Prato, G., Martínez-Plumed, F., & Delipetrev, B. (2020). *AI watch : defining artificial intelligence : towards an operational definition and taxonomy of artificial intelligence*. Publications Office. <https://doi.org/doi/10.2760/382730>
- Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Chapter 37 - Intelligent Tutoring Systems. In M. G. Helander, T. K. Landauer, & P. v Prabhu (Eds.), *Handbook of Human-Computer Interaction (Second Edition)* (pp. 849–874). North-Holland. <https://doi.org/https://doi.org/10.1016/B978-044481862-1.50103-5>
- Dartmouth, V. (2006). *Artificial Intelligence (AI) Coined at Dartmouth*. Vox of Dartmouth. <https://250.dartmouth.edu/highlights/artificial-intelligence-ai-coined-dartmouth>
- Davies, G., Otto, S., & Rüschoff, B. (2012). Historical perspectives on CALL. In M. Thomas, H. Reinders, & M. Warschauer (Eds.), *Contemporary Computer-Assisted Language Learning*, (Vol. 18, pp. 19–39). A&C Black.
- Digón, P. (2003). The Ley Orgánica de Calidad de la Educación: A Critical Analysis of the New Spanish Education Reform. *Revista Electrónica de Investigación Educativa*, 5(1). (pp. 57-71)
- Eckroth, J. (2021). *General Problem Solver (GPS)*. Ohio State University. <http://web.cse.ohio-state.edu/~stiff.4/cse3521/gps.html#sec-7>
- Elliot, A. (2019). *The Culture of AI: Everyday Life and the Digital Revolution*. Routledge.
- Ferreira, A., & Atkinson, J. (2009). Designing a Feedback Component of an Intelligent Tutoring System for Foreign Language. In *Research and Development in Intelligent Systems XXV* (pp. 277–290). Springer London. https://doi.org/10.1007/978-1-84882-171-2_20
- Figueras, N. (2012). The impact of the CEFR. *ELT Journal*, 66(4), 477–485. <https://doi.org/10.1093/elt/ccs037>
- Fogel, D. B., Liu, D., & Keller, J. M. (2016). *Fundamentals of Computational Intelligence*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781119214403>
- Forbus, K. D. (2010). AI and Cognitive Science: The Past and Next 30 Years. *Topics in Cognitive Science*, 2(3), 345–356. <https://doi.org/10.1111/j.1756-8765.2010.01083.x>

- Fox, M. A., Aschkenasi, C. J., & Kalyanpur, A. (2013). Voice recognition is here comma like it or not period. *Indian Journal of Radiology and Imaging*, 23(03), 191–194.
<https://doi.org/10.4103/0971-3026.120252>
- Fritz, A. (2021). *Image Recognition Guide*. Fritz Labs.
- Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, 279–289.
<https://doi.org/https://doi.org/10.1016/j.chb.2018.12.023>
- Fryer, L., & Carpenter, R. (2006). Bots as language learning tools. *LANGUAGE LEARNING & TECHNOLOGY*, 10(3), 8–14.
- Fulop, S. A. (2012). Mathematical Linguistics and Learning Theory. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 2110–2113). Springer US.
https://doi.org/10.1007/978-1-4419-1428-6_891
- Gamlo, N. (2019). The Impact of Mobile Game-Based Language Learning Apps on EFL Learners' Motivation. *English Language Teaching*, 12(4), 49.
<https://doi.org/10.5539/elt.v12n4p49>
- Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning and Technology*, 15, 2–11.
- González, J., & Wagenaar, R. (2008). *Universities' contribution to the Bologna Process - An introduction*.
- Gu, Y., Jin, Z., & Chiu, S. C. (2015). Active learning combining uncertainty and diversity for multi-class image classification. *IET Computer Vision*, 9(3), 400–407.
<https://doi.org/10.1049/iet-cvi.2014.0140>
- Guan, C., Mou, J., & Jiang, Z. (2020). Artificial intelligence innovation in education: A twenty-year data-driven historical analysis. *International Journal of Innovation Studies*, 4(4).
<https://doi.org/10.1016/j.ijis.2020.09.001>
- Guo, L., Wang, D., Gu, F., Li, Y., Wang, Y., & Zhou, R. (2021). Evolution and trends in intelligent tutoring systems research: a multidisciplinary and scientometric view. *Asia Pacific Education Review*, 22(3), 441–461. <https://doi.org/10.1007/s12564-021-09697-7>
- Haristiani, N. (2019). Artificial Intelligence (AI) Chatbot as Language Learning Medium: An inquiry. *Journal of Physics: Conference Series*, 1387(1), 012020.
<https://doi.org/10.1088/1742-6596/1387/1/012020>
- Heift, T., & Vyatkina, N. (2017). Technologies for Teaching and Learning L2 Grammar. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 26–44). Wiley.
<https://doi.org/10.1002/9781118914069.ch3>

- Hodges, C. B., Moore, S., Lockee, B. B., Trust, T., & Bond, M. A. (2020). The Difference Between Emergency Remote Teaching and Online Learning. *Scholarly Works, Institute for Society, Culture, and Environment (ISCE)*, 1–12.
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Hubbard, P. (2017). Technologies for Teaching and Learning L2 Listening. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 93–106). Wiley. <https://doi.org/10.1002/9781118914069.ch7>
- Hwang, G.-J. (2003). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, 40(3), 217–235. [https://doi.org/10.1016/S0360-1315\(02\)00121-5](https://doi.org/10.1016/S0360-1315(02)00121-5)
- Jain, A. K., Duin, R. P. W., & Mao, J. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1), 4–37. <https://doi.org/10.1109/34.824819>
- Johnson, B. G., Phillips, F., & Chase, L. G. (2009). An intelligent tutoring system for the accounting cycle: Enhancing textbook homework with artificial intelligence. *Journal of Accounting Education*, 27(1), 30–39. <https://doi.org/10.1016/j.jaccedu.2009.05.001>
- Johnson, W. L., & Valente, A. (2009). Tactical Language and Culture Training Systems: Using AI to Teach Foreign Languages and Cultures. *AI Magazine*, 30(2), 72. <https://doi.org/10.1609/aimag.v30i2.2240>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. <https://doi.org/https://doi.org/10.1016/j.bushor.2018.08.004>
- Kim, H., & Yeonhee, K. (2012). Exploring smartphone applications for effective mobile-assisted language learning. *Multimedia-Assisted Language Learning*, 15(1), 31–57.
- Kim, T.-K. (2020). Short Research on Voice Control System Based on Artificial Intelligence Assistant. *2020 International Conference on Electronics, Information, and Communication (ICEIC)*, 1–2. <https://doi.org/10.1109/ICEIC49074.2020.9051160>
- Kubrick, S. (1968). *2001: A Space Odyssey*. Metro Goldwyn Mayer.
- Kukulska-Hulme, A. (2009). Will mobile learning change language learning? *ReCALL*, 21. <https://doi.org/10.1017/S0958344009000202>
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. University of Chicago Press. <https://doi.org/10.7208/chicago/9780226471013.001.0001>
- Lang, F. (1927). *Metropolis*. UFA.

- Lee, S.-M., & Park, M. (2020). Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning*, 33(8), 936–959. <https://doi.org/10.1080/09588221.2019.1602545>
- Levy, M. (1997). *Computer-Assisted Language Learning: Context and Conceptualization*. Clarendon Press.
- Li, P., & Lan, Y.-J. (2021). Digital Language Learning (DLL): Insights from Behavior, Cognition, and the Brain. *Bilingualism: Language and Cognition*, 1–18. <https://doi.org/DOI:10.1017/S1366728921000353>
- Liaw, M.-L., & English, K. (2017). Technologies for Teaching and Learning L2 Reading. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 62–76). John Wiley & Sons, Ltd. <https://doi.org/https://doi.org/10.1002/9781118914069.ch5>
- Liskov, B. (1978). Session Details: LISP Session. In *History of Programming Languages*. Association for Computing Machinery. <https://doi.org/10.1145/3243999>
- Lokman, A. S., & Amedeen, M. A. (2019). Modern Chatbot Systems: A Technical Review. In K. Arai, R. Bhatia, & S. Kapoor (Eds.), *Proceedings of the Future Technologies Conference (FTC) 2018* (pp. 1012–1023). Springer International Publishing.
- Lucas, G. (1977). *Star Wars*. 20th Century Fox.
- Marcus, G., Davis, E., & Aaronson, S. (2022). *A very preliminary analysis of DALL-E 2*.
- Martín-Monje, E., Arús, J., Rodríguez, Arancón, P., & Calle-Martínez, C. (2014). *REALL: Rubric for the evaluation of apps in language learning*.
- Mayor, A. (2018). *Gods and Robots*. Princeton University Press. <https://doi.org/10.1515/9780691185446>
- McCarthy, J. (1959). Programs with Common Sense. *The First International Symposium on Artificial Intelligence*.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *AI Magazine*, 27(4). (pp. 12-14)
- McClelland, J., & Cleeremans, A. (2009). *Connectionist Models*.
- Miettinen, K. (1999). *Evolutionary algorithms in engineering and Computer Science: Recent advances in genetic algorithms, evolution strategies, evolutionary programming, genetic programming and Industrial Applications*. John Willey & Sons.
- Minsky, M. (1974). *Frames: A Framework for Representing Knowledge*. MIT- AI Laboratory. <https://web.media.mit.edu/~minsky/papers/Frames/frames.html>

- Nwana, H. S. (1990). Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, 4(4), 251–277. <https://doi.org/10.1007/BF00168958>
- Odo, C. (2018). *Adapting Learning Activities Selection in an Intelligent Tutoring System to Affect* (pp. 521–525). https://doi.org/10.1007/978-3-319-93846-2_98
- OpenAI. (2022). *DALL·E 2*. OpenAI. <https://openai.com/dall-e-2/>
- Oppy, G., & Dowe, D. (2021). The Turing Test. In Z. Edward N. (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter). Metaphysics Research Lab, Stanford University.
- Orden ECD/489/2016, (2016). <https://educa.aragon.es/documents/20126/868873/ORDEN+CURRICULO+SECUNDARIA+2016.pdf/cf9e8c58-4ae0-886b-9311-9863eddf9c5?t=1593156016565>
- Otto, S. E. K. (2017). From Past to Present. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 10–25). John Wiley & Sons, Ltd. <https://doi.org/https://doi.org/10.1002/9781118914069.ch2>
- Otto, S. E. K. (2017). From Past to Present. In *The Handbook of Technology and Second Language Teaching and Learning* (pp. 10–25). Wiley. <https://doi.org/10.1002/9781118914069.ch2>
- Oxford University Press. (2022). Artificial Intelligence. *Oxford Dictionary*. Retrieved May 6, 2022, from <https://www.oxfordlearnersdictionaries.com/definition/english/artificial-intelligence?q=artificial+intelligence>.
- Paullada, A., Raji, I. D., Bender, E. M., Denton, E., & Hanna, A. (2021). Data and its (dis)contents: A survey of dataset development and use in machine learning research. *Patterns*, 2(11), (pp. 1-14) 100336. <https://doi.org/https://doi.org/10.1016/j.patter.2021.100336>
- Pham, X. L., Pham, T., Nguyen, Q. M., Nguyen, T. H., & Cao, T. T. H. (2018). Chatbot as an Intelligent Personal Assistant for Mobile Language Learning. *Proceedings of the 2018 2nd International Conference on Education and E-Learning*, 16–21. <https://doi.org/10.1145/3291078.3291115>
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Pradhan, P., Costa, L., Rybski, D., Lucht, W., & Kropp, J. P. (2017). A Systematic Study of Sustainable Development Goal (SDG) Interactions. *Earth's Future*, 5(11), 1169–1179. <https://doi.org/10.1002/2017EF000632>
- Puebla, C., Fievet, T., Tsopanidi, M., & Clahsen, H. (2022). Mobile-assisted language learning in older adults: Chances and challenges. *ReCALL*, 34(2), 169–184. <https://doi.org/10.1017/S0958344021000276>

- Qian, H., Dou, Z., Zhu, Y., Ma, Y., & Wen, J.-R. (2021). Learning Implicit User Profile for Personalized Retrieval-Based Chatbot. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*, 1467–1477. <https://doi.org/10.1145/3459637.3482269>
- Quasthoff, U., Goldhahn, D., & Eckart, T. (2014). Building Large Resources for Text Mining: The Leipzig Corpora Collection. In C. Biemann & A. Mehler (Eds.), *Text Mining: From Ontology Learning to Automated Text Processing Applications* (pp. 3–24). Springer International Publishing. https://doi.org/10.1007/978-3-319-12655-5_1
- Raja, & Nagasubramani, P. C. (2018). Impact of modern technology in education. *Journal of Applied and Advanced Research*, 3(1), 33–35.
- Replika. (2022). *Building a compassionate AI friend*. Luka Inc.
- Reynoso, R. (2021, May 25). *A complete history of artificial intelligence - G2*. A Complete History of Artificial Intelligence. Retrieved April 23, 2022, from <https://www.g2.com/articles/history-of-artificial-intelligence>
- Richards, J., & Rodgers, T. (2001). Major language trends in twentieth-century language teaching. In *Approaches and Methods in Language Teaching* (2nd ed.). Cambridge University Press.
- Ross, P. (1987). Intelligent tutoring systems. *Journal of Computer Assisted Learning*, 3(4), 194–203. <https://doi.org/10.1111/j.1365-2729.1987.tb00331.x>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- Sagiroglu, S., & Sinanc, D. (2013). Big data: A review. *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. <https://doi.org/10.1109/CTS.2013.6567202>
- Sánchez Torres, J. (2014). *Los papeles que desempeñan el “auxiliar de conversación” y el “profesor-coordinador” en centros bilingües español/inglés de Sevilla. Un estudio empírico de casos*. <https://idus.us.es/handle/11441/31276#.YxNBzA0EMdc.mendeley>
- Schank, R. C. (1972). Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology*, 3(4), 552–631. [https://doi.org/https://doi.org/10.1016/0010-0285\(72\)90022-9](https://doi.org/https://doi.org/10.1016/0010-0285(72)90022-9)
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424. <https://doi.org/DOI: 10.1017/S0140525X00005756>
- Secretaría General UPV. (2022). *Propuesta de acuerdo relativa a la ordenación e integración de las competencias transversales en los títulos oficiales de la Univeristat Politècnica de València* .

- Shannon, C. E. (1988). Programming a Computer for Playing Chess. In *Computer Chess Compendium* (pp. 2–13). Springer New York. https://doi.org/10.1007/978-1-4757-1968-0_1
- Sun, Y., & Gao, F. (2020). An investigation of the influence of intrinsic motivation on students' intention to use mobile devices in language learning. *Educational Technology Research and Development*, 68(3), 1181–1198. <https://doi.org/10.1007/s11423-019-09733-9>
- Suta, P., Lan, X., Wu, B., Mongkolnam, P., & Chan, J. (2020). An Overview of Machine Learning in Chatbots. *International Journal of Mechanical Engineering and Robotics Research*, 9(4), 502–510.
- Swift, J. (2008). *Gulliver's Travels*. Oxford University Press. <https://doi.org/10.1093/owc/9780199536849.001.0001>
- Tafazoli, D., & Golshan, N. (2014). Review of Computer-Assisted Language Learning: History, Merits & Barriers. *International Journal of Language and Linguistics*, 2(Special Issue: Teaching English as a Foreign/Second Language), 32–38.
- Taylor, C. R. (2004). Consumer Privacy and the Market for Customer Information. *The RAND Journal of Economics*, 35(4), 631–650. <https://doi.org/10.2307/1593765>
- Taylor, I. R. P. (2003). The Computer in School: Tutor, Tool, Tutee. *Contemporary Issues in Technology and Teacher Education*, 3(2), 240–252.
- The United Nations. (2015). *Do you know all 17 SDGs?* United Nations Department of Global Communications.
- Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching*, 45(2), 143–179. <https://doi.org/10.1017/S0261444811000528>
- Tono, Y. (2019). Coming Full Circle —From CEFR to CEFR-J and back. *CEFR Journal - Research and Practice*, 1, 5–67. <https://doi.org/10.37546/JALTSIG.CEFR1-1>
- TURING, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, LIX (236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Warwick, K., & Shah, H. (2016). Can machines think? A report on Turing test experiments at the Royal Society. *Journal of Experimental & Theoretical Artificial Intelligence*, 28(6), 989–1007. <https://doi.org/10.1080/0952813X.2015.1055826>
- Wees, D. (2009, December 3). *Paper use in schools*. The Reflective Educator.
- Weik, M. H. (1961). The ENIAC Story. *Ordnance*, 45(244), 571–575. <http://www.jstor.org/stable/45363261>

- Weizenbaum, J. (1966). *ELIZA—a computer program for the study of natural language communication between man and machine*. *Communications of the ACM*, 9(1), 36-45.
- Welin, C. W. (1979). Scripts, plans, goals and understanding, an inquiry into human knowledge structures: Roger C. Schank and Robert P. Abelson Hillsdale: Lawrence Erlbaum Associates, 1977. 248 pp. *Journal of Pragmatics*, 3(2), 211–217. [https://doi.org/https://doi.org/10.1016/0378-2166\(79\)90031-6](https://doi.org/https://doi.org/10.1016/0378-2166(79)90031-6)
- Winkler, R., & Söllner, M. (2018). Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. *Academy of Management Annual Meeting (AOM)*.
- Woolf, B. (2015). AI and education: Celebrating 30 years of marriage. *CEUR Workshop Proceedings*, 1432, 38–47.
- Yamazaki, K. (2014). Toward integrative CALL: A progressive outlook on the history, trends, and issues of CALL. *Tapestry Journal: : Digital Literacies for English Learners*, 6, 45–59.
- Yiu, C. (2012). *The Big Data Opportunity*. Policy Exchange.
- Zamir, A. R., & Shah, M. (2010). Accurate Image Localization Based on Google Maps Street View. In P. Kostas Daniilidis & N. P. Maragos (Eds.), *Computer Vision -- ECCV 2010* (1st ed., Vol. 6314, pp. 255–268). Springer Berlin. https://doi.org/10.1007/978-3-642-15561-1_19

8. ANNEXES

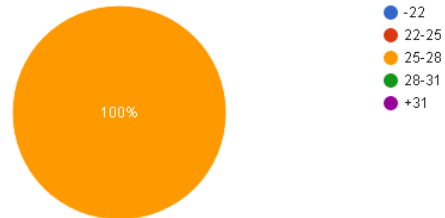
Annexe 1

Teacher previous questionnaire

1. Por favor, marque el rango de edad al que pertenece

 Copiar

1 respuesta



2. Indique su formación no obligatoria (Grado medio o superior, grado universitario, master...)

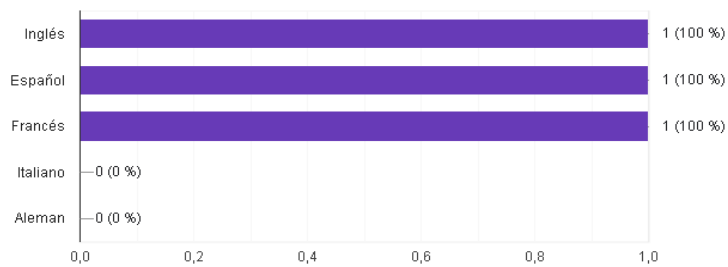
1 respuesta

Grado en Estudios Ingleses, Máster en Educación Secundaria Obligatoria, Máster en Enseñanza de Español como Lengua Extranjera

3. Indique los idiomas que habla (a todos los niveles)

 Copiar

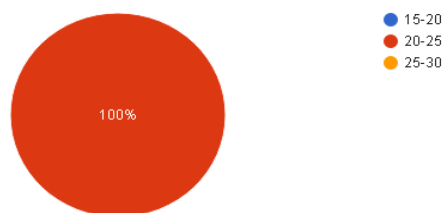
1 respuesta



4. Indique la cantidad media de alumnos y alumnas en sus clases

 Copiar

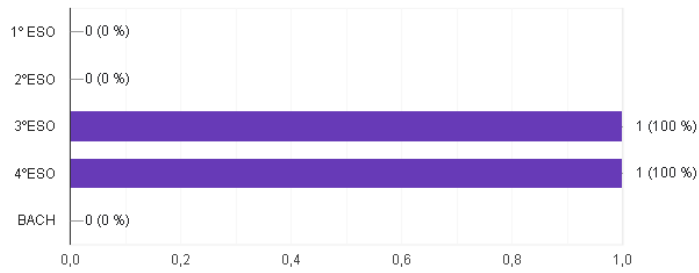
1 respuesta



5. ¿A que nivel de ESO da clase?

[Copiar](#)

1 respuesta



6. De aquellas que haya marcado en la respuesta anterior, ¿Cual es el nivel general de inglés en la clase? (A1, A2,B1,B2,C1,C2)

1 respuesta

A2/B1

7. Indique la ley educativa que sigue actualmente

1 respuesta

LOMCE/LOMLOE

8. Indique la importancia que cree que se le da a cada una de estas aptitudes en las clases de inglés

[Copiar](#)



9. ¿Que tipos de ejercicios realiza con mas frecuencia en su clase?

1 respuesta

Ejercicios del libro de texto (completar huecos, rephrasings), fichas con ejercicios extra de gramática similares, listenings con preguntas de verdadero/falso y respuestas breves

10. Cual cree que es la competencia que menos se desarrolla y practica.

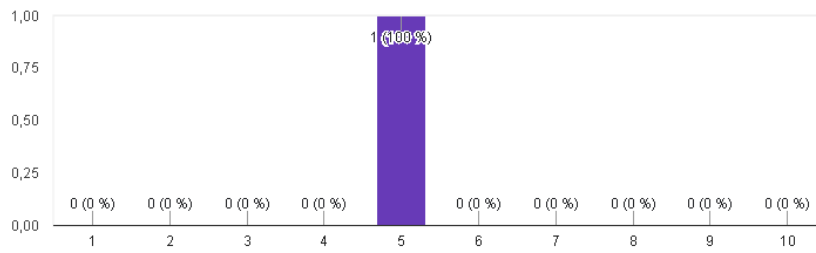
1 respuesta

Writing

11. Valora el uso de la tecnología en sus clases

 Copiar

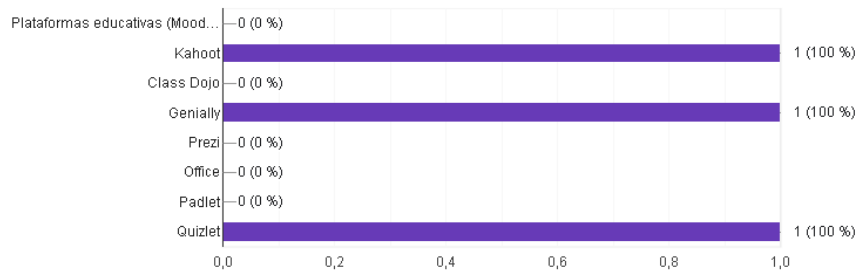
1 respuesta



12. ¿Que aplicaciones y herramientas de la siguiente lista usa o ha usado alguna vez en la clase?

 Copiar

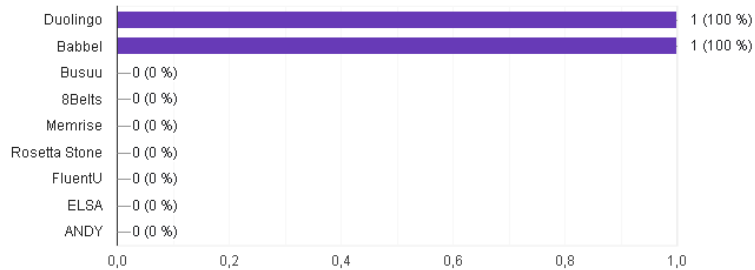
1 respuesta



13. ¿Cuales de las siguientes aplicaciones de aprendizaje de idiomas conoce?

 Copiar

1 respuesta



14. De las aplicaciones anteriores ¿Ha usado alguna alguna vez? ¿Cual?

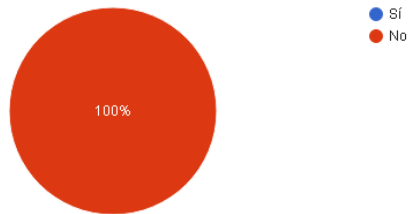
1 respuesta

Duolingo

15. ¿Ha usado alguna vez una aplicación similar en un contexto de clase?

 Copiar

1 respuesta



16. Muchísimas gracias por contestar a este cuestionario, por favor realiza cualquier pregunta o comentario que puedas tener.

0 respuestas

Aún no hay respuestas para esta pregunta.

Annex 2 Students' previous knowledge test

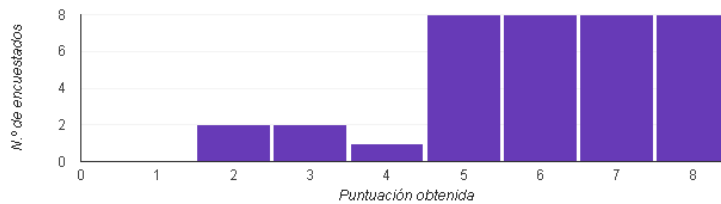
Estadísticas

Normal
6/8 puntos

Valor medio
6/8 puntos

Intervalo
2-8 puntos

Distribución de las puntuaciones totales



Preguntas en las que se suele fallar con frecuencia ?

Pregunta

Respuestas correctas

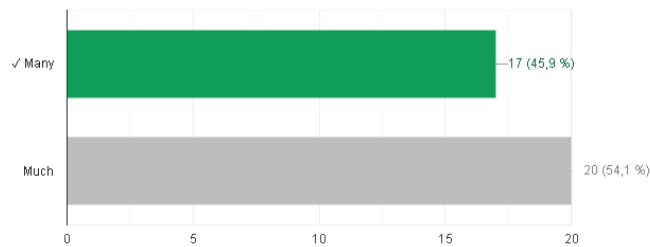
People had to jump over the fence to get to the match, _____ more weren't allowed to enter.

17/37

People had to **jump** over the fence to get to the match, _____ more weren't allowed to enter.

[Copiar](#)

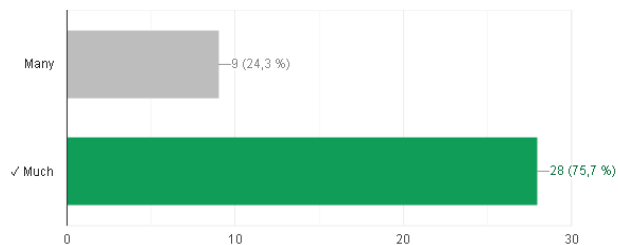
17 de 37 respuestas correctas



How _____ money does the new car cost?

[Copiar](#)

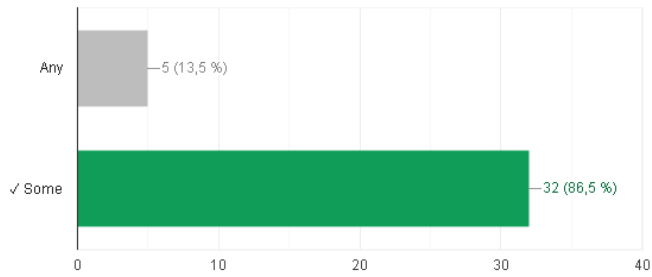
28 de 37 respuestas correctas



_____ people prefer to use the bus, in my case, I prefer the taxi.

 Copiar

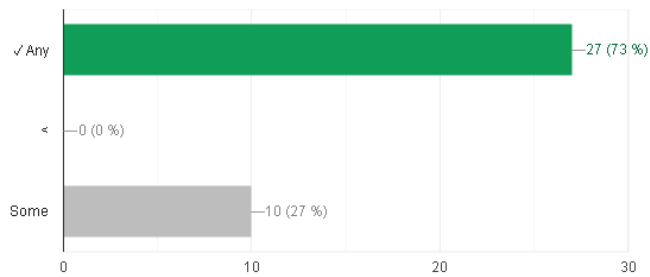
32 de 37 respuestas correctas



Please go outside, you can't stay here ___ longer

 Copiar

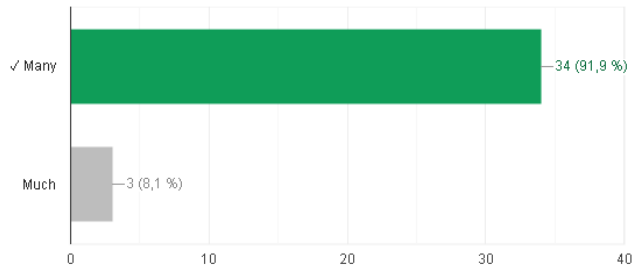
27 de 37 respuestas correctas



My family owns ___ properties in Italy

 Copiar

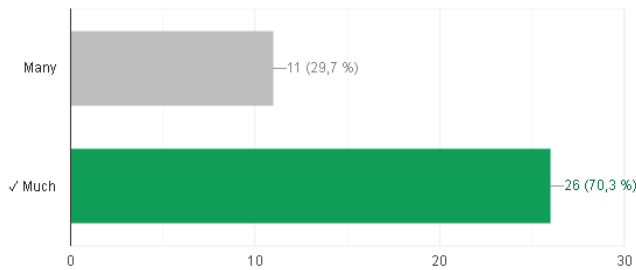
34 de 37 respuestas correctas



We don't need ___ milk for this recipe

 Copiar

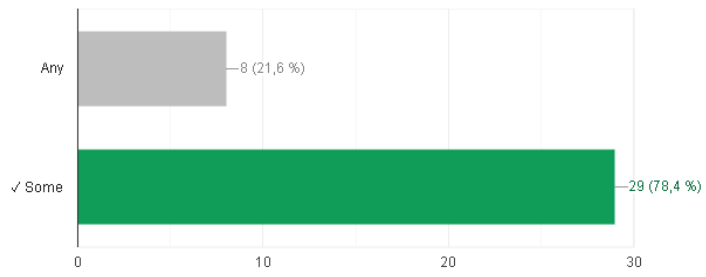
26 de 37 respuestas correctas



I really want to buy ___ white t-shirts

[Copiar](#)

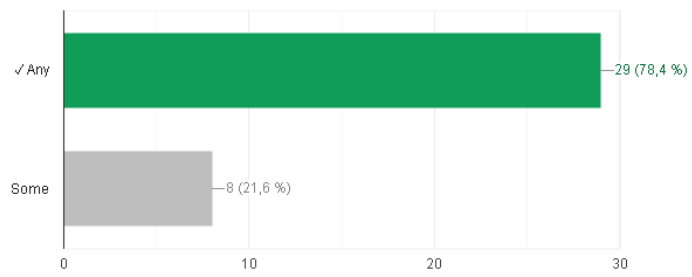
29 de 37 respuestas correctas



Are you going to this indie festival? I have never heard of ___ of these groups.

[Copiar](#)

29 de 37 respuestas correctas



Annex 3 Andy App features

The image displays two screenshots of the Andy App chat interface, showing a conversation between a user and the app. The app's navigation bar includes icons for Chat, Exercise, Grammar, and Stats, along with an 'Unlock Premium' button. The chat messages are as follows:

Left Screenshot:

- User: Hello 😊
- App: Hello
- User: My name is Andy. What is your name?
- App: My name is Jaime
- User: Nice to meet you, Jaime
- User: How are you?
- App: I'm fine and you?
- User: I'm fine, thanks for asking 😊
- App: Where are you from?
- User: It's hard to say :) My server is in Europe, but I consider myself a citizen of the world 🌍
- App: I'm from Zaragoza, Spain
- User: I've been to Madrid once. Beautiful city!
- App: Did you like it?
- User: Oh, I want to visit it now 😊

Right Screenshot:

- User: Hello 😊
- App: Let's learn some words! As always, you can write 'STOP' to exit the game 😊
- App: STANDARD
A standard is what people consider normal or good.
Buttons: Got it, I know this word, Explain
- App: This older model TV is below our store's standards.
Buttons: Got it, Explain
- App: Great!
- App: SESSION
A session is a meeting at a regular time in order to discuss and solve an issue.
Buttons: Got it, I know this word

Bottom Screenshot (Continuation of the right one):

- User: Cool, I've added it to your KNOWN words list 😊
- User: Now, try to remember the definition of the WORD
- User: After I say the definition, tell me how hard or easy this word is for you
- App: STANDARD
Definition
Buttons: Very bad, Explain, Easy, OK
- User: A standard is what people consider normal or good.
- App: "Some - Any" lesson please
- User: Sure
- App: In general we use SOME (also somebody/someone/something) in positive sentences and ANY (also anybody etc.) in negative sentences and questions.
We sold SOME books.
We didn't sell ANY books.
Did we sell ANY books?
- User: "Much - Many" lesson please
- User: Sure
- App: OK. The rule is very simple: you should use "many" with countable nouns: "many cats".
And "much" with uncountable nouns: "much water". You can't count water, but can count cats: 1 cat, 2 cats, 3 cats, etc. Do you understand?
- User: Easy
- App: OK, let's start the test
- App: There are so ___ poor people in the world.
Buttons: many, much
- User: A session is a meeting at a regular time in

Chat Exercise Grammar Stats Chat Exercise Grammar Stats

Unlock Premium > Unlock Premium >

"Some - Any" lesson please

OK, let's start the test

___ people like working alone, but I like teamwork.

any

some

some

Great job!

He's lazy. He never does ___ work.

any

some

some

We use ANY (also anybody etc.) in negative sentences.

During our holidays we visited ___ very interesting places.

any

some

some

Super!

Sure

In general we use SOME (also somebody/someone/something) in positive sentences and ANY (also anybody etc.) in negative sentences and questions.

We sold SOME books.

We didn't sell ANY books.

Did we sell ANY books?

"Much - Many" lesson please

Sure

OK. The rule is very simple: you should use "many" with countable nouns: "many cats".

And "much" with uncountable nouns: "much water". You can't count water, but can count cats: 1 cat, 2 cats, 3 cats, etc. Do you understand?

Yes

OK, let's start the test

There are so ___ poor people in the world.

many

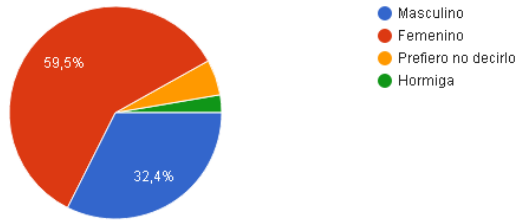
much

Annex 4 Student's Final Questionnaire

Indica tu genero

37 respuestas

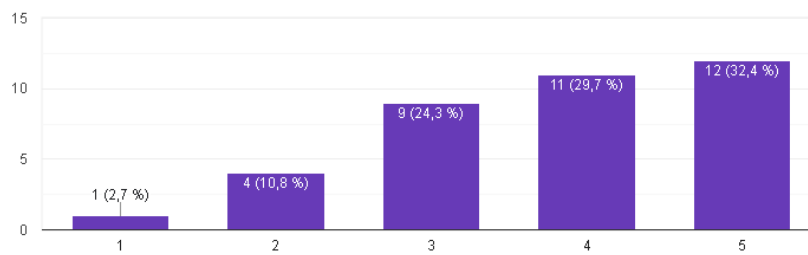
[Copiar](#)



1. ¿Te gusta el inglés?

37 respuestas

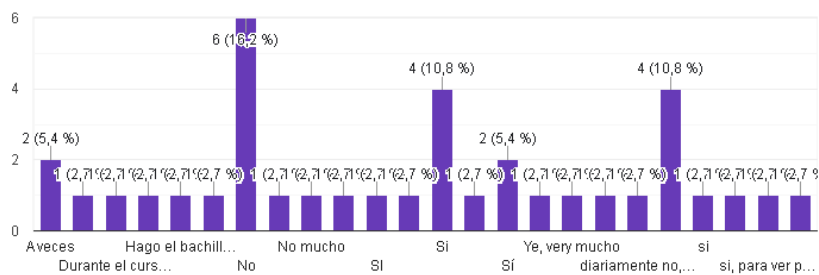
[Copiar](#)



2. ¿Sueles utilizar el inglés en tu vida fuera del colegio? (Academia, ver series, películas, libros, hablar con amigos...)

37 respuestas

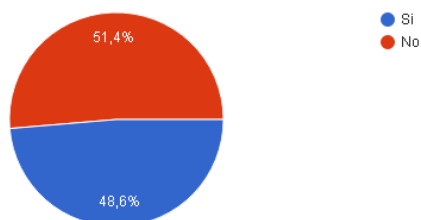
[Copiar](#)



3. ¿Utilizas o has utilizado alguna aplicacion de aprendizaje de idiomas?

37 respuestas

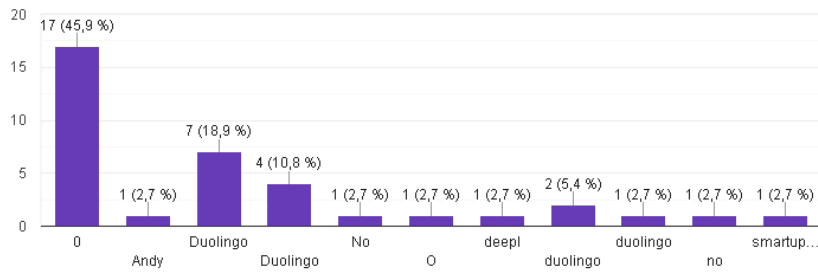
[Copiar](#)



4. En caso positivo, ¿Cual? Si no, pon 0

[Copiar](#)

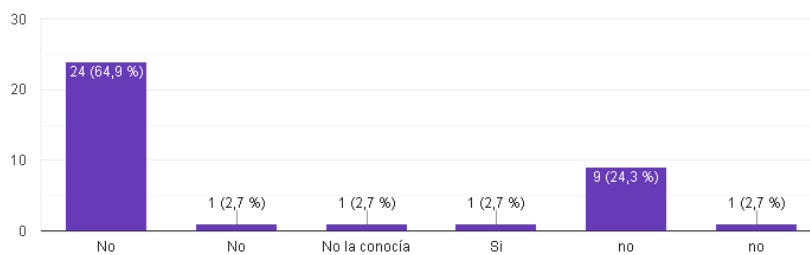
37 respuestas



5. ¿Conocías la aplicación que hemos usado? ANDY

[Copiar](#)

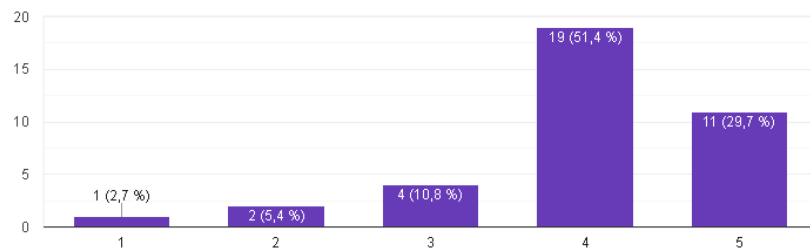
37 respuestas



6. Valora de 1 a 5 el diseño de la aplicación

[Copiar](#)

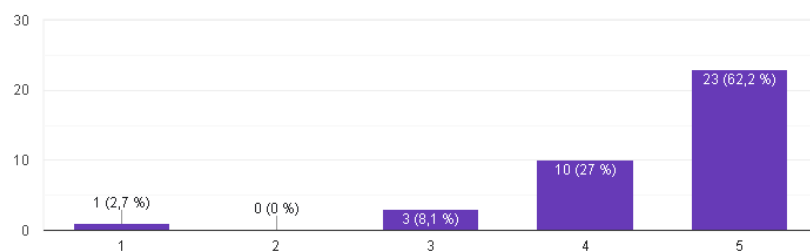
37 respuestas



7. Valora de 1 a 5 la facilidad de los ejercicios realizados.

[Copiar](#)

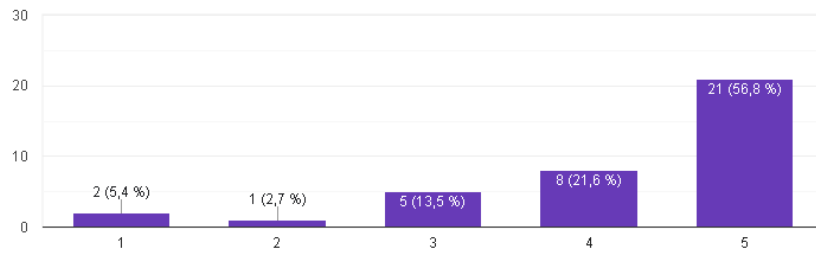
37 respuestas



8. Valora de 1 a 5 las correcciones de la aplicación

 Copiar

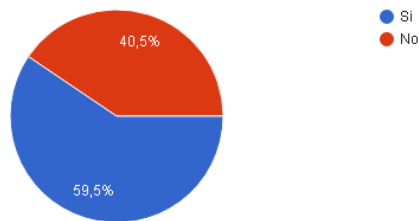
37 respuestas



La conversación con Andy ¿te ha parecido natural?

 Copiar

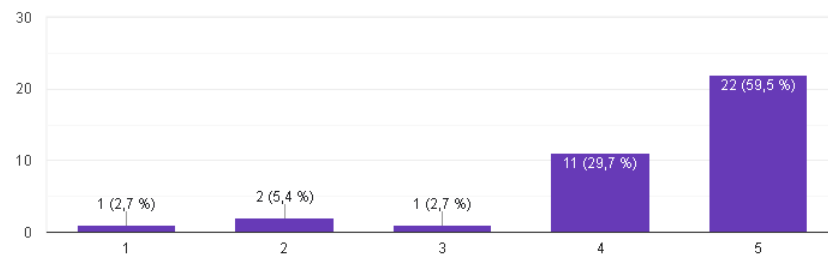
37 respuestas



9. ¿Has encontrado fácilmente los ejercicios propuestos?

 Copiar

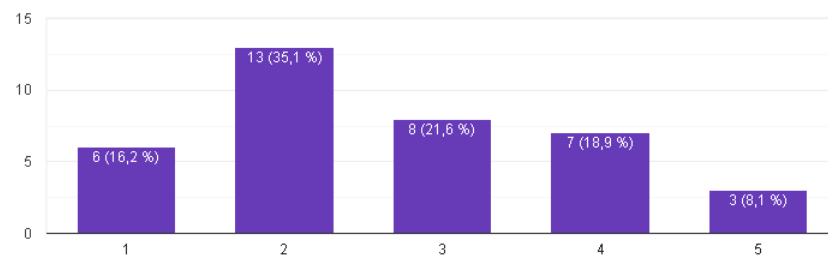
37 respuestas



10. ¿Como de probable es que sigas usando la aplicación para mejorar tu inglés?

 Copiar

37 respuestas



Annex 5
QR presentation

Android & iOS



Prueba previa



Encuesta posterior

