International Journal of English Language and Literature Studies
ISSN(e): 2306-0646
ISSN(p): 2306-9910
DOI: 10.55493/5019.v14i2.5460
Vol. 14, No. 2, 134-149.

URL: www.aessweb.com

Metaphor comprehension in AI and humans: Insights from English language learners and Alexa's NLP capabilities

Esraa Hantouleh[1]
Abdel Rahman Mitib Altakhaineh[2]
Kosay Moneer Alshewiter[3+]

[1,2,3]*University of Jordan, Amman, Jordan.*
[1]*Email: e.hantouleh@ju.edu.jo*
[2]*Email: a.altakhaineh@ju.edu.jo*
[3]*Email: k.alshewiter@ju.edu.jo*

(+ Corresponding author)

Article History
Received: 20 March 2025
Revised: 4 June 2025
Accepted: 23 June 2025
Published: 9 July 2025

Keywords
AI
Alexa
CMT
Metaphor comprehension
Metaphors
NLP.

ABSTRACT

This study investigates the types of English metaphors that Alexa can accurately interpret and compares her metaphor comprehension processes to those of humans. The study involved eighteen English language students from the University of Jordan, all with Academic IELTS band score of 6.5 or above. Data were collected through a metaphor interpretation test administered to human participants and the Alexa virtual assistant. The test consisted of fifteen metaphors divided into three categories which are as follows: five orientational, five ontological, and five structural metaphors. The results revealed that Alexa surpassed the human participants in interpreting metaphors, achieving an accuracy rate of 93.3%, compared to 64.8% for the students. Although Alexa achieved a higher metaphor interpretation rate, she exhibited difficulty with structural metaphors, which involve abstract concepts. Similarly, human participants managed a success rate of 74%, which was the lowest among the three categories. This study indicates that structural metaphors are challenging for both AI and humans. The findings highlight the complexities of interpreting structural metaphors and the limitations shared by humans and AI. These insights contribute to the broader discourse on AI's ability to process figurative language and offer valuable implications for advancements in NLP and human-computer interaction.

Contribution/ Originality: This study investigates the extent to which humans and Alexa can interpret English metaphors accurately employing a metaphor interpretation test. This paper not only fills the existing gap in the literature but also reveals the similarities and differences between human and AI cognitive abilities.

1. INTRODUCTION

The current study is a comparison of Alexa's comprehension of English metaphors with humans' comprehension of them. It highlights the difference between human and AI cognition with a specific emphasis on metaphorical reasoning and interpretation. According to Lakoff and Johnson (2003) this paper explores three types of metaphors, namely, orientational metaphors, structural metaphors and ontological metaphors. They hypothesize that metaphors are not merely rhetorical devices but are deeply ingrained in human cognition, shaping thought and perception. According to Lakoff and Johnson (2003) orientational metaphors are expressions that have special orientations like *happy is up* and *sad is down.* In addition, they introduce the concept of structural metaphors and look at them as ones where one thing is metaphorically structured in terms of another, such as conceptualizing that *ideas are food* (Lakoff & Johnson, 2003). Finally, they introduce ontological metaphors. Those metaphors view things like events and emotions as entities and substances. For example, proposing that *inflation is an entity* gives us a way of referring to it metaphorically (Lakoff & Johnson, 2003).

However, research on AI-powered personal assistants like Alexa remains limited, particularly in the domain of metaphor comprehension. Studies have explored AI's syntactic and semantic processing capabilities while fewer have investigated its ability to handle figurative language within the framework of cognitive semantics. Therefore, testing their ability to decode metaphors provides useful information about such assistants' understanding of metaphorical language. Furthermore, identifying the distinction between human and AI comprehension across different metaphor types enhances the understanding of the contrast between human and artificial intelligence.

Metaphor comprehension is a cognitively demanding process that involves multiple layers of interpretation, requiring individuals to integrate linguistic, contextual, and conceptual knowledge simultaneously (Fuyama, 2023). Artificial intelligence technologies, such as Alexa have been designed to reflect human linguistic abilities. Metaphor comprehension involves different cognitive processes, such as understanding context, evoking emotions, and merging cultural frameworks (Holyoak & Stamenković, 2018). This cognitive activation enables individuals to understand the literal/direct and figurative meanings of metaphors simultaneously (Fuyama, 2023). This process is central for human understanding of metaphors and while similar models have been attempted in AI systems like Alexa, they face significant difficulties. As Maas (2023) pointed out, metaphors do more than organize language; they are integral in shaping and developing a conceptual system for AI. Nevertheless, existing technologies still experience difficulties in integrating the cultural and emotional dimensions that metaphors convey (Dentella, Günther, Murphy, Marcus, & Leivada, 2024). For AI models to accurately understand metaphors, they must connect literal and abstract domains (Barnden & Lee, 2001). Yet, López, Quesada, and Guerrero (2018) and Zhao, Wu, Zhou, Wang, and Jia (2022) argue that AI systems like Alexa demonstrate strength in interpreting literal language but often struggle when it comes to understanding figurative language.

The Conceptual Metaphor Theory (CMT) states that metaphors aim to draw mappings between concrete (source domains) and abstract (target domains) concepts (Lakoff & Johnson, 2003). Consequently, familiarity with these domains is crucial to understand different types of metaphors. Such domain knowledge typically comes from specific shared cultural and contextual experiences and backgrounds which enable individuals in conversations to infer the metaphorical meaning. For AI to interpret metaphors proficiently, it must build these mappings dynamically rather than relying on static lexical associations. Lakoff and Johnson (2003) framework (CMT) offers a theoretical basis for analyzing how Alexa processes structural, ontological, and orientational metaphors and how Alexa's cognitive processes compare to those of humans. Fundamentally, the study seeks to determine how effectively Alexa interprets metaphors and to identify specific instances where its performance is different from that of human participants.

This study aims to develop a set of criteria for evaluating AI metaphor comprehension, examining factors, such as contextual adaptability, accuracy of semantic interpretations, and the system's interpretative flexibility. If Alexa succeeds in guessing the meaning of an orientational, structural, and ontological metaphor in a sentence context, then it can be considered competent in understanding figurative language; otherwise, it is considered less effective compared to human comprehension. This will be explored through investigating the difference in interpreting English metaphors by humans and Alexa. The study aims to answer the following research questions:

1. What types of metaphors (orientational, structural, and ontological) does Alexa handle successfully?
2. How does Alexa's metaphor comprehension compare to humans' cognitive processes?

While some studies have looked into AI's performance in language comprehension, research on metaphor interpretation remains limited. Existing studies have primarily focused on literal language processing, overlooking the intricacies of figurative speech interpretation by AI models. Moreover, a direct comparison between AI and human competence in metaphor comprehension, particularly using Lakoff and Johnson (2003) classification of metaphors remains absent. This study contributes to the broader discourse on AI's potential in mimicking human cognitive processes in natural language understanding by addressing this gap.

2. GENERAL BACKGROUND

2.1. Metaphors in CMT

Metaphors are very commonly used in everyday language (Lakoff & Johnson, 2003) to the extent that we use them without being aware. Lakoff and Johnson (1980) first developed Conceptual Metaphor Theory (CMT) where metaphors were viewed as connections between various concepts. In 2003, the theory was further developed based on findings from cognitive neuroscience. These insights demonstrate how sensorimotor experiences form the basis of metaphorical thinking. Following this advancement, CMT has been extensively adopted in cognitive linguistics, pragmatics, and artificial intelligence research to explore how abstract thought is shaped by embodied experiences. Within the framework of this theory, a metaphor is conceptualized as a mental mechanism that connects a source domain (a concrete domain) and a target domain (a more abstract domain). According to Lakoff and Johnson (2003) in their version of CMT, the cognitive approach to metaphor changed the way people viewed metaphor.

Lakoff and Johnson (2003) categorized metaphors into three categories: First, orientational metaphors, which are space- oriented. This is noted through the use of such words as up, down, front, and back in this metaphor. For example, when someone says they are *feeling down* today it means they are sad (Lakoff & Johnson, 2003). The use of orientational metaphors is grounded in sensorimotor experiences which are closely tied to how humans perceive bodily movements and spatial positioning. Neural evidence proves it. Functional Magnetic Resonance Imaging (FMRI) by Rapp, Leube, Erb, Grodd, and Kircher (2004) shows that metaphors recruit brain regions, including the parietal lobe that are involved in spatial reasoning.

Second, structural metaphors include concepts that are metaphorically structured in terms of others to understand them. For instance, when someone tries to *digest information*, they metaphorically picture understanding information as a physical process. This type of metaphor integrates both abstract and semantic domains (Lakoff & Johnson, 2003). This happens through engaging the left inferior frontal gyrus in the interpretation process. Psycholinguistic research indicates that our brain processes metaphors by activating neural networks connected to the source domain and the target domain at the same time. Cognitive processing of metaphors requires activating schemas in our memory that enable us to make comparisons between different concepts, which makes abstract concepts easier to comprehend (Rapp et al., 2004).

Third, in ontological metaphors, physical characteristics are added to abstract concepts. For example, when someone states that an argument lies *on a shaky foundation*, it means that it is based on weak evidence (Lakoff & Johnson, 2003). The neural correlates of these metaphors suggest that areas of the brain involved in object recognition, like the ventral visual pathway are turned on during their comprehension. Here, we conceptualize abstract domains as concrete objects, using embodied experiences to embed abstract reasoning in tangible frameworks through cognitive processing (Lakoff & Johnson, 2003). Ontological metaphors are particularly significant in AI language processing as they pose a challenge for models that rely on pattern recognition rather than conceptual embodiment. AI systems struggle to assign tangible attributes to abstract entities, often leading to misinterpretations of figurative expressions.

Metaphor comprehension has been a hot topic of interest in cognitive linguistics and psycholinguistics. Most of the recent research on metaphor and other figurative devices comprehension was done with the help of CMT (Aseel Zibin, 2016) which was the first study to be done on this topic. Moreover, more studies started focusing on this issue (Altakhaineh & Shahzad, 2020; Aseel Zibin, Altakhaineh, & Hussein, 2020; A. Zibin & Solopova, 2024). Our brain goes through multiple cognitive processes to interpret each type of metaphor as explained previously gaining huge interest in psycholinguistics. Studies in psycholinguistics point out that this process begins with rapid lexical access (activating literal meaning) through excluding unrelated literal interpretations depending on the context and ends with a complete understanding of metaphorical meanings (Rapp et al., 2004; Shutova, 2010). After discussing how people understand metaphors, we must ask, how different is this process in machines? Do they go through cognitive processes similar to our brains, or are they programmed software that never thinks and still never makes mistakes? Many researchers have attempted to answer these questions by investigating AI's comprehension of metaphors in the following sub-section:

2.2. Alexa's Natural Language Processing

A study by Kumar et al. (2017) provided a comprehensive explanation of how Alexa's machine learning model works although studies investigating Alexa's processing of metaphors are limited. Despite creating a design that enhances Alexa's capabilities, the study provides an extensive overview of how it interprets spoken language. Kumar et al. (2017) use the example of Alexa skills to explain how their model works. Alexa skills are spoken language understanding (SLU) subsystems comprised of automatic speech recognition (ASR) and natural language understanding (NLU) processes. When a speaker talks about a skill, Alexa's ASR recognizes their speech and turns it into text. Then, Alexa classifies the interaction into a specific suitable intent. Furthermore, the order is filled in a slot according to a pre-defined schema by the developers of that skill. Finally, the structured request arrives at the developer, who returns the text to be turned into speech again by Alexa's text-to-speech (TTS) feature, resulting in the playback of an audio response from Alexa and potentially showing graphics on Alexa's screen if the used version has a screen (Kumar et al., 2017).

Hoy (2018) examined how various voice assistants, including Alexa, function. He explained that Alexa, like Google Assistant operates as software on a dedicated speaker device. These assistants record the user's voice and send it to a server once they continuously listen for a particular wake word. The server processes the audio, interprets it as a command, and provides Alexa with the most relevant response. This cloud-based processing allows for rapid responses but also limits Alexa's ability to understand nuanced figurative expressions without explicit programming.

2.3. Previous Studies on AI's Comprehension of Metaphors

The study of metaphor comprehension by AI models has received significant interest in psycholinguistics as more people start using these tools. Following the presentation of Lakoff and Johnson (2003) foundational work on CMT, several studies were conducted to compare human cognitive processes with computational models (Barnden & Lee, 2001; Dentella et al., 2024; Ge, Mao, & Cambria, 2023; López et al., 2018; Shutova, 2010). Barnden and Lee (2001) were among the first to examine how AI systems might process metaphors. Their work uncovered the difficulties that these computational models have in mapping figurative language. A significant problem for future researchers to examine remains open by these models' perceived development of systems that can combine both the concrete and abstract realms to emphasize the metaphors' multiple meanings.

Shutova (2010) advanced this area of inquiry by introducing computational models for metaphor processing through natural language processing (NLP) grounded in Lakoff and Johnson's Conceptual Metaphor Theory (CMT). Her study emphasized the way AI attempts to mimic humans by using large-scale linguistic inputs to interpret metaphors. This study was significant in integrating modern computational models with foundational cognitive theories.

López et al. (2018) examined how voice assistants process metaphorical language, including Alexa in light of growing reliance on AI in domestic contexts. Their findings revealed that such systems primarily depend on rule-based algorithms and data-driven models, making them less effective at handling complex or unfamiliar metaphorical constructs. This study reflects the huge cognitive gap between humans and machines.

Ge et al. (2023) offered a valuable contribution through a comprehensive review investigating AI's engagement with metaphorical language. Their analysis was structured into four distinct phases. In the initial phase, metaphor identification, the AI model recognized figurative language. The researchers observed that these systems often rely on contextual approaches, such as rule-based methods like those suggested by Lakoff and Johnson (2003) with WordNet being a prime example. However, they also discussed modern machine learning and deep learning models like BERT, which are reported to perform better with figurative language in natural language processing.

In the second phase, metaphor interpretation, AI systems aimed to extract metaphorical meaning by employing conceptual mapping akin to Lakoff and Johnson (2003) model or by creating neural embedding to link literal and figurative representations. The third phase was metaphor generation where new metaphors were created using traditional methods like fill-in-the-blank techniques as well as modern methods that identify patterns in large datasets.

The final phase, metaphor application demonstrated how AI employs metaphors across various fields. This included using metaphors in sentiment analysis to make AI systems sound more human, generating creative metaphors for specific sectors like advertising, and aiding language learners in grasping abstract concepts in their target language. Ge et al. (2023) concluded that while this four-stage process represents a notable advance in how machines interpret metaphors, AI still has a long way to go in understanding context and culturally related metaphors. Similarly, Dentella et al. (2024) draw comparable conclusions, highlighting that AI often struggles with figurative meaning.

The development from rule-based approaches to modern neural models in AI for understanding metaphors study has been helpful in understanding the cognitive processes involved. Previous studies have shown that the CMT introduced by Lakoff and Johnson (2003) is a good theoretical framework for the analysis of metaphors in cognitive linguistics and to some extent in computational methods. However, based on the observation made by Ge et al. (2023), this study attempts to investigate the ability of Alexa to understand English metaphors in context and compare it with that of advanced English learners. This paper provides new insights into the similarities and differences between the humans' cognitive processes in comparison to those of the machine by investigating how Alexa interprets different types of metaphors in comparison to advanced English learners. Choosing advanced English learners with a minimum score of 6.5 in the Academic IELTS as a sample for this study (see section 3) guarantees creating a competition between humans and Alexa in this study. As a result, it provides us with a deep understanding of the extent to which machines may or may not surpass our processing abilities when it comes to figurative language. Therefore, this paper contributes to understanding whether Alexa's natural language processing (NLP) model is capable of understanding non-literal language in context or it fails to do so, which means that the creators of the model may need to improve it in the future depending on the findings.

Significant gaps remain in understanding how voice assistants like Alexa comprehend and interpret non-literal expressions despite the growing interest in AI's ability to process figurative language, particularly metaphors. Previous studies by Barnden and Lee (2001), Shutova (2010), López et al. (2018) and Ge et al. (2023) have explored AI metaphor comprehension in general while few have focused specifically on commercially available AI assistants such as Alexa which millions of users interact with daily.

Most existing research has concentrated on rule-based models. Machine learning and deep learning approaches to metaphor interpretation highlight AI's struggles with figurative meaning due to a lack of contextual awareness and conceptual understanding (Dentella et al., 2024; Ge et al., 2023). However, no study has directly compared Alexa's metaphor comprehension to that of human participants, particularly advanced English learners. Such a comparison is crucial in determining whether AI models are approaching human-like metaphor processing or if they still require significant refinement. Moreover, although Kumar et al. (2017) and Hoy (2018) have offered detailed technical explanations of Alexa's NLP framework, there is a lack of research investigating Alexa's ability to process metaphors in real conversational contexts. AI's ability to engage with figurative language is essential for improving natural language interactions, yet no empirical studies have systematically evaluated how Alexa interprets metaphors within psycholinguistic frameworks, such as Lakoff and Johnson (2003) Conceptual Metaphor Theory (CMT).

To address this research gap, the current study aims to examine Alexa's comprehension of English metaphors in context, comparing it with the metaphor interpretation abilities of advanced English learners. This research provides new insights into the strengths and limitations of current NLP models and highlights areas where AI still lags behind human cognition by investigating the similarities and differences between human and AI metaphor processing. The outcomes of this research aim to inform the advancement of AI technologies, suggesting enhancements for voice-activated assistants in their capacity to manage figurative expressions with greater accuracy.

3. METHODOLOGY

3.1. Sample

The sample of this study includes two groups of participants, namely, Alexa as the AI participant and 18 PhD students with a minimum band score of 6.5 or higher in Academic IELTS. The participants were sampled conveniently from a population of 23 students who were accepted to pursue a PhD at the Faculty of Foreign Languages at the University of Jordan, as per statistics provided by the registration unit on the 15th of Dec 2024. The sample covers 78.3% of the whole population which is considered a sufficient percentage for studies that have small populations where surveying at least 75% of them to achieve accuracy (Bartlett, Kotrlik, & Higgins, 2001). The inclusion of PhD students with advanced language proficiency is critical to ensuring that the comparisons made between human comprehension and Alexa's are valid, as it establishes a high benchmark of metaphor comprehension for the human participants.

3.2. Instruments

3.2.1. The Interviews

First, a structured interview with Alexa, Alexa was asked to provide the meaning of 15 metaphors involving Lakoff and Johnson (2003) orientational, structural, and ontological metaphors in contextualized sentences with five instances in each category (see Appendix 1 that presents the test questions and Table 1 on the next page). The interview lasted for almost 15 minutes during which Alexa was asked the questions and was given the time to answer until it finished.

The interview with Alexa was audio recorded. Before conducting the interview, a pilot study with different 3 items that were not included in the interview was conducted to determine the best way to ask Alexa the questions and obtain the answers. During the pilot study, Alexa showed some limitations, such as having to repeat the word *Alexa* before every new question other than that Alexa did not recognize that the question is directed to it. Therefore, every question during the interview was prefixed with the word "Alexa." The pilot study was essential to refining the interview process, addressing Alexa's limitations, and ensuring that the questions were structured in a way that the AI could respond appropriately.

Second, the same type of interview including the same questions was conducted with each of the human participants. Every interview lasted 10-15 minutes on average and all interviews were audio recorded for the purpose of data analysis. Regarding ethical considerations, the participants' consent was gained through disclosing the purpose of collecting their answers and noting the private storage and analysis processes of processing the data. In addition, the participants were asked to sign a consent form found in Appendix 2 which includes a general idea about the research and has a researcher copy and a participant copy in Arabic and English (see Appendix 2). Finally, an IRB approval was granted from the University of Jordan to conduct this particular study and investigate this particular sample. The number of the approval is (356/2025). Three experts in linguistics from the University of Jordan looked at the questions of the interviews to achieve validity for the current paper. The experts confirmed that the instrument is clear, comprehensive, and covers all areas under investigation. The expert review of the interview questions added a layer of validation to the methodology, ensuring that the questions were well-designed and comprehensive in their coverage of the metaphors under study.

3.3. Data Analysis

After recording the interviews, the data collected were partially transcribed by taking notes of each participant's answer to each question while listening to the audio. The notes included whether Alexa and every other participant answered correctly or incorrectly to each question. The data provided by Alexa and human participants during the interviews were analyzed in relation to accuracy. The answers were categorized as *successful* or *unsuccessful* based on whether the meaning of each metaphor was guessed correctly. This categorization relied solely on Lakoff and Johnson (2003) explanation of the meaning of every metaphor investigated in this paper. In other words, if Alexa or one of the other participants answered with the correct meaning of the metaphor in question correctly within the context where it's mentioned then their answer was marked as successful, other than that it was marked as unsuccessful. Then, the percentage of Alexa's successful and unsuccessful interpretations was calculated. Moreover, frequencies and percentages of humans' correct answers were calculated. Finally, a calculation of the total percentage as explained by Triola (2015) was done to calculate the overall percentage of correct answers by humans (see section 4). The focus on categorizing responses as either successful or unsuccessful based on Lakoff and Johnson (2003) framework ensures that the analysis remains consistent with the established conceptual metaphor theory and allows for a clear comparison between AI and human interpretations.

3.4. Limitations

A notable limitation of the present study is the analysis of only 15 metaphors, 5 from each type of metaphor presented by Lakoff and Johnson (2003) namely structural, orientational, and ontological metaphors. Table 1 shows the types of metaphors and the examples on each that are investigated in this paper.

Table 1. Types of metaphors in CMT and instances of them under investigation

Number	Types of metaphors	Metaphors under investigation
1.	Structural metaphors	a. Shot down my argument.
		b. Wasting my time.
		c. Digest this new information.
		d. Hit a dead end.
		e. Playing by my own rules.
2.	Orientational metaphors	a. Feeling down.
		b. On top of things.
		c. Gone up.
		d. Look forward.
		e. Looking up.
3.	Ontological metaphors	a. Shaky foundation.
		b. Full of brilliant ideas.
		c. Hold onto this anger.
		d. Stuck in a boring meeting.
		e. Poured his heart.

While 15 metaphors is a sufficient number to achieve saturation defined as a point where more data will not further develop the findings but rather repeat what has been previously revealed (Glaser, Strauss, & Strutzel, 1968) testing more examples of each type of metaphor may provide more generalizable findings. However, the use of 15 metaphors in context was influenced by practical constraints, including the time frame within which the study was conducted and the limited duration of Alexa's session (the duration of the machine's focus span on one topic). This limited scope may influence the generalizability of the findings.

4. FINDINGS AND DISCUSSION

4.1. Findings

4.1.1. Alexa's answers

A test was administered to both humans and Alexa to investigate the metaphors it handles successfully as described in the methodology section. Table 2 shows the frequency and percentage of Alexa's correct interpretations of metaphors.

Table 2. Frequency and percentage of Alexa's correct interpretations

Number	Types of metaphors	Metaphors under investigation	Alexa's interpretation
1.	Structural metaphors	a. Shot down my argument.	Unsuccessful
		b. Wasting my time.	Successful
		c. Digest this new information.	Successful
		d. Hit a dead end.	Successful
		e. Playing by my own rules.	Successful
2.	Orientational metaphors	a. Feeling down.	Successful
		b. On top of things.	Successful
		c. Gone up.	Successful
		d. Look forward.	Successful
		e. Looking up.	Successful
3.	Ontological metaphors	a. Shaky foundation.	Successful
		b. Full of brilliant ideas.	Successful
		c. Hold onto this anger.	Successful
		d. Stuck in a boring meeting.	Successful
		e. Poured his heart.	Successful
Frequency of correct interpretations			14
Percentage of correct interpretations			93.3%

4.1.2. Human Participants' Answers

After calculating the frequency and percentage of Alexa's successful interpretations of metaphors, the same test was employed to explore humans' competence in interpreting metaphors. Table 3 shows the frequency and percentage of humans' correct interpretations.

Table 3. Frequency and percentage of humans correct interpretations

Number	Types of metaphors	Metaphors under investigation	Students' interpretation	
			Frequency of correct interpretations	Percentage of correct interpretations
1.	Structural metaphors	a. Shot down my argument.	12	67%
		b. Wasting my time.	16	89%
		c. Digest this new information.	17	94%
		d. Hit a dead end.	10	56%
		e. Playing by my own rules.	12	67%
2.	Orientational metaphors	a. Feeling down.	12	67%
		b. On top of things.	9	50%
		c. Gone up.	17	94%
		d. Look forward.	16	89%
		e. Looking up.	17	94%
3.	Ontological metaphors	a. Shaky foundation.	17	94%
		b. Full of brilliant ideas.	18	100%
		c. Hold onto this anger.	9	50%
		d. Stuck in a boring meeting.	12	67%
		e. Poured his heart.	13	72%

4.1.3. Total Percentage of Correct Answers Calculation

After calculating the frequency and percentage of correct metaphor interpretations by each of the human participants individually, we need to calculate the overall percentage of the students' correct answers to the test in order to compare this percentage with Alexa's percentage of correct interpretations. Triola (2015) percentage calculation approach allows us to do so using the following formula:

Percentage= (total correct Answers/ total possible answers) ×100

The calculation of the total correct answers as per Table 3 is as follows:

12 + 16 + 17 + 10 + 12 + 12 + 9 + 17 + 16 + 17 + 17 + 18 + 9 + 12 + 13 = **175**

Calculation of the total possible answers:

Total possible answers= total number of interviewees × total number of questions

= 18 × 15 = 270

Total percentage= (175/270) × 100 = 0.648 × 100 = 64.8%

Based on the above, Alexa succeeded in interpreting 93.3% of the examined metaphors while the sampled students were able to interpret 64.8% of them. Alexa showed high competence in comprehending the three types of metaphors. In contrast, the sampled students, who are advanced learners of English showed much lower competence in doing so. Orientational metaphors (feeling down, on top of things, gone up, look forward, and looking up) were the most successfully understood by both Alexa and humans due to their relatively straightforward conceptual mappings (Lakoff & Johnson, 2003). On the other hand, structural and ontological metaphors revealed variations. Examples of these findings and a discussion of these results are presented in the following sub-section.

4.2. Discussion

4.2.1. The Findings of Alexa's Answers

As previously mentioned, a structured interview including 15 questions on metaphors in context was conducted with Alexa. Alexa is highly efficient in interpreting figurative language based on the results obtained from this interview, which demonstrate Alexa's competence in interpreting metaphors with a 93.3% accuracy rate. For example, when asked, *what does the metaphor in I'm feeling down today, but yesterday I was on top of the world mean?* Alexa answered, "Sadness. "Feel down" is a verb usually defined as to feel depressed or unhappy." Alexa's proficiency in all of the orientational metaphors stems from their reliance on spatial relationships, which are easily encoded in AI systems (Lakoff & Johnson, 2003). This success may reflect programming efficiency rather than genuine comprehension. Shutova (2010) highlighted this point of view, proposing that machine learning models, Alexa being one of them, employ NLP to mimic humans depending on large-scale linguistic inputs. Another example is when Alexa was asked, *what does the metaphor in I am playing by my own rules to win at life mean?* The answer was, "Belt conveyor. From forbes.com: You are now playing by your rules and not allowing yourself to be put on a conveyor belt of someone else's expectations." Hoy (2018) notes that Alexa, among other voice assistants is always connected to the Internet which means that all interactions with the device are sent to a central computing system that processes the user's voice commands and provides the assistant with the proper response. Shutova (2010) and Ge et al. (2023) studies on AI's processing of metaphors support the concept that these devices depend on NLR, which Kumar et al. (2017) confirmed that Alexa uses among other deep learning models like BERT to process users' requests and commands. Kumar et al. (2017) state that after commands are processed and Alexa is ready to respond, Alexa's TTS feature turns its response into speech again and plays it as audio for the user. These findings prove that Alexa's system somehow simulates or mimics our brain's cognition in terms of having a clear procedure or process to interpret metaphors.

However, Alexa failed to interpret the structural metaphor, *shot down my argument* successfully and provided a unrelated interpretation that is not literal despite being highly competent in comprehending all metaphors under investigation, including structural ones. At the same time, it is not metaphorical or figurative. When asked, *what does the metaphor, He shot down my argument with solid evidence mean?* Alexa answered with, "Argument. From thesaurus net: When someone is said to "shoot holes in" a particular argument or idea, it signifies their ability to successfully discredit or undermine its credibility by providing counterexamples, contradictory evidence, or rational objections." Alexa was asked this question three times during the pilot, and it kept giving the same exact answer. The AI-powered voice assistant did not recognize the metaphor in the sentence and answered with an explanation of a different expression.

The expression *shoot holes in something* was not found as an entry in the Cambridge online dictionary at all. Instead, an entry for *pick holes in something* was found categorized as an idiom and defined as finding mistakes in someone's speech or actions to show that it is incorrect or not good (Cambridge Dictionary, 2024). Although close in meaning, these two expressions remain different in meaning. *Shooting down an argument* means to refute it Lakoff and Johnson (2003). *Shooting holes in an argument* is a combination of metaphor. Alexa was asked about an idiom that refers to pointing out flaws in it to weaken it. Structural metaphors, such as "shot down my argument," require abstract reasoning and cultural context which are beyond Alexa's current capabilities. This shortcoming of Alexa supports López et al. (2018) argument that voice assistants although attempting to mimic humans' cognitive processes still have to refer to predefined rules and patterns that limit their ability to comprehend complex or more abstract metaphors.

From these findings, it becomes clear that while AI systems like Alexa are highly competent at interpreting figurative language, they still fall short when trying to interpret abstract metaphors. This is justified by Shutova (2010), López et al. (2018) and Fuyama (2023) with the AI's dependence on the datasets it has rather than thinking about metaphors. On the other hand, human metaphor comprehension makes use of context, culture, emotional resonance, and cognitive flexibility to interpret metaphors (Lakoff & Johnson, 2003). It is argued that to interpret the meaning of metaphors, we integrate context and activate certain areas in our brains, like the inferior frontal and middle temporal gyri (Rapp et al., 2004).

4.2.2. The Responses Provided by Human Participants

The responses introduced in section 3 were distributed among PhD students of English Literature and English Linguistics who achieved a band score of at least 6.5 in the academic IELTS and showed a variation in processing metaphors. The sample's accuracy in metaphor comprehension was 64.8%, distributed in different percentages depending on the type of metaphor. For instance, orientational metaphors were the most successfully comprehended by the sampled students, with a success rate of nearly 79%, meaning that approximately 71 out of 90 answers to these metaphors were correct. For example, 17 out of the 18 students (94% of them) answered correctly to the metaphors *gone up* and *looking up*, the answers were all given the meanings of "increased significantly" and "improving or becoming better," respectively. In contrast, the least successfully interpreted metaphor in this category was *on top of things* with 50% of the students misinterpreting it into things like "I am overwhelmed with work tasks" (33%) and "I am the supervisor" (17%) while the others provided the correct meaning, which is "I am in control of the situation at work." Lakoff and Johnson (2003) note that the use of this type is rooted in sensorimotor experiences, recruiting brain regions like the parietal lobe, which is responsible for spatial reasoning. Rapp et al. (2004) FMRI experiments conclude that areas like the parietal lobe in the brain are responsible for understanding this type of metaphors as they are specialized in spatial reasoning. This explains finding out that most of the human participants succeeded in interpreting this type of metaphor.

The second most successfully interpreted type of metaphor by humans is ontological metaphors with 69 correct answers out of a total of 90 answers, resulting in an approximate accuracy rate of 77%. For example, all 18 respondents answered correctly when asked about the meaning of the metaphor *full of brilliant ideas* in the context of *his head is full of brilliant ideas*, all answering with things that give the meaning of "his mind is rich with creative thoughts." On the lower end, *hold onto anger* was the most misinterpreted metaphor in this category with half of the participants providing the correct answer, "retain feelings of anger " while the rest mainly answered, "letting go of anger" and "inability to control anger" in equal percentages of 19.5% (4 answers) each. Finally, 2 participants (11%) provided the literal answer of physically holding anger (Alazazmeh & Zibin, 2023). The neural correlates of these metaphors demonstrate that object recognition occurs through activating the ventral visual pathway in order to achieve comprehension. Here, we process abstract domains as if they were concrete object domains (Lakoff & Johnson, 2003). Rapp et al. (2004) note that cognitive processing of such metaphors requires activating schemas in our memory that enable us to comprehend abstract concepts easily. This explains the small number of participants (2) who interpreted these ontological metaphors literally.

The least successfully interpreted metaphor type by human respondents is the structural metaphors with around 74% of the participants choosing the correct interpretation, which means that 67 out of 90 answers to the questions on this type of metaphor were correct. An example of this is when 17 out of 18 participants (94%) answered correctly when asked about the meaning of digesting *new information*, answering, "I need to process and understand the information." On the other hand, only 56% (10 out of 18 participants) of the students answered correctly to the meaning of *"hit a dead end"* with "our relationship has stopped progressing," while the rest of the answers were "our relationship is stuck in a dead-end road" and "our relationship is no longer exciting."

Lakoff and Johnson (2003) suggest that understanding this type of metaphor involves our brain connecting abstract ideas with concrete meanings. Rapp et al. (2004) provide neurological evidence arguing that the left inferior frontal gyrus becomes active when we interpret metaphors. This indicates that certain memory frameworks are triggered during metaphor comprehension, helping us relate different concepts. It is a process extremely tied to how we grow cognitively. Children first understand their physical surroundings, and over time, they begin to relate these basic experiences to more abstract ideas through metaphorical thinking (Lakoff & Johnson, 2003). In this sense, metaphor comprehension develops progressively as we build up more complex mental structures from simpler ones.

Research into human metaphor comprehension lends strong support to Lakoff and Johnson (2003) neural theory of metaphor. According to this view, metaphors are shaped by physical neural pathways that connect distinct regions of the brain, each responsible for different conceptual areas. These links are formed through repeated co-activation of related experiences, enabling the transfer of meaning from familiar, concrete domains to more abstract ones. It is a flexible, situation-sensitive process. Our interpretations often shift depending on the immediate context and our lived experiences.

This view is supported by Rapp et al. (2004) FMRI study, which explored how the brain responds to both literal and metaphorical sentences. They found that the left inferior frontal gyrus showed higher levels of activity when participants processed metaphors, proposing this area plays a key role in drawing semantic connections across conceptual domains. It highlights the collaborative effort of various brain regions in interpreting figurative language, particularly those involved in language use and complex reasoning.

Further insights from Rapp et al. (2004) and Shutova (2010) show that metaphor comprehension starts with quick access toward meanings. Initially, we tap into the literal sense and then gradually rule out irrelevant interpretations based on the context before arriving at the metaphorical meaning. This filtering process often includes emotional and social cues, especially in metaphors rooted in shared cultural experiences or personal memory. These findings make it clear that metaphor understanding goes beyond basic semantic processing other cognitive factors like emotion, memory, and social awareness are also deeply involved.

5. CONCLUSION AND RECOMMENDATIONS

Alexa and humans depend on complex cognitive processes to comprehend metaphors based on the results of the study. Alexa surprisingly surpasses advanced English language learners' abilities in interpreting metaphors.

Alexa's machine learning model is highly efficient as it refers to both memory and internet search to provide the user with correct interpretations. However, it faces difficulties when it comes across metaphors that require abstract analysis and the use of context to be comprehended. This highlights the importance of context-awareness in metaphor interpretation, something that human cognition typically excels at but which remains a challenge for AI systems.

Regarding humans' metaphor comprehension competence, it was significantly lower than that of Alexa's, although they are highly proficient in English. The study indicates that it is easier for humans to comprehend orientational metaphors than to process structural metaphors. It seems that both Alexa and humans face challenges in understanding structural metaphors which indicates that they may share some cognitive processes. This suggests that human cognition struggles with interpreting abstract concepts just like Alexa. However, human cognition might still possess a certain level of flexibility and nuance in processing metaphors influenced by social and emotional context which may not be fully replicated in Alexa's AI model. This is due to the possibility of factors like the participants' native language, being Arabic in this study, intervening in the interpretation of metaphors. The study might also benefit from considering the participants' cultural backgrounds as different cultures may have different metaphorical constructions. Especially given that English and Arabic are from two separate language families and have completely different structures. It is recommended to study more instances of metaphors from all categories to achieve a more comprehensive understanding of where exactly the weaknesses lie. In addition, analyzing the effectiveness of various types of machine learning models in interpreting metaphors across languages and cultures would provide valuable insights. It is also recommended to investigate the comprehension of the same metaphors used in this study among native speakers of English to discover whether they achieve the same success rate as Alexa or are more or less competent. Further research could also explore the role of emotional and contextual cues in both human and AI metaphor comprehension to better understand the cognitive mechanisms at play.

Funding: This study received no specific financial support.
Institutional Review Board Statement: The Ethical Committee of the University of Jordan, Jordan has granted approval for this study on 4 February 2025 (Ref. No. 356/2025).
Transparency: The authors state that the manuscript is honest, truthful, and transparent, that no key aspects of the investigation have been omitted, and that any differences from the study as planned have been clarified. This study followed all writing ethics.
Competing Interests: The authors declare that they have no competing interests.
Authors' Contributions: All authors contributed equally to the conception and design of the study. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Alazazmeh, H. M., & Zibin, A. (2023). The conceptualization of anger through metaphors, metonymies and metaphtonymies in Jordanian Arabic and English: A contrastive study. *Cognitive Semantics, 8*(3), 409-446. https://doi.org/10.1163/23526416-bja10037

Altakhaineh, A. R. M., & Shahzad, N. M. (2020). Using pictures in teaching metaphorical expressions to Arabic-speaking EFL learners. *The Asian Journal of Applied Linguistics, 7*(1), 32-44.

Barnden, J. A., & Lee, M. G. (2001). *Metaphor and artificial intelligence*. East Sussex, UK: Psychology Press.

Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal, 19*(1), 43-50.

Cambridge Dictionary. (2024). *Pick holes in something. In Cambridge English dictionary*. Retrieved from https://dictionary.cambridge.org/vi/dictionary/english/pick-holes-in. [Accessed December 20, 2024]

Dentella, V., Günther, F., Murphy, E., Marcus, G., & Leivada, E. (2024). Testing AI on language comprehension tasks reveals insensitivity to underlying meaning. *Scientific Reports, 14*, 28083. https://doi.org/10.1038/s41598-024-79531-8

Fuyama, M. (2023). Does the coexistence of literal and figurative meanings in metaphor comprehension yield novel meaning?: Empirical testing based on quantum cognition. *Frontiers in Psychology, 14*, 1146262. https://doi.org/10.3389/fpsyg.2023.1146262

Ge, M., Mao, R., & Cambria, E. (2023). A survey on computational metaphor processing techniques: From identification, interpretation, generation to application. *Artificial Intelligence Review, 56*, 1829-1895. https://doi.org/10.1007/s10462-023-10564-7

Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing Research, 17*(4), 364. https://doi.org/10.1097/00006199-196807000-00014

Holyoak, K. J., & Stamenković, D. (2018). Metaphor comprehension: A critical review of theories and evidence. *Psychological Bulletin, 144*(6), 641–671. https://doi.org/10.1037/bul0000145

Hoy, M. B. (2018). Alexa, Siri, Cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly, 37*(1), 81-88. https://doi.org/10.1080/02763869.2018.1404391

Kumar, A., Gupta, A., Chan, J., Tucker, S., Hoffmeister, B., Dreyer, M., . . . Rastrow, A. (2017). Just ASK: Building an architecture for extensible self-service spoken language understanding. *arXiv preprint arXiv:1711.00549*. https://doi.org/10.48550/arXiv.1711.00549

Lakoff, G., & Johnson, M. (1980). Conceptual metaphor in everyday language. *Journal of Philosophy, 77*(8), 453–486.

Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago, IL: University of Chicago Press.

López, G., Quesada, L., & Guerrero, L. A. (2018). *Alexa vs. Siri vs. Cortana vs. Google assistant: a comparison of speech-based natural user interfaces*. Paper presented at the Advances in Human Factors and Systems Interaction: Proceedings of the AHFE 2017 International Conference on Human Factors and Systems Interaction, July 17− 21, 2017, The Westin Bonaventure Hotel, Los Angeles, California, USA 8.

Maas, M. M. (2023). AI is like… A literature review of AI metaphors and why they matter for policy. *AI Foundations Report, 2*.

Rapp, A. M., Leube, D. T., Erb, M., Grodd, W., & Kircher, T. T. (2004). Neural correlates of metaphor processing. *Cognitive Brain Research, 20*(3), 395-402. https://doi.org/10.1016/j.cogbrainres.2004.03.017

Shutova, E. (2010). *Models of metaphor in NLP*. Paper presented at the Proceedings of the 48th annual meeting of the association for computational linguistics.

Triola, M. F. (2015). *Elementary statistics*. Boston, MA: Pearson Education.

Zhao, J., Wu, M., Zhou, L., Wang, X., & Jia, J. (2022). Cognitive psychology-based artificial intelligence review. *Frontiers in Neuroscience, 16*, 1024316. https://doi.org/10.3389/fnins.2022.1024316

Zibin, A. (2016). The comprehension of metaphorical expressions by Jordanian EFL learners. *Sage Open, 6*(2), 2158244016643144.

Zibin, A., Altakhaineh, A. R. M., & Hussein, E. T. (2020). On the comprehension of metonymical expressions by Arabic-speaking EFL learners: A cognitive linguistic approach. *Topics in Linguistics, 21*(1), 45-61.

Zibin, A., & Solopova, O. A. (2024). Metaphors across languages, cultures and discourses: A research agenda. *Russian Journal of Linguistics, 28*(1), 7-32. https://doi.org/10.22363/2687-0088-37837

APPENDICES

Appendix 1. Metaphor interpretation test.

Structural metaphors

1. What does the metaphor in *He shot down my argument with solid evidence* mean?
2. What does the metaphor in *You're wasting my time with your idle talk* mean?
3. What does the metaphor in *I need to digest this new information before deciding what to choose* mean?
4. What does the metaphor in *We've hit a dead end in our relationship* mean?
5. What does the metaphor in *I am playing by my own rules to win at life* mean?

Orientational metaphors

6. What does the metaphor in *I'm feeling down today, but yesterday I was on top of the world* mean?
7. What does the metaphor in *I am actually on top of things at work* mean?
8. What does the metaphor in *Prices have gone up dramatically recently* mean?
9. What does the metaphor in *We both need to look forward and stop dwelling on the past* mean?
10. What does the metaphor in *Things are finally looking up after a tough year* mean?

Ontological metaphors

11. What does the metaphor in *Their argument rests on a shaky foundation* mean?
12. What does the metaphor in *His head is full of brilliant ideas* mean?
13. What does the metaphor in *I cannot hold onto this anger any longer* mean?
14. What does the metaphor in *I am stuck in a boring meeting* mean?
15. What does the metaphor in *He poured his heart into his project* mean?

Appendix 2. Consent form

عمادة البحث العلمي
DEANSHIP OF SCIENTIFIC RESEARCH
Deanship of Scientific Research

الجامعة الأردنية
THE UNIVERSITY OF JORDAN
عمادة البحث العلمي

Participant's copy

Consent and Authorization for Adult Participation in the Study
Metaphor Comprehension in Alexa's Talk

You are invited to participate in a research study led by researchers at the **School of Foreign Languages / The University of Jordan**. Please read this document before deciding whether to participate.

- **What is the purpose of this study?**

This study aims to: First, identify the types of metaphorical expressions that Alexa can successfully process. Second, understand how Alexa's comprehension compares to human comprehension. Third, explore how Alexa's limitations in understanding metaphorical expressions affect interactions between Alexa and its users.

- **What is a metaphorical expression?**

These are expressions where a word or phrase is used to refer to something or an action to which it does not literally apply, thus carrying symbolic or comparative meaning.

- **What is your role?**

You will be asked to perform tasks that include answering questions about the meanings of metaphorical expressions presented to you.

- **What are the direct benefits of your participation?**

There are no direct benefits. However, your participation may contribute to a better understanding of how artificial intelligence devices comprehend metaphorical expressions in comparison to humans.

- **Are there any risks involved?**

There are no risks. Additionally, all the information you provide in this form will remain confidential.

- **How will the data/samples be collected?**

Through a test that measures your comprehension of metaphorical expressions.

- **How will the data/samples be stored?**

By recording responses without including any identifying information about you as a participant.

- **What if we need additional information?**

We will not require any additional information.

- **What if you decide to withdraw from the study later?**

You can inform us by phone or email (provided below), and you will not suffer any physical or moral harm.

عمادة البحث العلمي
DEANSHIP OF SCIENTIFIC RESEARCH
Deanship of Scientific Research

الجامعة الأردنية
THE UNIVERSITY OF JORDAN
عمادة البحث العلمي

- **What if you have any questions or inquiries?**

You may contact any of the following individuals:
Esraa Hantouleh e.hantouleh@ju.edu.jo +962799161251
Or the study supervisor, **Dr. Abdelrahman Al-Takhaineh**, Professor at the School of Foreign Languages / The University of Jordan:
Phone: +962775338655 (Extension: 24777)
Email: a.altakhaineh@ju.edu.jo

عمادة البحث العلمي
DEANSHIP OF SCIENTIFIC RESEARCH
Deanship of Scientific Research

الجامعة الأردنية
THE UNIVERSITY OF JORDAN
عمادة البحث العلمي

Participant's number: ...

Researcher's copy

I, the undersigned, hereby consent to participate in the study titled (Metaphor Comprehension in Alexa's Talk). I am aware of the procedures involved, the potential benefits, and the possible risks. I am also informed about the individuals I can contact if I have any questions or concerns. I have received a signed copy of this consent form.

□ I agree to participate in the study.
□ I agree to allow the researchers to use my responses and data in future research studies, while retaining all my rights, provided that the necessary approvals are obtained from research ethics committees.

…………………	……………………………………	……………………………………
Date	Signature	Participant's name

……………………………………	……………………………………
Alternative phone number	Phone number
……………………………………	……………………………………
Phone number	A phone number to call if you are not reachable on your main number

…………………	……………………………………	……………………………………
Date	Signature	Witness's name
…………………	……………………………………	……………………………………
Date	Signature	Researcher's name