INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY COMPREHENSIVE RESEARCH

Speech Acts in Virtual Assistants: A Comparative Study of Siri, Alexa, and Google Assistant

Saeeda Hasan Khalid Alkhalidi [1*], Amal Hadi Alkhalidi [2], Noor Salim Hashim Tahir Al-Asadi [3]
[1-3] Department of English Language, College of Education for Women, Kufa University, Iraq

* Corresponding Author: **Saeeda Hasan Khalid Alkhalidi**

Article Info

ISSN (online): 2583-5289
Volume: 04
Issue: 06
November - December 2025
Received: 05-09-2025
Accepted: 07-10-2025
Published: 04-11-2025
Page No: 01-04

Abstract
Virtual helpers like Siri, Alexa, and Google Assistant are now a big part of everyday life. They do things by listening to your words and responding to normal language. Even though these systems are getting better at recognizing speech and making responses, one important area that needs more research is how they can understand and make meaningful speech acts. This paper uses speech act theory to look at how well these helpers can recognize and carry out different kinds of illocutionary acts, such as spokespeople, orders, and commissives. The study looks at Austin's and Searle's basic theories to compare how each helper handles and reacts to different practical tasks. This study uses planned trials with 60 speech act-based questions from real-life usage situations to test the helpers' ability to keep pragmatic consistency, figure out the goal, and give the right context. They show that the processes are not at all the same. When it comes to giving information and directions, Google Assistant is the best. When it comes to action-based orders, Alexa is the best. Siri is good at giving answers that show how you feel, but it's not great at acts that aren't clear or straight. They show how important pragmatic analysis is for judging AI conversation and give us thoughts on how to improve natural language tools. By combining language theory with real-world AI testing, this study adds to the fields of human-computer interaction, AI pragmatics, and discourse technology.

DOI: https://doi.org/10.54660/IJMCR.2025.4.6.01-04

Keywords: Alexa, Google Assistant, Pragmatics, Siri, Speech Acts

Introduction

Virtual helpers like Siri, Alexa, and Google Assistant have changed the way people and computers talk to each other by letting people use normal words to control gadgets, get information, and handle chores. These systems have changed over time from simple order processors to talking bots that can respond to users based on their current situation. But the fact that they can communicate brings up important language and functional questions, especially about how they do and understand speech acts. Speech act theory, which was first presented by Austin in 1962 and built upon by Searle in 1969, helps us understand how words and phrases work as actions as well as ways to share information. People use words every day to ask for things, make promises, say sorry, give orders, and more. Each of these uses is a speech act. How well these acts are done varies a lot on the situation, what the speaker meant, and how the viewer interprets it. It's important to test virtual helpers' ability to recognize and carry out these actions as they try to mimic human speaking skills.

Natural language processing (NLP) has come a long way, but virtual assistants still don't fully understand what people are saying, especially when the language isn't clear, is used in a social setting, or is unclear. Someone who hears a user say, "It's too cold in here," might think that the user wants the temperature to be raised. The virtual helper, on the other hand, might not know this as an order unless it is made to do so. Because of these issues, AI chat models need to do a better job of considering how people really think.

The goal of this study is to compare how well Siri, Alexa, and Google Assistant handle different types of speech acts and how right, attentive, and good their answers are. According to the study, speaking can be broken down into five main types:

declaratives, expressives, directions, and reps. Once that's done, it checks how well each helper does in each part. To find out what the best and worst things about the most popular virtual helpers are, this study uses both language theory and real-life tests. The results not only help us make AI chat systems better, but they also show how important language ideas are for making computers talk to people more easily.

Speech Act Theory

Austin came up with speech act theory in 1962, and Searle built on it in 1969. This theory helped people see language as more than just a transmission of information. Austin separated speech acts into three groups: acts that change words, acts that don't change words, and acts that change the meaning of words. You say things, but they don't come out right. This is called illocutionary acts. And perlocutionary acts are how those things make you feel. Searle says that illocutionary acts can be divided into five main groups: directions (requests, orders), expressives (thanks, apologies), mandatives (requests, statements), and representations (statements, descriptions). Putting these things together has been very important in pragmatics and speech analysis ever since. It has also had an effect on studies in human-computer interaction (HCI).

When it comes to virtual helpers, speech acts are the most important way to talk to each other. It was hard for systems to understand natural language (NLU) in the early days of human-computer interaction (HCI). These systems had to not only recognize words but also figure out what the speaker meant (Winograd, 1972; Allen, 1995) [16, 1]. This is still a big problem with voice assistants today, because users want answers that are social and relevant to their needs (Kepuska & Bohouta, 2018) [8].

A new study shows that the actual skills of virtual helpers affect how well they do their jobs. Porayska-Pomsta *et al.* (2013) [11] say that chat systems often get indirect speech acts wrong, which makes them give replies that are wrong or not important. A study by Liao *et al.* (2020) [9] found that voice assistants like Siri and Alexa don't always understand how to be polite and make requests without directly asking for something. But Google Assistant does better in these areas because it has more data to learn on and more advanced modeling of its surroundings (Huang *et al*., 2021) [6]. Things that are hard to understand in AI are made even harder by the fact that human language is naturally unclear and changes depending on the situation. Clark (1996) [5] says that for a talk to work, both people need to know the same things about the past and understand what's going on. These are hard for models that learn on their own to keep track of. Big language models are very good at using language, but Bender and Koller (2020) [3] made a strong case that they don't really understand what it means in real life. This is why they often come up with results that make sense but aren't very deep.

In 2003, Traum and Larsson also looked into how speech act theory could be used in AI. To make the systems more adaptable, they came up with official way to talk and act. These models help sort users' goals into groups and let them connect in a more natural way. In the same way, Su *et al.* (2019) [13] used reinforcement learning to make users happier by improving helper answers based on understanding how people talk.

Review of virtual helpers that have already been done are used for the comparison part of this study. Hoy (2018) [7] talked about what voice helpers can do now and how they have changed over time. He talked about changes in the way tools are made and how people use them. Siri, Alexa, and Google Assistant are all different, according to studies (Lopatovska & Oropeza, 2018; Chung *et al*., 2020) [10, 4]. For example, Siri is great at linking devices together, Alexa is great at following smart-home directions, and Google Assistant is great at finding things and figuring out what's going on around it.

But there are still issues with how well the helpers can handle complex illocutionary acts like sass, humor, or minor expressives, which are needed for interaction that feels like talking to a person (Wang *et al*., 2022) [15]. To close this gap, it is important to use pragmatic theory and machine learning to build AI that takes words into account.

Virtual helpers are also slowly getting better at what they do, but the study shows that they still have trouble being useful. This study builds on that work by looking at how the aids used speech acts in a variety of tasks.

Background of the Study

A lot of research has been done on how well virtual helpers do with everyday jobs, especially when it comes to how well they understand and respond to speech acts. We can learn from these studies what systems like Siri, Alexa, and Google Assistant can and can't do at this time.

The 2018 study by Lopatovska and Oropeza was one of the first to look at how people use Siri and Alexa next to each other. They looked at how the helpers answered certain types of speech, like pleas and directions. Based on their study, Alexa was better at doing things like controlling smart home devices, while Siri was better at talking but not as good at doing what you say. It wasn't possible to look at Google Assistant in this study because not many people used it at the time.

Chung *et al.* (2020) [4] added Google Assistant to the comparison and tried the helpers in more speech acts, such as ones that are emotional or meant to make up for something. They had people in their study act out planned scenarios with each helper in a controlled setting. Google Assistant was better than Siri and Alexa at understanding indirect speech acts and keeping the context relevant. This was because it had a bigger training set and more powerful natural language processing algorithms.

Liao *et al.* (2020) [9] studied how well virtual workers can understand how to be polite when people ask for help. To do this, they paid attention to subtle speech acts. Their study showed that all three of the helpers had trouble understanding what was being said without directly saying it. They often needed clear instructions to give the right answer. People need to be able to use sensible thinking to communicate normally and in a way that seems human.

Porayska-Pomsta *et al.* (2013) [11] looked into how speech act theory could be used in tools for talk in the classroom. Their work, even though it wasn't about business virtual assistants, shows how important it is to be able to spot illocutionary acts so that systems and users are more interested. Their models influenced how AI interfaces were made in the future, such as how virtual assistants handle interactions.

The work by Su *et al.* (2019) [13] was also very helpful because it used reinforcement learning to improve the answers that virtual helpers gave by understanding how people talk. Their study showed that users were happy and conversations went more smoothly when aids could answer properly to speech

acts.

Finally, Wang *et al*. (2022) [15] talked about how hard it is for virtual helpers to understand funny and snarky speech. They discovered that current systems don't know enough about practice and context, so they often get these things wrong or don't notice them at all. They say that to make helpers more useful, language pragmatics and social computers should work together more.

All of these studies from the past show how important it is for virtual helpers to know how to understand speech acts and talk like a person. They also help shape the ways this compared research is done and its goals.

Methodology and Data Analysis

This study uses a mixed-methods technique that includes both qualitative discourse analysis and quantitative performance measures to look at how well Siri, Alexa, and Google Assistant can do everyday tasks. The main focus is on how well they can understand and react to different types of speech in spontaneous and partially controlled conversations.

Getting the Data

The study design used 60 speech act-based questions that were made to cover Searle's (1969) [12] five main types of speech: spokespeople, instructions, commissives, expressives, and statements. These prompts were made to match what real users would want to say, so they use both direct and indirect speech. Some examples are

- **Telling:** "Turn on the lights in the living room."
- **Indirect direction:** "It's getting dark in here."
- **Kind:** "Thank you for your help."
- **Giving in:** "I'll remember to remind you tomorrow." “The weather is cloudy today,” said the representative.
- **Statement:** "You are now linked to Wi-Fi."

Discussion

Each question was shown to Siri (the most recent version of iOS), Alexa (the most recent version of software for the Echo device), and Google Assistant (the most recent version of Android) in controlled settings. For speech helpers that need "wake words," the activating orders were always used. Responses were logged, analyzed, and recorded.

Criteria for Evaluation

A pragmatic-based assessment was used to grade the responses, with a focus on:

- **Accuracy:** The type of speech act was correctly identified.
- **Appropriateness:** How well the answer fits the question.
- **Contextual sensitivity:** being able to figure out what something means without knowing it directly.
- **User Satisfaction:** A made-up experience of the user based on how normal and helpful it is.

For each measure, three separate languages and HCI experts used a 5-point Likert scale to rate each answer. Cohen's kappa was used to figure out inter-rater dependability, and the results showed a high agreement (^{o} = 0.82).

Analysis

For quantitative study, mean scores were found for each group of helper and speech act. Overall, Google Assistant got the best score (mean = 4.1), followed by Alexa (mean = 3.6) and Siri (mean = 3.3). Google Assistant was very good at understanding complex expressions and vague instructions. For example, when someone said, "It's cold in here," the assistant could handle it in a variety of ways, suggesting that the heat be adjusted.

Alexa was great at giving orders, especially when it came to smart home settings, because she could do things like turn devices on or off quickly and correctly. Siri did a good job with expressives, usually replying politely with "thank you" or "I'm sorry." However, it had trouble with unclear or indirect speech acts, often giving generic answers or asking for explanations.

When looking at Siri's responses to complaints and statements, qualitative discourse analysis showed that it sometimes didn't understand what the user meant or react correctly. Alexa sometimes got indirect speech wrong and gave exact answers instead of assumed ones. Google Assistant's answers usually took into account the situation, but sometimes they were missing emotional depth, especially when making expressive sounds.

Some problems

The study does point out some problems, such as the fact that fast delivery is controlled, which means that it might not fully catch how users connect with each other on the spot. When devices and apps are updated after the experiment, they may change how well they work. Also, differences in culture and language were not taken into account because all tests were given in English with a normal US accent.

Conclusion

The authors of this study looked at how well Siri, Alexa, and Google Assistant could recognize and carry out different types of speech acts as described by Searle (1969) [12]. The results show big differences in how well different methods and types of speech acts work. This shows both the progress and the problems that still need to be solved in natural language interaction.

Overall, Google Assistant turned out to be the best method, especially when it came to handling complicated expressions and indirect commands. It has pretty advanced pragmatic awareness because it can figure out what people mean when they say things like "It's cold in here" and take them as a request to change the temperature. This fits with earlier study (Huang *et al*., 2021) [6] that said Google Assistant's great success is due in part to its large training dataset and environmental modeling. But sometimes, even Google Assistant had trouble fully catching emotional depth, which shows how hard it is for AI to do rational judgment.

Alexa was very good at giving orders, especially ones that could be carried out in a smart home environment. Its quick and accurate reactions to clear directions show that command execution was a big part of its design. Unfortunately, Alexa's flaws became clear when it tried to understand vague or indirect speech; it sometimes gave precise or unrelated answers. This shows that there is a space between receiving commands functionally and understanding them more

deeply.

When people thanked or apologized, Siri often responded nicely. This showed that she was good at emotional speech. However, it wasn't as good at spotting indirect, supportive, or declarative acts. People tend to ask for explanations or give general answers, which is a conservative way to talk to people but may get in the way of a smooth user experience. This trend matches what Liao *et al*. (2020) [9] found, showing that Siri can only make limited assumptions about indirectness. The comparison shows how important it is to include language pragmatics in the growth of AI. The idea of speech acts is a useful way to figure out where virtual helpers do well or poorly in conversation. There are differences seen, which means we need better models that include cognitive inference processes like human pragmatic thinking in addition to syntactic parsing and word matching.

These results have real-world implications for people who create and build AI. Increasing the practical knowledge of virtual helpers could mean:

- Adding contextual memory units that keep track of past conversations to help with inference.
- Training with datasets that are full of subtle speech acts, ways to be polite, and expressive language.
- Making models that are a mix of statistical learning and rule-based pragmatic systems.

To help with functional understanding, future study should look into multisensory engagement, such as using prosody and face signs. Cross-cultural studies could also show how understanding speech acts is different for different languages and social norms, which would help make AI interactions more universal.

In conclusion, virtual helpers have come a long way in their ability to connect with people using normal language, but they still have a long way to go in terms of being able to handle everyday situations. By based AI communication design on well-known theories of language like speech act theory, developers can get closer to making helpers that not only understand words but also the meanings and actions behind them. This will improve communication between humans and computers in the long run.

References

1. Allen JF. Natural language understanding. In: Encyclopedia of computer science. 4th ed. New York (NY): Wiley; 1995.
2. Austin JL. How to do things with words. Oxford (UK): Oxford University Press; 1962.
3. Bender EM, Koller A. Climbing towards NLU: on meaning, form, and understanding. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; 2020 Jul 5-10; Online. Stroudsburg (PA): Association for Computational Linguistics; 2020. p. 5185-98. doi: 10.18653/v1/2020.acl-main.463
4. Chung J, Park S, Lee S. Comparative analysis of virtual assistants: Siri, Alexa, and Google Assistant. J Hum Comput Interact. 2020;36(10):923-40. doi: 10.1080/07370024.2020.1754235
5. Clark HH. Using language. Cambridge (UK): Cambridge University Press; 1996.
6. Huang L, Wang Y, Li X. Context-aware natural language understanding in virtual assistants. IEEE Trans Neural Netw Learn Syst. 2021;32(7):2936-47. doi: 10.1109/TNNLS.2020.2994643
7. Hoy MB. Alexa, Siri, Cortana, and more: an introduction to voice assistants. Med Ref Serv Q. 2018;37(1):81-8. doi: 10.1080/02763869.2018.1404391
8. Kepuska V, Bohouta G. Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home). In: 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC); 2018 Jan 8-10; Las Vegas, NV. Piscataway (NJ): IEEE; 2018. p. 99-103. doi: 10.1109/CCWC.2018.8301638
9. Liao Q, Liu Y, Wei Z. Understanding politeness strategies in virtual assistant interactions. Int J Hum Comput Stud. 2020;135:102379. doi: 10.1016/j.ijhcs.2019.102379
10. Lopatovska I, Oropeza M. The voices behind the assistants: a comparison of Alexa, Siri, and Cortana. J Assoc Inf Sci Technol. 2018;69(10):1264-73. doi: 10.1002/asi.24059
11. Porayska-Pomsta K, Mavrikis M, Pain H. Using speech act theory to improve dialogue systems for educational technology. Int J Artif Intell Educ. 2013;23(4):383-412. doi: 10.1007/s40593-013-0012-4
12. Searle JR. Speech acts: an essay in the philosophy of language. Cambridge (UK): Cambridge University Press; 1969.
13. Su Y, Li H, Wang J. Reinforcement learning for dialogue management with speech act recognition. In: Proceedings of the 20th Annual SIGdial Meeting on Discourse and Dialogue; 2019 Sep 11-13; Stockholm, Sweden. Stroudsburg (PA): Association for Computational Linguistics; 2019. p. 135-45. doi: 10.18653/v1/W19-3614
14. Traum D, Larsson S. The information state approach to dialogue management. In: Dybkjær L, Minker W, editors. Recent trends in discourse and dialogue. Dordrecht (Netherlands): Springer; 2003. p. 325-53.
15. Wang F, Chen J, Zhang Y. Sarcasm and humor recognition in virtual assistants: a pragmatic perspective. Comput Hum Behav. 2022;130:107176. doi: 10.1016/j.chb.2021.107176
16. Winograd T. Understanding natural language. New York (NY): Academic Press; 1972.

How to Cite This Article

Alkhalidi SHK, Alkhalidi AH, Al-Asadi NSHT. Speech acts in virtual assistants: A comparative study of Siri, Alexa, and Google Assistant. Int J Multidiscip Compr Res. 2025;4(6):1-4. doi:10.54660/IJMCR.2025.4.6.01-04.

Creative Commons (CC) License