Article

Toward Smart Communication Components: Recent Advances in Human and AI Speaker Interaction

Hyejoo Kim [1], Sewoong Hwang [1], Jonghyuk Kim [2] and Zoonky Lee [1,*]

1   Graduate School of Information, Yonsei University, Seoul 03722, Korea; hyejoo.kim@yonsei.ac.kr (H.K.); indimoa@gmail.com (S.H.)
2   Division of Computer Science and Engineering, Sunmoon University, Asan-si 31460, Korea; jonghyuk@sunmoon.ac.kr
*   Correspondence: zlee@yonsei.ac.kr; Tel.: +82-02-2123-4528

Abstract: This study aims to investigate how humans and artificial intelligence (AI) speakers interact and to examine the interactions based on three types of communication failures: system, semantic, and effectiveness. We divided service failures using AI speaker user data provided by the top telecommunication service providers in South Korea and investigated the means to increase the continuity of product use for each type. We proved the occurrence of failure due to system error (H1) and negative results on sustainable use of the AI speaker due to not understanding the meaning (H2). It was observed that the number of users increases as the effectiveness failure rate increases. For single-person households constituted by persons in their 30s and 70s or older, the continued use of AI speakers was significant. We found that it alleviated loneliness and that human-machine interaction using AI speaker could reach a high level through a high degree of meaning transfer. We also expect AI speakers to play a positive role in single-person households, especially in cases of the elderly, which has become a tough challenge in the recent times.

Keywords: AI speaker; speech recognition; system and semantic failure; effectiveness communication; expectancy disconfirmation; single-person households; human-machine interaction

1. Introduction

According to the Market Insights Reports 2022, the global artificial intelligence (AI) speaker market has achieved continuous growth, reaching 5.08 billion dollars in 2021 and is expected to reach 8.71 billion dollars by 2022, as shown in Table 1. The report projects the market to reach 21.94 billion dollars by 2027, at a compound annual growth rate of 26.10% from 2022 to 2027. The supply of AI speakers surpassed 100 million in 2018, and is predicted to go beyond 200 million by 2022. This is a remarkable achievement only eight years after Amazon's Echo with Alexa was released in November 2014. The forecast for the future demand for AI speakers is reassuring—that they will become as important as smartphones.

Table 1. Global AI speaker market by vendor (shipments in millions of units).

Vendors	Market Share (2021)	Shipments (2020)	Shipments (2021)	Growth Rate
Amazon	26.6%	33.6	42.4	8.8%
Google	17.3%	23.8	27.6	3.8%
Baidu	15.6%	19.4	24.8	5.4%
Alibaba	12.6%	17.1	20.0	2.9%
Apple	9.6%	7.3	15.2	7.9%
Xiaomi	6.3%	10.6	10.0	−0.6%
Others	12.0%	18.9	19.1	0.2%
Total	100.0%	130.7	159.1	28.4%

AI speaker is a service platform designed to communicate and process user commands by combining speech recognition and text analysis technologies [1]. In 1954, when the speech recognition technology was first studied, it could not be commercialized owing to its poor recognition rate. AI speakers based on speech recognition technology began to be commercialized when Siri was installed in iPhone 4S in 2011. This has grown into a service provided in most IT products and mobile devices [2]. The mechanism of AI speaker is shown in Figure 1. The AI speaker transmits commands from the speech of the user and voice transmission through speech to text (STT) and conversation recognition stages [3]. Conversely, services are delivered to users through text to speech (TTS) and voice transmission [4].
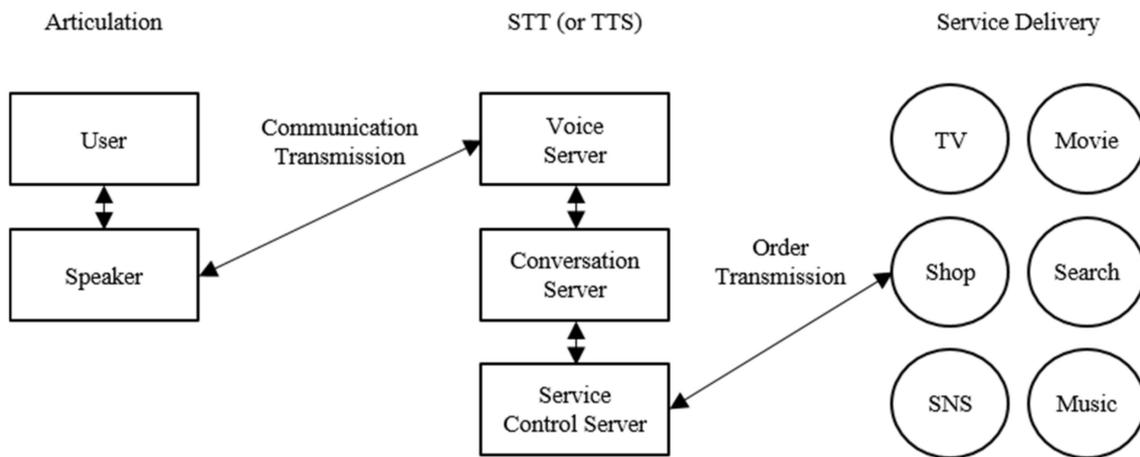

Figure 1. Operating process of AI speaker.

As extended technology and processing stages are required, the completion of the service is still lacking. Despite the lack of the technology readiness level, the services provided by AI speakers are expanding, and various companies are attempting to replace the existing services with new services [5]. As shown in Figure 2, the report published by voicebot.ai suggests that the most frequently used monthly services by the US consumers are music, searching the web, weather checking, and timer and alarm services [6].
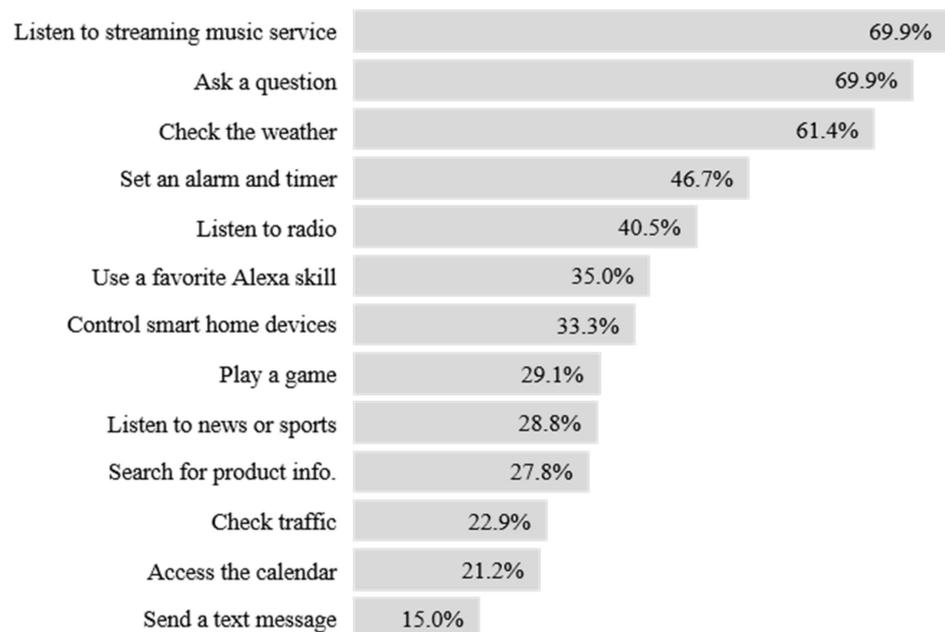

Figure 2. AI speaker use case frequency.

This paper comprises four chapters. Section 2 covers the mathematical theory of communication, which deals with the technical (or system) and semantic communication issues required for the interaction between humans and AI speakers and detects the issues of communication effectiveness. In addition, we examine the expectation of the customers and disconfirmation with the acceptance and non-use of specific technologies. Finally, we investigate the mechanism and current status of AI speaker development. In Section 3 we perform statistical analysis based on the user of AI speaker services. We conduct quantitative analysis using statistical techniques and descriptive statistics to establish and verify the hypotheses on which this study is based, according to prior studies. In Section 4, we present the expected effects of using the results of our study. Finally, in Section 5, we discuss the implications and limitations of the study and propose a future study that builds upon our findings.

2. Background

2.1. Shannon–Weaver Model of Communication: A Mathematical Theory of Communication

Shannon and Weaver provided rigorous and formal solutions to technical problems on which the information theory is based. Initially, Shannon and Weaver focused only on technical communication, intentionally excluding all accessible accidents related to semantic and effectiveness communication [7]. However, as communication develops into Internet of things (IoT) that connect humans and machines across various levels of intelligence and enables new services, semantic and effectiveness communication has become a core concept that can no longer be ignored [8]. The model proposed by Shannon and Weaver is an approach that views communication as a simple linear process, transmits as much information as possible in each path, and evaluates the information. Furthermore, this model quantifies the degree of complexity of information delivery by linking the information concept to the total amount of information and the amount of selectable information. That is, noise may occur in the communication process, apart from the issues of redundancy and optimality [9]. This may be engineering noise or semantic noise that may occur in the information interpretation process [10]. Shannon and Weaver divided the problems in communication into technical problems and semantic problems along with effectiveness problems according to the problem level [11]. The technical problem involves how accurately the signal is transmitted, the semantic problem involves whether the symbol conveys the meaning intended, and the effectiveness problem is whether the received meaning is performed as expected [12].

First, in terms of technical problems, noise must be eliminated or minimized because the noise generated in a limited channel reduces the efficiency of overall information transfer [13]. Furthermore, the noise in the information design of the print media includes decorations that make it difficult to read data, unnecessary visual devices, complex patterns that cause optical errors, vocabularies that make it difficult to convey meaning, and inappropriate images [14].

Second, according to the semantic problem, communication between humans involves the exchange of information, in which the word information is associated with meaning. The information conveyed when passing a concept from a source to a destination is a relevant aspect, not how the message is delivered to the destination [15]. An accurate semantic communication occurs when the concepts associated with the message sent by the source are correctly interpreted by the destination. This does not necessarily mean that the entire bit sequence used to transmit the message is decoded without errors. In other words, one of the main reasons that semantic levels offer significant performance gains relative to purely descriptive levels is that they leverage the sharing of prior knowledge between source and destination [16]. This knowledge can be human language or a formal language (at a more general level) consisting of a set of logical rules that enable entities and receivers to correct errors that occur at the symbol level. An interesting aspect of the semantic problem is that it arises from interactions between different languages. In the semantic problem, the noise serves as a clue to infer and predict the meaning of the message

delivered to the receiver. The number of selection conditions of the recipient is reduced when the amount of information is small in the information delivery process, such that the probability of the message being selected is high [17]. Conversely, if the amount of information is large, the number of message selection conditions increases. This makes it less likely that the recipient will select a particular message, and it is difficult to predict the meaning conveyed by the message.

Third, the effectiveness or goal-oriented problem is performed to achieve a common goal through communication between interacting entities. The basic system specification uses a number of resources (e.g., energy and computation) to precisely achieve the desired goal within a given time constraint. A communication system that enables interaction between goals and related entities should be defined to focus on goal-related specifications and constraints [18]. For example, any information that is not strictly related to attaining a goal can be ignored. The efficiency level is the level responsible for the efficient management of goal-oriented communication. The effectiveness problem is undoubtedly the most important virtue in any information design, regardless of the nature and purpose of information. However, if this is overemphasized, the role of information design may be limited only to solving technical problems of communication. Early research on the effectiveness problem mainly focused on how to visualize data, and qualitative aspects of information such as images and narratives were mainly studied [19]. Appropriate expression of entropy and redundancy in information design can achieve information delivery by arousing interest and increasing the level of involvement to actively interpret information [20]. Various studies have suggested the possibility of actively utilizing noise in information design by interpreting the noise in the communication process as "noise as an interest factor" and "surplus as a persuasion factor" from the audience perspective. Particularly, play, storytelling, and interaction with information surplus are more effective because information design requires user participation and interrelationship in a multimedia environment.

Shannon and Weaver proposed that communication is composed of three levels as follows [21]:

Low level of the stack (The technical problem): How accurately can the communication symbols be transmitted.

Middle level of the stack (The semantic problem): How precise the transmitted symbols convey the desired meaning.

High level of the stack (The effectiveness problem): How effective the received meaning affects conduct in the desired way.

2.2. Expectation Disconfirmation Theory

Generally, people react differently in terms of satisfaction even when they use the same product or service, implying that product performance (i.e., quality perceived by consumers) is determined by the expectation of the consumers in addition to the objective function of the product. In other words, consumers do not determine their level of satisfaction with a product based only on the performance level of the product but compare the product performance with their initial expectations to determine their satisfaction [22]. In the expectancy disconfirmation paradigm, the most studied concept as a comparative criterion is expectation. However, there is no clear conceptual consensus on what expectations mean. Several researchers have conceptualized expectations as "perceptions of the likelihood of some event" or "perceptions of the probability of occurrence of some event". By contrast, other researchers view expectations as a concept that includes the "estimation of the likelihood of a specific event" and "evaluation of the good or bad of that event" [23].

Expectations include predictive expectation, desired expectation, and normative expectation. Predictive expectation is wherein the performance is up to a certain extent whereas desires expectation is that it is desirable to have a certain level of performance [24]. Normative expectations imply that performance should be up to some extent. There are four types of expectations: ideal, predictive, natural, and minimum acceptable expectations. There are studies that divide consumer expectations into predictive and normative expec-

tations. Predictive expectation is the prediction of the consumer regarding the expected frequency of problem occurrence, and normative expectation is defined as the normative evaluation of how often problems occur [25,26]. As shown in Figure 3, the initial interest in expectancy disconfirmation research in marketing investigated how expectations, rather than expectations and satisfaction, affect perceived consumer performance. These studies were primarily concerned with whether the effect of expectations on the perceived performance is based on the assimilation theory or contrast theory. At the beginning of the studies, contradictory research results were obtained. However, in later studies, it was reported that consumer expectations generally have a positive effect on product perception [27].
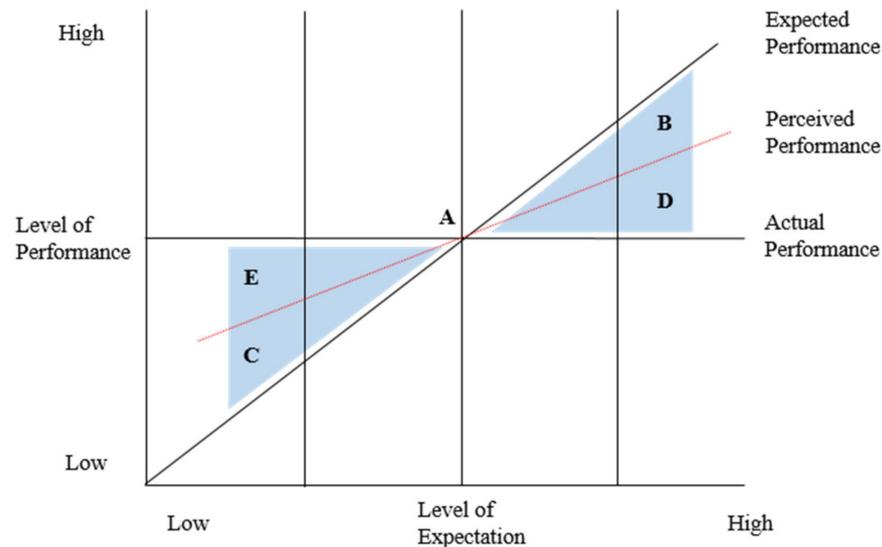

Figure 3. Schematic of the expectancy disconfirmation theory.

Early research on consumer satisfaction mainly focused on analyzing how expectations affect user satisfaction rather than the effect of expectancy disconfirmation on the performance. Most of these studies have attempted to explain the effect of expectation as an assimilation effect based on the cognitive dissonance theory of Festinger [28]. For example, a consumer who purchases a product with high expectations will feel psychologically uncomfortable when the performance of the product fails to meet the expectations. As it is impossible to increase the performance of the product, to solve the psychological discomfort (cognitive dissonance), they try to satisfy themselves based on their high expectations, thereby increasing the level of satisfaction [29].

Expectancy disconfirmation is based on product expectations and perceived performance. Therefore, this theory can accurately explain how expectations and perceived performance affect consumer satisfaction. This theory is generally known as the main theory that explains consumer satisfaction based on expectations [30]. If the product performance is higher than the expectations of the consumer, the level of satisfaction rises, and the consumer becomes dissatisfied if the product performance is lower than the expectation owing to the disappointment effect. When product performance is judged to be lower than expected, better than expected, or equal to expectation, it is called negative disconfirmation, positive disconfirmation, or simple confirmation, respectively. Therefore, in the case of simple confirmation and positive confirmation, the consumer is satisfied whereas the consumer is dissatisfied in the case of negative confirmation [31].

The expectancy disconfirmation theory hypothesizes expectations tied to perceived performance, leading to post-purchase satisfaction. This effect is mediated through positive or negative disconfirmation between expectation and satisfaction; satisfaction occurs when the service expectation is exceeded whereas the customer is dissatisfied if the expectation is not met. As shown in Figure 4, the expectation of the present time point (t) is connected

with the disconfirmation of the future time point (t + 1) and affects the choice of the consumer [32].
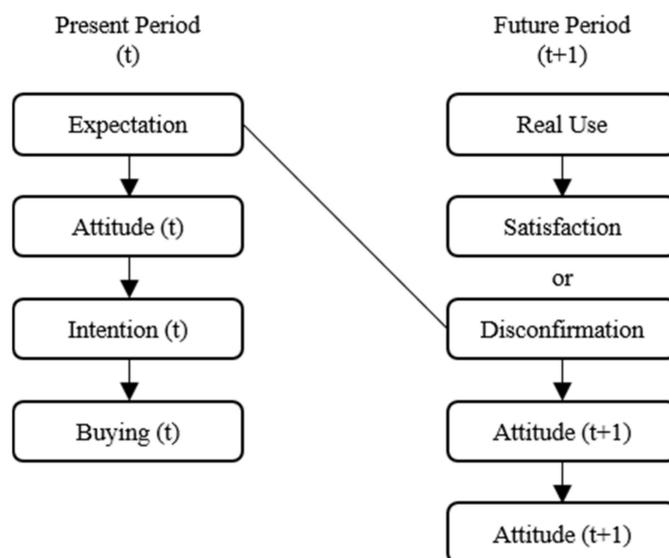
Figure 4. Consumer satisfaction process by time series.

If a good reputation is created by appropriately responding to changes in the demands and expectations of major stakeholder groups and maintaining harmony, then building consumer trust and confidence will naturally follow. Therefore, it is necessary for any company to pay sufficient attention to its key stakeholders and strive to meet their expectations [33]. By meeting stakeholder expectations, a company can anticipate and respond to potential crises in advance, build trust in the organization, and have the results of the company confirmed by stakeholders. The expectations and interests of stakeholders are constantly evolving; therefore, companies should conduct regular stakeholder monitoring and dialogue to keep pace with these developments, and thus, the stakeholder feedback can be obtained [34,35].

2.3. Technical Background of AI Speaker

IoT is one of the representative information technologies leading the fourth industrial revolution. The core technology of IoT consists of the collection, processing, and management of data coming through sensors, wired or wireless communication and network infrastructure, security technology to prevent information leakage, and software that can connect various technologies [36]. Therefore, although IoT is a service industry related to individual consumption, it has a complex structure in which large industries in various fields are complementary to each other. AI speakers focus on convenience functions in general life based on IoT [37,38]. Companies that design AI speakers focus on content businesses that consider human accessibility and intimacy as well as the functional aspects of devices. They have even started to formulate the name of AI assistant service. The AI speaker is a voice command device with a built-in virtual assistant, which provides interactive work and hands-free activation [39]. The core of the virtual assistant service is human–machine interaction through question answering (QA). All AI speakers are equipped with voice recognition technology by default, but a special wake word is required to link the device so that it can communicate with the server [40]. AI collects and analyzes user commands to provide information and services tailored to the situation. In other words, the AI speaker collects the command, compares it with other commands stored in the cloud, and recognizes it when a user inputs a command to the AI speaker. The service is operated in such a way that the suitable data are retrieved from the big data and appropriate information is provided to the user through the recognized command [41,42].

This enables the industry to launch AI services that can satisfy accessibility and interactivity. Furthermore, an AI speaker is composed of a speech recognition technology that converts human voice into text data that can be recognized by a computer and a natural language processing technology that can process the converted data [43,44]. The convergence of these two speech-language processing technologies is not only used in the AI speaker industry, but also in the wearable device and artificial intelligence industry (which is based on human intimacy). AI speakers are based on voice assistants or voice user interfaces (VUIs), such as Google Assistant and Amazon Alexa. VUIs receive language as input information through a speech recognizer and output voice information through speech synthesis or pre-recorded audio. There is a growing interest in integrating the voice assistant with various devices, such as AI speakers, smartphones, and smart TVs, to improve the voice recognition rate. The error rate is 5% when a person hears and transcribes the conversation over the phone. In 2017, the interactive agent recognition error rate reached 5%, a level similar to that of humans [45]. Figure 5 shows the basic model of acoustic echo cancellation. When a user tries a voice command while the music is playing on the speaker, the AI speaker removes the music signal by applying acoustic echo cancellation technology and only accepts the voice command signal without distortion. For the acoustic echo cancellation technology to work ideally, the reference signal (R) for signal processing and the signal (Rˆ) reproduced through the AI speaker and input back to the microphone must exactly match. In this case, even if the reproduced music signal is much larger than the voice command, the AI speaker accurately recognizes only the voice command signal [46].
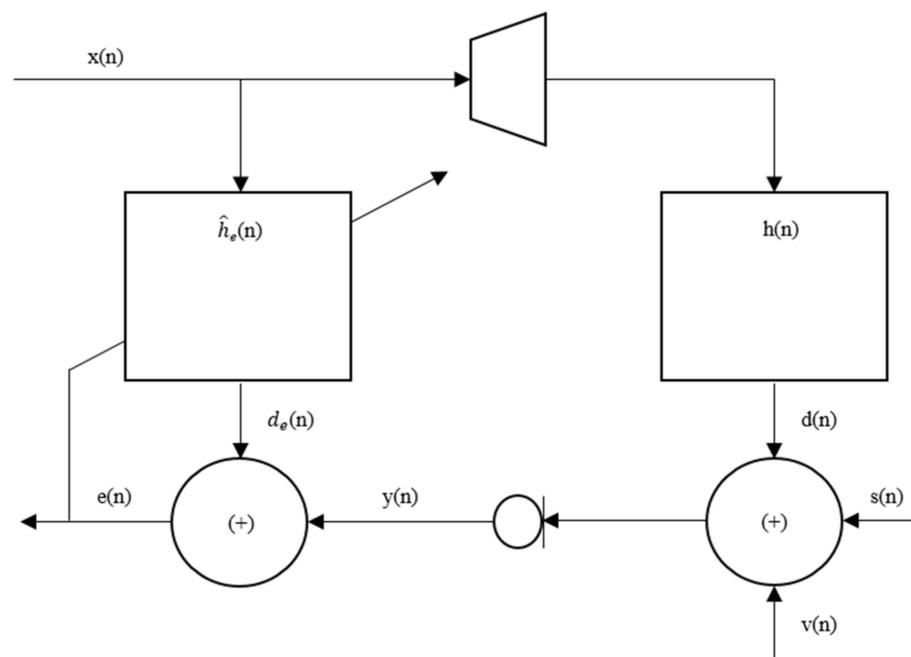
Figure 5. Block diagram of the acoustic echo cancellation system.

The most basic principle for acoustic echo cancellation is to estimate the characteristics of the acoustic path, which should be modeled as accurately as possible to implement a pattern similar to the actual echo. An adaptive filter should be used as the characteristics of the acoustic path generally vary according to time and surrounding conditions [47,48]. The situation of the music interference seriously affects the voice recognition performance. The quality of acoustic echo cancellation depends on the speed of convergence and the accuracy of the adaptive filter. The echo signal is typically modeled as a convolution of the musical signal and the impulse response of the audio path [49].

3. Methodology

3.1. Research Methods

IT corporations, which develop and sell AI speakers, have been persistently making colossal investments in marketing to survive in the competitive market. This emphasizes the importance of understanding consumer behavior in developing and expanding the services that enable the consumers to lock-in the company product. Therefore, we focus on discussing and analyzing the following two issues: (1) How the initial experience of the user affects the continued use of the speaker and (2) how the individual characteristics of the user affect consumption. We seek to provide insight into what customer behavior characteristics companies should consider to increase sustainability in the future.

In this study, we used real consumer data on AI speaker obtained from Korea Telecom (KT), which is one of the most representative broadband companies in South Korea. The company launched an AI speaker service, Giga-Genie, which is installed in all-in-one set-up box to control Internet Protocol Television (IPTV). The content of the communication appears on the TV screen which improves the function by helping consumers know how their words are recognized. As the service comes along with the IPTV registration, there are consumers who unwillingly registered with the AI speaker, and this could provide a meaningful consequence because we specifically used three months' data of the customers registered in December 2017.

To analyze the initial user experience, we started with the section that is highlighted by the users. Tracing back to the resource above, voicebot.ai, the most important function that consumers focus on is the understandability of the communication between the AI speaker and users. To improve this, failure to maintain communication should be reviewed. The three aspects of the success in communication consist of technical (or system), semantic, and effectiveness aspects as defined by the mathematical theory of communication in Section 2. As shown in Table 2, we analyzed communication failure between AI speakers and users in these three aspects.

Table 2. Classification of the typical failure experiences of AI speaker users.

Type	Request Message (By Customer)	System Message	Results
System Comm. Failure	Do "something."	The network connection failed to cancel the service. Please try again later.	Cancellation of service
	How is the weather in Kangnam, Seoul?	I cannot access the weather information. Please try again later.	A sudden service failure
	Turn on the Netflix service.	A subscription to the service is required.	Requirement to sign-in to additional service
Semantic Comm. Failure	How is the weather in Kangnam, Seoul?	I cannot find the name of the location. Please ask another location.	Rejection to unavailable service request or incomprehensible request
	Play the music of "ABBA".	I will turn on the music of "Bach".	Searching for other results with similar pronunciation.
	Search for the movie 'Smile'.	I'll smile with you too, ha-ha-ha.	error in recognizing the meaning of words
Effectiveness Comm. Failure	Search for the movie "Squid Game".	Should I search contents related to "Squid Game" in TV or in web browser?	Inducement of the user's choice amongst possible actions
	Call Mr. Smith.	Do you want to call Mr. Smith?	Confirmation of the user's will
	I want to make a call.	Let me know the contact number you would like to make a call.	Request for the additional information

We observed and analyzed the log data of 27,308 AI speaker users for three months from December 2017 to February 2018. We used variables that consider individual attributes to quantify the exposure level of the user environment along with basic demographic information. We defined exposure level as the TV viewing time as the IPTV and set-up box are all-in-one service. The gender and age of the subscribers were used for personal profile information. The failure in each field was measured by the failed percentage of the

request attempt. We classified independent variables (Xs) as the percentage of three types of failures with the moderating variables considering the first day, accumulated three days, one week, and two and more weeks. We used two more types of independent variables: exposure level of the user environment and demographic information such as age and gender. Finally, we defined two types of dependent variables (Ys) to display continuous usage of the speaker as shown in Table 3. We considered a discrete dependent variable (Y1) and a continuous dependent variable (Y2), both of which indicate continuous use as a dependent variable, simultaneously. We defined the continuous use of discrete Y1 as 1 for more than five days and 0 for less than five days. In the case of continuous Y2, the date and number of use cases were quantitatively constructed.

Table 3. Operational definition of variables.

Variables		Definition
Independent Vars. (Xs)		Percentage of three types of failures with system, semantic, and effectiveness. the exposure level of the user environment, and demographic information (ex. age and gender)
Moderating Vars.		Period of customer failure with first-day, accumulated three-days, one-week, and two and more weeks.
Dependent Var.	Discrete Var. (Y1)	Sustainable use of AI speaker, with one for more than five days, and 0 for less than five days
	Continuous Var. (Y2)	Sustainable use of AI speaker, with the date of use, and the number of use cases

3.2. Hypothesis

We discovered that systematic (or technical) service error that occurs in the verbal request owing to the circumstances where the service is not ready, the user should sign-in to the additional service, or voice recognition technology is not working, leads to a decline in the customers' trust in the voice service. This eventually hinders customers from purchasing further subscription to the AI speaker. Thus, we assumed hypothesis 1 as a system failure.

Hypothesis 1 (H1). *If no service is technically provided in response to consumer request, the continued use of AI speakers will be negatively affected.*

The provision of the wrong service refers to the incident where the result of the verbal request of the user is dissatisfying, leading to another request within 5 s. This refers to unrelated emotional chat or some other reason due to the absurdness of the situation. The circumstance where a different type of service from what the consumer requested is provided can be explained, for example, as in not providing music service when the consumer has requested for it. We argued that the disharmony in the service during the initial experience within the first day, first three days, first week, and first two weeks or more, leads to the absence of consumer needs. Therefore, we assumed the second hypothesis as a semantic failure.

Hypothesis 2 (H2). *Providing a kind of service that is different from the one ordered by the consumer will have a negative impact on the continued use of the AI speaker.*

If the AI speaker cannot recognize the verbal request at once, or requires additional information on the verbal request, the impression on its continued use is negative. For example, when the user commands, "turn on Beatles", and the speaker replies, "If you want to listen to the music of Beatles, say 'Turn on the music of Beatles'". This is a typical example of the circumstance requesting a change in the method of verbal request or additional information. Replying with sentences such as "I couldn't understand" or

"Repeat your request" lowers the likelihood of continuing the subscription. Therefore, we assumed a third hypothesis as an effectiveness failure.

Hypothesis 3 (H3). *If the consumer does not recognize the commands at once or has requested for additional information, the continued use of the AI speaker is negatively affected.*

In addition to the three hypotheses, we hypothesized that the degree of continued use of AI speakers varies according to the degree of exposure to the usage environment and demographic information. Accordingly, we established two more hypotheses.

Hypothesis 4 (H4). *The continuity of usage depends on the mechanical time the users stay at home.*

Hypothesis 5 (H5). *The continuity of the usage differs based on the age, range, and gender of the users.*

Figure 6 schematically illustrates the causal relationships for the above five hypotheses, and each independent variable is classified according to the customer failure period, which is a moderator variable.
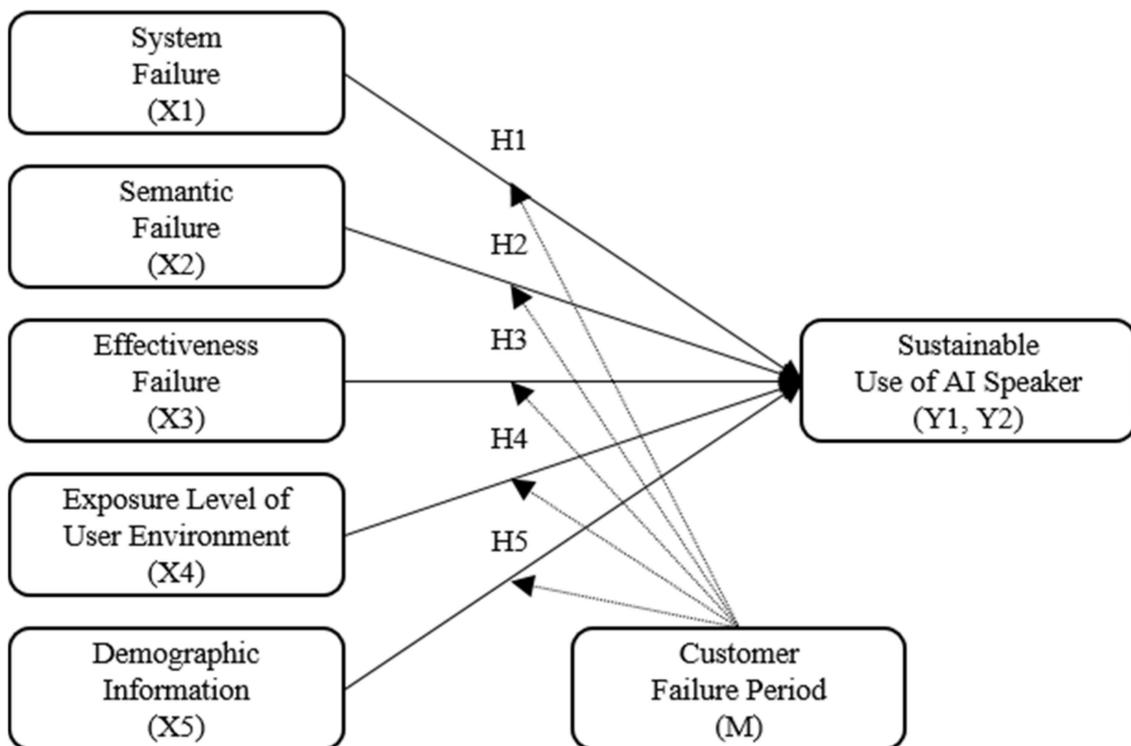

Figure 6. Causality of the hypotheses.

4. Results

To test the first hypothesis, as shown in Figure 7, we explored the continuous use of the AI speaker for each customer failure period related to system failure. We found that the sustained use was negatively affected if a system failure occurred in any period and that the most significant difference in sustained use was for users with a period of two weeks or more.
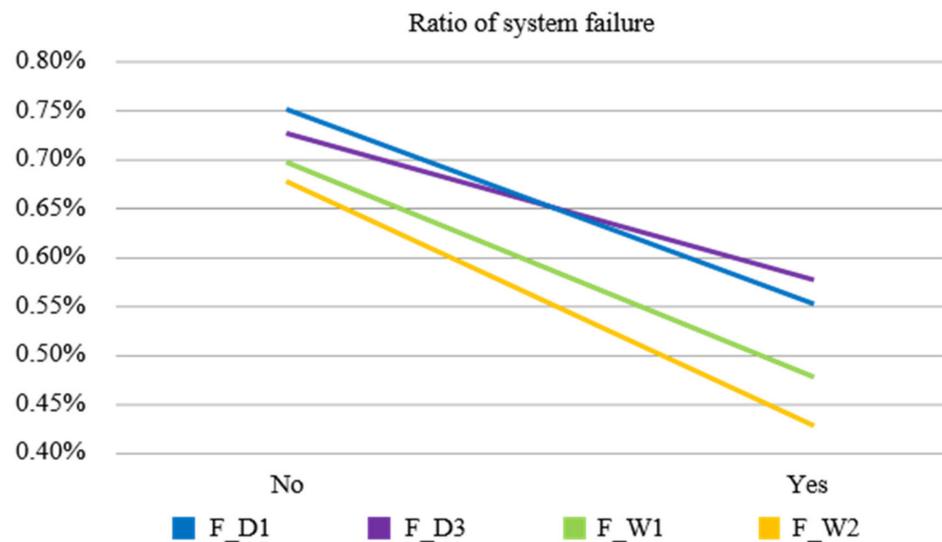
Figure 7. Ratio of system failure.

We explored the continuous use of AI speakers for each failure period of semantic failure-related customers to test the second hypothesis. As shown in Figure 8, we found that the occurrence of semantic failure across all periods had a negative effect on sustained use. The largest difference in continued use was the users over a period of two weeks.
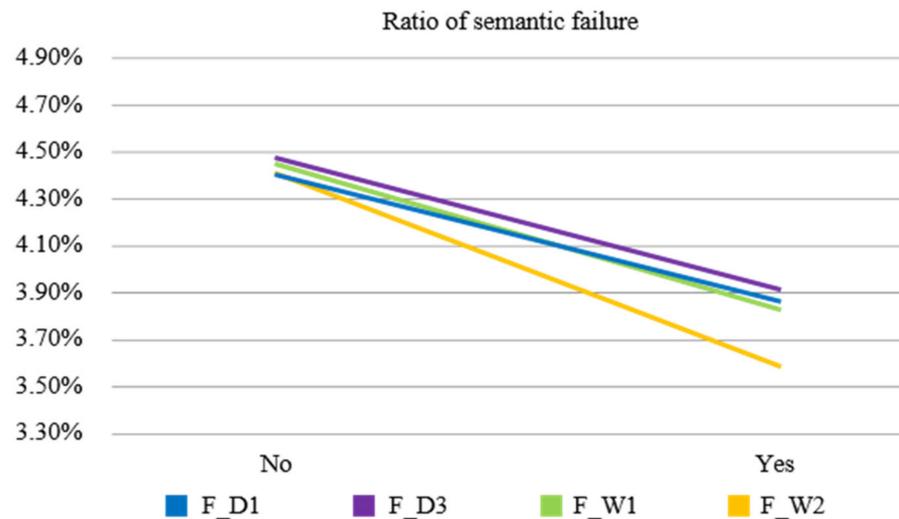

Figure 8. Ratio of semantic failure.

Finally, to test of the third hypothesis, we explored the continuous use of the AI speaker for each failure period related to the effectiveness failure. We reached conclusions that were contrary to the results of the previous hypotheses. That is, the negative effect on continuous use did not increase even if the effectiveness failure occurred in almost all periods, as shown in Figure 9. Rather, most of the cumulative users, except for 1-day users, showed a positive will to use despite the failure. Moreover, we discovered that, as with the previous hypotheses, the biggest difference in sustained use was for users over a period of two weeks.
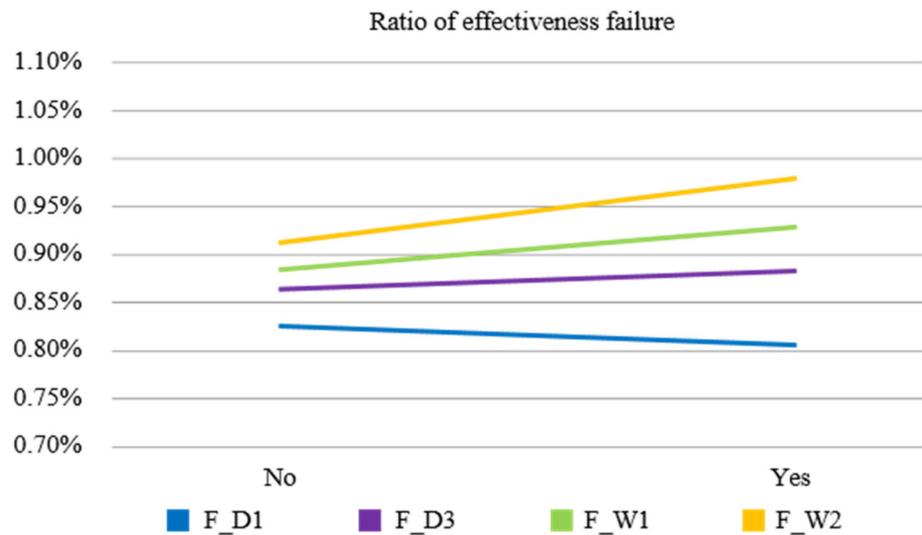
Figure 9. Ratio of effectiveness failure.

We searched for each failure variable by date for customers using the AI speaker for more than five days a week for three consecutive months. Based on all graphs, we found that the period that showed the greatest difference was the failure rate over two weeks. Therefore, we performed a logistic regression analysis for Y1 and regression analysis for Y2 with reference to the failure rate of two weeks to test each hypothesis. Tables 4 and 5 lists the results.

Table 4. Result of analysis 1.

Hypothesis	Y1	Y2
H1	0.0007 **	0.0096 **
H2	<0.0001 **	<0.0001 **
H3	0.721 (N.S.)	0.0575
H4	<0.0001 **	0.0042 **

(Significant Level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 5. Result of analysis 2.

Components of H6		Case of Y1 = 1	Y2 (Days)
	F	9.23%	10.2
Gender	M	9.37%	10.0
	Diff.	N.S.	N.S.
	Under 20	5.05%	7.6
	20s	8.07%	9.9
	30s	12.31%	11.2
	40s	5.79%	8.7
Age	50s	4.21%	7.5
	60s	9.45%	9.1
	>70s	12.31%	10.8
	Diff.	<0.0001 **	<0.0001 **

(Significant Level: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

We statistically proved that the failure due to system error (H1) and not understanding the meaning (H2) resulted in negative results on continued use of the AI speaker when two dependent variables were applied. However, in the case of H3, contrary to H1 and H2, the higher the failure rate, the greater the number of cases used. We proved that the additional queries and re-words of H3 provide a positive experience for people to continue the conversations.

We found no statistically significant differences between sexes. However, found a difference in usage according to age. We found that sustained use was significantly higher for those in their 30s and 70s or older. The common tendency between the two age groups is that there are mainly single-person household; therefore, the lack of a conversation partner could increase conversation with the AI speaker, which explains the positive result of continued use.

5. Conclusions

From the results of this study, unlike the system and semantic communication failure, effectiveness failure has a positive effect on the continuous use of AI speakers. This proves that human–machine interaction can reach a high level through a high degree of meaning transfer. AI speakers are expected to play a positive role in the case of single-person households, especially the elderly; such cases are ubiquitous. This study improves the continued customer use of products based on data from the top telecommunication service companies in South Korea and proves the three types of communication functions academically. However, short-term data may cause a problem in that the failure type is simple. To address these problems, we plan to include type diversification along with customer groups in future research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shalini, A.; Jayasuruthi, L.; VinothKumar, V. Voice recognition robot control using android device. *J. Comput. Nanosci.* **2018**, *15*, 2197–2201. [CrossRef]
2. Lee, W.; Seong, J.J.; Ozlu, B.; Shim, B.S.; Marakhimov, A.; Lee, S. Biosignal sensors and deep learning-based speech recognition: A review. *Sensors* **2021**, *21*, 1399. [CrossRef] [PubMed]
3. Yadav, S.; Kaushik, A. Do You Ever Get Off Track in a Conversation? The Conversational System's Anatomy and Evaluation Metrics. *Knowledge* **2022**, *2*, 55–87. [CrossRef]
4. Jeong, H.D.J.; Ye, S.K.; Lim, J.; You, I.; Hyun, W. A computer remote control system based on speech recognition technologies of mobile devices and wireless communication technologies. *Comput. Sci. Inf. Syst.* **2014**, *11*, 1001–1016. [CrossRef]
5. Xie, Y.; Li, F.; Wu, Y.; Wang, Y. HearFit: Fitness monitoring on smart speakers via active acoustic sensing. In Proceedings of the IEEE INFOCOM 2021-IEEE Conference on Computer Communications, Vancouver, BC, Canada, 10–13 May 2021; pp. 1–10. [CrossRef]
6. Youn, M.A.; Lim, Y.; Seo, K.; Chung, H.; Lee, S. Forensic analysis for AI speaker with display Echo Show 2nd generation as a case study. *Forensic Sci. Int. Digit. Investig.* **2021**, *38*, 301130. [CrossRef]
7. Strandberg, P.E. Automated system-level software testing of industrial networked embedded systems. *arXiv* **2021**, arXiv:2111.08312.
8. Yang, W.; Liew, Z.Q.; Lim, W.Y.B.; Xiong, Z.; Niyato, D.; Chi, X.; Cao, X.; Letaief, K.B. Semantic communication meets edge intelligence. *arXiv* **2022**, arXiv:2202.06471.
9. Semrau, D.; Lavery, D.; Galdino, L.; Killey, R.I.; Bayvel, P. The impact of transceiver noise on digital nonlinearity compensation. *J. Light Technol.* **2018**, *36*, 695–702. [CrossRef]
10. Dušek, O.; Howcroft, D.M.; Rieser, V. Semantic noise matters for neural natural language generation. *arXiv* **2019**, arXiv:1911.03905. [CrossRef]
11. Gillespie, D.J.; Schiffman, R. A critique of the Shannon-Weaver theory of communication and its implications for nursing. *Res. Theory Nurs. Pract.* **2018**, *32*, 216–225. [CrossRef]
12. Weng, Z.; Qin, Z. Semantic communication systems for speech transmission. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 2434–2444. [CrossRef]
13. Lippi, G.L.; Mørk, J.; Puccioni, G.P. Numerical solutions to the laser rate equations with noise: Technical issues, implementation and pitfalls. In *Nanophotonics*; SPIE: Bellingham, WA, USA, 2018; Volume 10672, pp. 82–95. [CrossRef]
14. Bergemann, D.; Morris, S. Information design: A unified perspective. *J. Econ. Lit.* **2019**, *57*, 44–95. [CrossRef]

15. Fedushko, S.; Benova, E. Semantic analysis for information and communication threats detection of online service users. *Procedia Comput. Sci.* **2019**, *160*, 254–259. [CrossRef]
16. Strinati, E.C.; Barbarossa, S. 6G networks: Beyond Shannon towards semantic and goal-oriented communications. *Comput. Netw.* **2021**, *190*, 107930. [CrossRef]
17. Maulud, D.H.; Zeebaree, S.R.; Jacksi, K.; Sadeeq, M.A.M.; Sharif, K.H. State of art for semantic analysis of natural language processing. *Qubahan Acad. J.* **2021**, *1*, 21–28. [CrossRef]
18. Michael, J.; Rumpe, B.; Zimmermann, L.T. Goal modeling and mdse for behavior assistance. In Proceedings of the 2021 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C), Fukuoka, Japan, 10–15 October 2021; pp. 370–379. [CrossRef]
19. Li, G.; Kou, G.; Peng, Y. A group decision making model for integrating heterogeneous information. *IEEE Trans. Syst. Man Cybern. Syst.* **2016**, *48*, 982–992. [CrossRef]
20. Lopez-Caudana, E.; Ramirez-Montoya, M.S.; Martínez-Pérez, S.; Rodríguez-Abitia, G. Using robotics to enhance active learning in mathematics: A multi-scenario study. *Mathematics* **2020**, *8*, 2163. [CrossRef]
21. Grech, N.; Brent, L.; Scholz, B.; Smaragdakis, Y. Gigahorse: Thorough, declarative decompilation of smart contracts. In Proceedings of the 2019 IEEE/ACM 41st International Conference on Software Engineering, Montreal, QC, Canada, 25–31 May 2019; pp. 1176–1186. [CrossRef]
22. Lee, C.P.; Hung, M.J.; Chen, D.Y. Factors affecting citizen satisfaction: Examining from the perspective of the expectancy disconfirmation theory and individual differences. *Asian J. Political Sci.* **2022**, 1–26. [CrossRef]
23. Nuradiana, S.; Sobari, N. Expectancy disconfirmation theory on millenials consumer behaviour in retail store. *ICORE* **2021**, *5*, 116–125.
24. Liu, F.; Lim, E.T.; Li, H.; Tan, C.W.; Cyr, D. Disentangling utilitarian and hedonic consumption behavior in online shopping: An expectation disconfirmation perspective. *Inf. Manag.* **2020**, *57*, 103199. [CrossRef]
25. Wang, X.; Zhou, R.; Zhang, R. The impact of expectation and disconfirmation on user experience and behavior intention. In *International Conference on Human-Computer Interaction*; Springer: Cham, Switzerland, 2020; pp. 464–475. [CrossRef]
26. Zhang, X.; Chen, Y. Admissibility and robust stabilization of continuous linear singular fractional order systems with the fractional order α: The $0 < \alpha < 1$ case. *ISA Trans.* **2018**, *82*, 42–50. [PubMed]
27. Ugaddan, R.G. Does performance management effectiveness matter? Testing the expanded expectations disconfirmation model of local disaster risk reduction. *Asia-Pac. Soc. Sci. Rev.* **2021**, *21*, 220–235.
28. Delly, M.C.; Kealesitse, B.; Moeti-Lysson, J.; Nametsegang, A. An Expectation Disconfirmation Analysis of Undergraduate Research Supervision: Opinions of Business Students at the University of Botswana. *Botsw. J. Bus.* **2021**, *13*. Available online: https://journals.ub.bw/index.php/bjb/article/view/1964 (accessed on 6 April 2022).
29. Fadel, K.J.; Meservy, T.O.; Kirwan, C.B. Information filtering in electronic networks of practice: An fMRI investigation of expectation (dis) confirmation. *J. Assoc. Inf. Syst.* **2022**, *23*, 491–520. [CrossRef]
30. Wang, C.; Liu, J.; Zhang, T. 'What if my experience was not what I expected?': Examining expectation-experience (dis) confirmation effects in China's rural destinations. *J. Vacat. Mark.* **2021**, *27*, 365–384. [CrossRef]
31. Dos Santos, M.A.; Baeza, S.; Lizama, J.C. The intention of attending a sporting event through expectation disconfirmation and the effect of emotions. In *Integrated Marketing Communications, Strategies, and Tactical Operations in Sports Organizations*; IGI Global: Hershey, PA, USA, 2019; pp. 223–240. [CrossRef]
32. Liu, J.; Shah, C. Investigating the impacts of expectation disconfirmation on web search. In Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, Glasgow, UK, 10–14 March 2019; pp. 319–323. [CrossRef]
33. Evangelidis, I.; Van Osselaer, S.M. Points of (dis) parity: Expectation disconfirmation from common attributes in consumer choice. *J. Mark. Res.* **2018**, *55*, 1–13. [CrossRef]
34. Matikiti, R.; Mpinganjira, M.; Roberts-Lombard, M. Antecedents and outcomes of positive disconfirmation after service failure and recovery. *J. Glob. Bus. Technol.* **2018**, *14*, 43–57.
35. Kaushik, A.; Loir, N.; Jones, G.J. Multi-view conversational search interface using a dialogue-based agent. In *European Conference on Information Retrieval*; Springer: Cham, Switzerland, 2021; pp. 520–524.
36. Bentley, F.; Luvogt, C.; Silverman, M.; Wirasinghe, R.; White, B.; Lottridge, D. Understanding the long-term use of smart speaker assistants. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–24. [CrossRef]
37. Ling, H.C.; Chen, H.R.; Ho, K.K.; Hsiao, K.L. Exploring the factors affecting customers' intention to purchase a smart speaker. *J. Retail. Consum. Serv.* **2021**, *59*, 102331. [CrossRef]
38. Wang, J.S. Exploring biometric identification in FinTech applications based on the modified TAM. *Financ. Innov.* **2021**, *7*, 1–24. [CrossRef]
39. Ashfaq, M.; Yun, J.; Yu, S. My smart speaker is cool! perceived coolness, perceived values, and users' attitude toward smart speakers. *Int. J. Hum.–Comput. Interact.* **2021**, *37*, 560–573. [CrossRef]
40. Kim, S.; Choudhury, A. Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study. *Comput. Hum. Behav.* **2021**, *124*, 106914. [CrossRef]
41. Smith, E.; Sumner, P.; Hedge, C.; Powell, G. Smart-speaker technology and intellectual disabilities: Agency and wellbeing. *Disabil. Rehabil. Assist. Technol.* **2020**, 1–11. [CrossRef] [PubMed]
42. Hu, K.H.; Hsu, M.F.; Chen, F.H.; Liu, M.Z. Identifying the key factors of subsidiary supervision and management using an innovative hybrid architecture in a big data environment. *Financ. Innov.* **2021**, *7*, 10. [CrossRef] [PubMed]

43. Hashemi, S.H.; Williams, K.; El Kholy, A.; Zitouni, I.; Crook, P.A. Measuring user satisfaction on smart speaker intelligent assistants using intent sensitive query embeddings. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Turin, Italy, 22–26 October 2018; pp. 1183–1192. [CrossRef]

44. Zhang, J.X.; Yang, G.H. Low-complexity tracking control of strict-feedback systems with unknown control directions. *IEEE Trans. Autom. Control.* **2019**, *64*, 5175–5182. [CrossRef]

45. Brause, S.R.; Blank, G. Externalized domestication: Smart speaker assistants, networks, and domestication theory. *Inf. Commun. Soc.* **2020**, *23*, 751–763. [CrossRef]

46. Choi, Y.; Demiris, G.; Thompson, H. Feasibility of smart speaker use to support aging in place. *Innov. Aging* **2018**, *2* (Suppl. S1), 560. [CrossRef]

47. Jung, H.; Oh, C.; Hwang, G.; Oh, C.Y.; Lee, J.; Suh, B. Tell me more: Understanding user interaction of smart speaker news powered by conversational search. In Proceedings of the Extended Abstracts of the 2019 chi Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; pp. 1–6. [CrossRef]

48. Kaushik, A.; Jones, G.J. A Conceptual Framework for Implicit Evaluation of Conversational Search Interfaces. *arXiv* **2021**, arXiv:2104.03940.

49. Ito, T.; Oyama, T.; Watanabe, T. Smart speaker interaction through ARM-COMS for health monitoring platform. In *International Conference on Human-Computer Interaction*; Springer: Cham, Switzerland, 2021; pp. 396–405. [CrossRef]