Marketing Letters (2026) 37:9
https://doi.org/10.1007/s11002-025-09806-1

ORIGINAL RESEARCH

Splurging with Alexa: How voicebots increase product upgrades

Sang Kyu Park[1] · **Yegyu Han**[2] · **Aner Sela**[3]

Received: 22 April 2025 / Accepted: 22 September 2025

Abstract
Do AI-enabled voicebots, such as Amazon Alexa and Google Assistant, influence consumer choice, and if so, how and why? We demonstrate that consumers are more likely to choose expensive upgrades over more basic options when shopping using voicebots, compared to screen-based or text-based online shopping interfaces. Eleven studies using real voicebots, including secondary data from an online vendor, suggest that this tendency to upgrade arises from the cognitive demands of interacting with voicebots, which compromise processing of cost information while not influencing processing of benefit information. This effect attenuates for voice interactions with humans, suggesting a potential boundary condition. The findings make important contributions in the emerging field of consumer-AI interaction.

Keywords Voicebots · Artificial intelligence · Human–machine interaction · Product upgrade · Splurging · Technology

1 Introduction

AI-enabled voicebots (i.e., devices with which consumers interact through speech) are widely used. Rather than click and type, consumers can now use their voice to order household supplies, electronics, snacks, and services such as event tickets or

✉ Yegyu Han
yegyuhan@khu.ac.kr

Sang Kyu Park
sangkyu.park@ust.hk

Aner Sela
aner.sela@warrington.ufl.edu

[1] Hong Kong University of Science and Technology, Hong Kong, Hong Kong SAR

[2] Kyung Hee University, Seoul, South Korea

[3] Warrington College of Business, University of Florida, Gainesville, FL, USA

Published online: 06 December 2025

ready-to-pickup food. Interacting with a voicebot leads consumers to focus on evaluable attributes (Munz & Morwitz, 2022), brace for misunderstandings (Melumad, 2023), and use emotion-related words (Schindler et al., 2024), but does it impact spending? If so, why?

We examine how voicebots influence upgrade decisions, or the tendency to choose premium over more basic options, compared with equivalent interactions using traditional input devices like screens, mice, and keyboards. We find that voicebots increase the choice likelihood of upgrades that provide additional benefits at a higher price. We suggest this happens because voicebots' auditory modality and perceived rigidity place a significant load on consumers' processing capacity (Bostrom & Waldhart, 1988; Leahy & Sweller, 2011). This, in turn, compromises diligent processing of price information, which requires analytical processing, but not reward information, which is processed intuitively (Hsee & Rottenstreich, 2004; Shiv & Fedorikhin, 1999), resulting in increased upgrading. Interestingly, we find that this effect is unique to voicebots. Verbally interacting with humans does not increase upgrading, likely because interpersonal conversations are less taxing and contain additional features that make them easier to process.

This research extends the nascent literature on voicebots (Dellaert et al., 2020; Hildebrand et al., 2020; Melumad, 2023) by showing how and why they influence upgrade decisions. It also extends research on modality and information processing, examining how modality influences cost–benefit trade-offs in human–machine interactions. Further, we identify an important distinction between interpersonal and human–machine communication, which have distinct effects despite their shared modality. Lastly, our findings have practical implications for firms, demonstrating the cues to which consumers are responsive when shopping using voicebots.

2 Theoretical Background

2.1 How voicebots impact information processing

Voicebots are voice-enabled AI that interact with consumers via two-way speech. Whether as standalone devices (e.g., Amazon Echo) or integrated within operating systems (e.g., Apple's Siri), they serve in diverse applications including playing music, ordering products and services, and providing information. Despite being adaptable (e.g., recognizing synonyms and accents), they are often constrained to perform within certain functional boundaries. For example, Alexa is programmed to function in a relatively predictable "if-this-then-that" dialog (Jeffress, 2018).

Voicebots' auditory modality and interactional limitations promote greater focus on evaluable, affect-rich attributes (Munz & Morwitz, 2022) and feeling-based words (Schindler et al., 2024), at the expense of more analytical processing. First, unlike written or visual information, auditory information is ephemeral and therefore places heavier demands on working memory (Bostrom & Waldhart, 1988; Leahy & Sweller, 2011). In contrast, text-based dialogs allow users to offload information from working memory without fear of losing it and process items carefully and analytically (Holtgraves, 2008; Rubin et al., 2000). Because voicebots are inherently

auditory, they are likely to tax working memory, compared with webpages or text-based chatbots.

Second, these processing challenges are amplified by how consumers perceive voicebots. Users often experience them as rigid, marked by "limited interactional abilities" (Fox & Gambino, 2021) and prone to misunderstanding, which increases anticipatory tension (Melumad, 2023). This added strain likely compounds the load already created by the auditory modality itself, further impeding the processing of analytical attributes beyond what occurs in typical conversations with humans.

Together, these findings suggest that voicebots may increase the weight of qualitative, inherently evaluable attributes in the decision process, rather than analytical or quantitative information. We extend this literature by examining how voicebots' influence on attention and processing affects upgrade decisions.

2.2 How cognitive load impacts upgrading

When considering upgrade options, consumers weigh the premium option's price against its enhanced benefits to determine the deal's attractiveness (Okada, 2006; Zeithaml, 1988). For example, choosing a front-row seat over a back-row seat requires evaluating whether the added comfort and enjoyment (i.e., the benefit) justifies the higher price. The less consumers attend to the price difference (vs. added benefit), the more likely they are to upgrade (Sela & LeBoeuf, 2017).

Assessing costs and benefits requires different types of processing. In upgrade decisions, cost is typically represented by a numerical price difference, which has no intrinsic meaning and must be traded-off with other attributes to be evaluated (Simonson & Tversky, 1992; Zeithaml, 1988). This requires deliberative analytical processing, where consumers assess both the price difference and its subjective worth in benefit terms (Deck & Jahedi, 2015; DeStefano & LeFevre, 2004).

Conversely, benefits—whether indulgent or practical in nature—involve attributes that are often intrinsically meaningful and therefore qualitatively evaluable (Hsee, 1996; Nowlis & Simonson, 1997). Benefit features such as "next-day delivery," "front-row seats," or "premium infotainment" are intrinsically associated with positive affect and therefore evaluable even without careful comparisons or analytical processing.

The cognitive dichotomy between processing of costs and benefits aligns with the notion that feelings-based processing reduces sensitivity to quantitative calculations (Hsee & Rottenstreich, 2004). Indeed, when cognitive resources are constrained, people rely on feelings more than analytical processing (Greifeneder et al., 2011; Shiv & Fedorikhin, 1999), as feelings are registered faster and without reflective deliberation (Clore et al., 1994; Pham et al., 2001; Zajonc, 1980). Because evaluating costs requires more analytical thinking than do benefits, we argue that the increased cognitive demands posed by interactions with voicebots hinder the processing of upgrade costs more than benefits. When presented with a choice between a front-row seat for a sports event priced at $174 and a back-row seat at $137, for example, reduced deliberation capacity is likely to hinder the ability to fully grasp the cost (i.e., calculating the price difference and assessing

whether it is worth the experiential benefit), whereas "front row seat" is intrinsically evaluable as attractive without further analysis (Hsee, 1996; Zajonc, 1980). Similarly, "next-day delivery" is an intuitive and readily comprehensible benefit without analytical thinking, whereas any price difference requires some calculation and analysis.

Based on our analysis, we predict that using voicebots reduces consumers' price sensitivity. As a result, in upgrade decisions, consumers become more likely to select higher-priced options with added benefits, as reduced price sensitivity shifts their focus toward the benefits.

We tested our propositions in two different representative upgrade contexts: spending more on products and services that offer greater benefits and paying additional charges to receive products faster. Since all our studies used voicebots, we a-priori excluded participants who encountered technical difficulties during the session. All experimental data and materials are publicly available (https://osf.io/5n3bj/?view_only=426d0ca1b2304db19e6f3b7f8a71948a).

2.3 Pilot study: Evidence from the field

We examined data from a large video-on-demand vendor (company identity remains confidential under a nondisclosure agreement) where customers could rent content using either a remote control or a voicebot connected to their TV (similar to Amazon Fire TV and Alexa). The dataset included VOD purchase information from 15,585 households in 2018, aggregated at the customer level over a 6-month period before and after each customer adopted the company's voicebot. Since customers adopted voicebots on different dates, any effect of adoption on purchase behavior cannot reflect external influences such as seasonality or macro-economic events.

As predicted, households spent 10.25% more, on average, on streaming content during the 6 months after voicebot adoption, compared with the 6 months prior to adoption when they were using their remote control ($F(1, 15{,}584) = 71.67$, $p < .001$, $\eta_p^2 = 0.005$). This result was unchanged when controlling for the number of pages browsed ($F(1, 15{,}582) = 70.54$, $p < .001$, $\eta_p^2 = 0.005$), casting doubt on the possibility that purchases increased because voicebots led to shorter or more impatient searches.

Supporting our upgrading prediction, the average cost of each movie rented was 21.81% higher after the adoption ($F(1, 15{,}584) = 622.3$, $p < .001$, $\eta_p^2 = 0.038$), despite the number of rentals being 3.24% lower ($F(1, 15{,}584) = 6.60$, $p = .01$, $\eta_p^2 < 0.01$). This indicates that customers purchased more expensive content after adopting voicebots. Both results remain unchanged when controlling for the pre-purchase time spent on browsing ($F(1, 15{,}582) = 607.84$, $p < .001$, $\eta_p^2 = 0.038$; $F(1, 15{,}582) = 5.05$, $p = .025$, $\eta_p^2 < 0.01$, respectively), contrary to an alternative impatience account.

These findings align with our prediction but they do not indicate causality due to the data's correlational nature. The following experiments test the causal relationship and mechanism more directly.

3 Study 1: Causal effects on upgrades in consequential choice

Study 1 examines the causal effect of voicebot interactions on upgrade likelihood, compared with similar interactions with screen-based web interfaces, using a fully consequential choice. Participants were endowed with money through a preliminary task and then used some of it to either buy a basic option (a cup of coffee) or upgrade to a coffee and snack combo.

3.1 Method

Undergraduate students ($M_{age} = 20.0$, 72.5% women; $N = 171$ after excluding 45 participants who had dietary restrictions for the products offered in the study and 8 participants who experienced technical difficulties), recruited from two different labs to expedite data collection, were randomly assigned to a between-subjects condition: voicebot or webpage. Consistent with recommendations made by Nelson et al. (2018), we preregistered both sessions, which followed an identical protocol (https://aspredicted.org/7VP_M9B; https://aspredicted.org/Z9V_3R4). Lab location did not interact with our independent variable ($\chi^2(1) = .08$, $p = .771$), so we combined the data in further analyses.

We gave all the participants a small sum of money ($3 USD or $25 HKD) and told them that they had the option to use the money to buy various items during the study, but could keep any unused funds for themselves. Then, we told participants that they could buy beverages and snacks from a major coffee chain and asked them to indicate if they drank coffee (those who said they never drank coffee were excluded from the analysis). Participants in the voicebot condition learned about the details of the offer as they interacted with the voicebot, which informed them auditorily that they could use the money to buy a medium brewed coffee from a major coffee chain for $1.50 (or $17 HKD) or upgrade to a combo that includes a cookie or a doughnut for $2.75 (or $23 HKD). For example, the voicebot said "You can buy a medium coffee from Dunkin Donuts for $1.50. You can also choose to upgrade it to a coffee and doughnut combo for only $2.75, which includes a medium coffee and a classic doughnut. Which would you like to choose?" (see Web Appendix A for details). Participants in the webpage condition read the same script in text form. We gave them their chosen option(s) along with any leftover change, in cash, as they left the lab.

3.2 Results

Consistent with our prediction, participants in the voicebot condition were more likely to upgrade to the combo option (74.1%), compared to the ones assigned to the webpage condition (60.0%; $\chi^2(1) = 3.80$, $p = .051$).

3.3 Discussion

Study 1 provides initial causal evidence that, compared with a webpage format, interacting with a voicebot leads consumers to choose more expensive upgrades offering enhanced benefits. We replicated this pattern in two additional contexts, reported as Study WA1 and Study WA2 in Web Appendices B and C). In Study WA1 ($N=191$), using an incentive-compatible design, voicebot users were more likely to choose a pricier expedited shipping option over a cheaper, slower one (44.4% vs. 30.0%; $\chi^2(1)=4.22$, $p=.040$). In Study WA2 ($N=175$), interacting with a voicebot increased selection of a more expensive stadium seat with a better view over a cheaper seat with a worse view (74.0% vs. 20.4%; $\chi^2(1)=50.31$, $p<.001$). Together, these results bolster the generalizability of our findings across decision types and contexts.

4 Study 2: Voicebot vs. Text-based chatbot

We conceptualized that the effect of voicebots on upgrading is driven by increased cognitive demand stemming from the ephemeral auditory modality of interaction. Study 2 tests this by adding a text-based chatbot. We developed a chatbot using the same platform on which the voicebot was built and scripted it to interact in the same way as the voicebot, only using text instead of speech. The chatbot condition allows us to test whether the effect is driven by modality or, rather, the dialogic nature of the interaction (preregistered at https://aspredicted.org/4NF_YHR).

4.1 Method

Undergraduate students ($M_{age}=20.5$, 60.6% women; $N=218$ after excluding 12 who had technical difficulties) completed the study in a lab. Participants were randomly assigned to one of three between-subjects conditions: voicebot vs. chatbot vs. webpage.

All participants saw an image of a new iPad Pro which they presumably wanted to buy (see Web Appendix D). Participants in the voicebot and chatbot conditions interacted with a bot to complete their order using voice or text dialog, respectively. The bot first asked which color of iPad the participant wanted and then described two shipping options: 14-day normal shipping for $2.75 vs. 2-day expedited shipping for $11.75. Participants indicated their choices by speaking or typing. Aside from interaction modality, bot dialogs and interactions were identical in these conditions. Participants in the webpage condition chose their preferred options on a computer screen without interacting with an AI-enabled bot. Information content was held constant across conditions.

4.2 Results

Participants in the voicebot condition were more likely to upgrade to expedited shipping (51.9%) than in the chatbot condition (23.8%; $\chi^2(1)=11.22$, $p<.001$) and webpage condition (23.2%; $\chi^2(1)=11.67$, $p<.001$; omnibus $\chi^2(2)=15.17$, $p<.001$,

See Fig. 1). There was no difference between the chatbot and webpage conditions ($\chi^2(1)=.01$, $p=.923$).

4.3 Discussion

Study 2 replicates the effect of voicebot interaction on upgrade likelihood, compared with screen-based and an AI-enabled text-based chatbot interface. The text-based chatbot condition underscores the role of auditory modality in our effect.

5 Study 3a: Processing price vs. Reward information

We theorized that voicebot interactions tax cognitive resources, consequently hindering processing of cost but not benefit information. Study 3a tests this mechanism by examining recognition accuracy for transaction cost and benefit details. We predicted that compromised processing of cost—but not benefit—information would mediate the effect on upgrades.

5.1 Method

Undergraduate students ($M_{age}=19.8$, 62.0% women; $N=234$, after excluding 25 participants who had technical difficulties) were randomly assigned to a voicebot or webpage condition.

Participants imagined buying a box of chocolates for themselves. We told those in the voicebot condition they could place the order using Amazon's Alexa. Alexa told participants that a 36-piece standard gift box cost $20.99, but they had the option to buy a luxury gift box originally priced at $36.99 and currently on sale for $34.99. Alexa further said, "The luxury gift box additionally offers truffles and chocolate

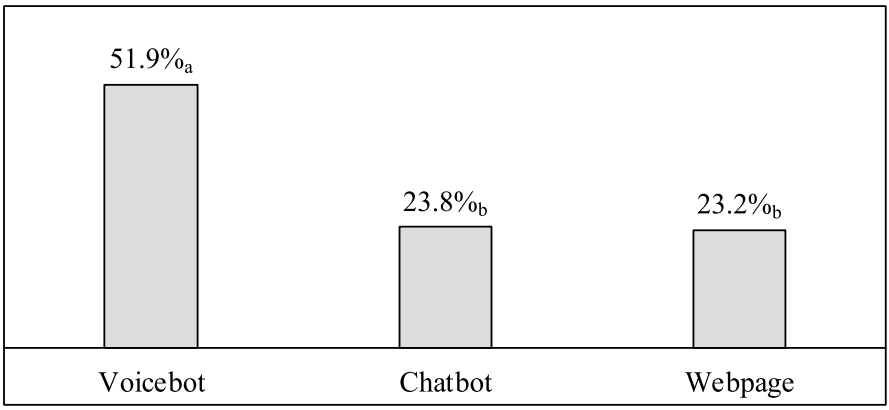


Fig. 1 Choice of expedited shipping (Study 2). *Note:* Percentages with different subscripts indicate significant difference $p<.001$

bonbons crafted with premium cacao beans from Ghana, Ecuador, and Madagascar, with premium flavors including crunchy hazelnut with gold flakes, strawberry mousse with vanilla filling, and key lime with a twist of lemon." Alexa then asked participants whether they would like to upgrade to the luxury gift box. Participants in the webpage condition read the same information on their screen and decided which box to purchase.

After deciding, participants answered six recognition questions about the cost (prices) and benefits (flavors) of the options they were presented with, by selecting from five possible responses for each question (see Web Appendix E). The number of correct cost responses and benefit responses, respectively, served as separate measures of information processing.

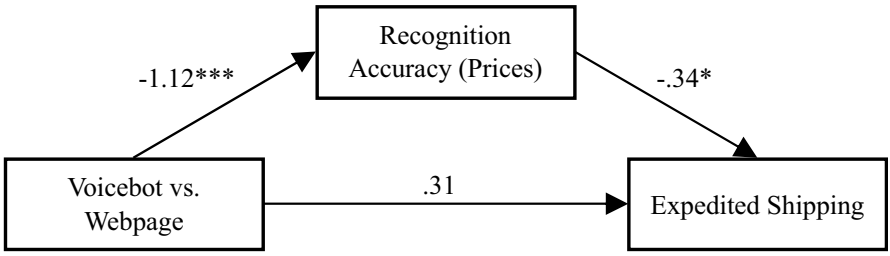
5.2 Results

Participants in the voicebot condition were more likely to upgrade to the luxury chocolate box (voicebot = 51.5% vs. webpage = 35.0%; $\chi^2(1) = 6.36$, $p = .012$). Furthermore, they recognized price information less accurately than those in the webpage condition ($M_{voicebot} = 1.35$, SD = .94 vs. $M_{webpage} = 2.47$, SD = .81; $F(1, 232) = 95.41$, $p < .001$), but there was no difference in recognizing benefit information ($M_{voicebot} = 1.51$, SD = .84 vs. $M_{webpage} = 1.51$, SD = .93; $F(1, 232) < .01$, $p = .961$).

Recognition accuracy for price information mediated the effect of interaction modality on upgrading (Indirect effect = .31, SE = .19, 95% CI [.02,.77], 5000 bootstrapped samples). See Fig. 2.

6 Study 3b: Mediation by sensitivity to price differences

Study 3b examines an alternative operationalization of price processing: perceived magnitude of price differences. Greater elaboration begets more extreme judgments (Park & Sela, 2020; Tesser et al., 1995; see Judd & Brauer, 1995 for a review). Thus, we predicted voicebot interactions would diminish perceived price differences, by decreasing elaboration on price information.



*$p < .05$, ***$p < .001$.

Fig. 2 Mediation analysis (Study 3a)

6.1 Method

Undergraduate students (M_{age} = 21.0, 46.3% women; N = 216 after excluding 22 participants who had technical difficulties) were randomly assigned to a voicebot or webpage condition. The design was identical to Study 3a, with two exceptions. First, we used a different product (i.e., Bose headphones; see Web Appendix F). Second, after choosing a shipping option, participants indicated how big the price difference between standard and expedited shipping seemed (0 = very small, 100 = very big).

6.2 Results

Replicating our previous results, participants were more likely to upgrade to expedited shipping when using a voicebot (34.3%) than a webpage (22.2%; $\chi^2(1) = 3.93$, $p = .048$). Participants in the voicebot condition also perceived the price difference between the two options as subjectively smaller ($M_{voicebot} = 46.83$, SD = 27.89 vs. $M_{webpage} = 58.84$, SD = 26.38; $F(1, 214) = 10.55$, $p = .001$). Perceived price difference, in turn, mediated the effect of interacting with a voicebot on upgrading (Indirect effect = .70, SE = .24, 95% CI [.27, 1.23], 5000 bootstrapped samples). See Fig. 3.

6.3 Discussion

Studies 3a and 3b provide convergent evidence for our proposed mechanism: interacting with voicebots compromises the processing of price—but not benefit—information, which in turn increases upgrading. Web Appendices G and H report two additional process studies that bolster the cost processing account. In Study WA3 (N = 197), we observe a similar mediation pattern using unaided price recall accuracy as a mediator. Study WA4 (N = 256), using a moderation-of-process approach, shows that when participants' analytical processing is constrained, the difference between web and voicebot conditions disappears.

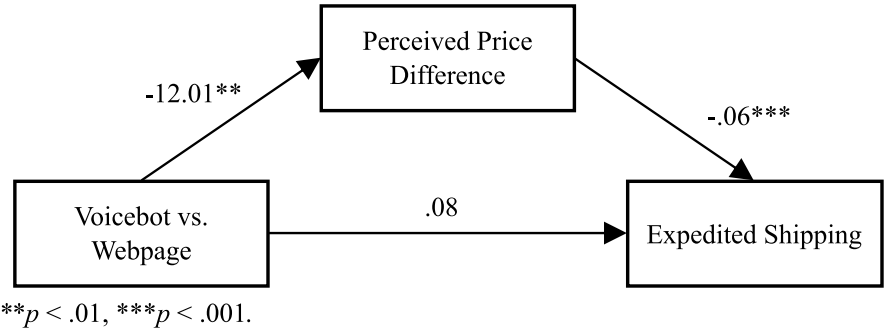


Fig. 3 Mediation analysis (Study 3b)

7 Study 4: Human interactions as a boundary condition

Study 4 tests whether the effect observed with voicebots extends to human agents. While both rely on auditory interaction, human-to-human conversations are experienced as easier and more natural. Compared to voicebots, human interlocutors engage in more adaptive conversational behaviors (i.e., interactive alignment)—such as adjusting pace and vocabulary, repeating information, and proactively seeking clarification—to facilitate mutual understanding and share cognitive load between participants (Clark & Brennan, 1991; Garrod & Pickering, 2004; Pickering & Garrod, 2004). Just as important, people *expect* human conversations to be flexible and accommodating, which reduces anticipatory tension relative to the rigidity often associated with voicebot interactions (Melumad, 2023). Consequently, we expected conversations with humans to be less cognitively taxing, despite sharing the same auditory modality.

As a preliminary test, we asked 100 consumers to rate how cognitively demanding they found interactions with voicebots compared to humans (e.g., "Compared to daily conversations with others, when I speak to voice assistants, I feel like I need to be more robotic," "…I need to pay more attention to the conversation," "… I cannot focus on the information they are saying as carefully as I would like to"; *1 = not at all agree, 4 = neither agree nor disagree, 7 = completely agree*; $\alpha = .72$; see Web Appendix I for a full list of items). Supporting our prediction, participants rated voicebot interactions as significantly more stressful and taxing, well above the scale's midpoint (4.96; $t(99) = 9.56$, $p < .001$).

To stress-test our prediction, we included two versions of the human agent condition: a baseline condition and an *enhanced* condition in which participants were prompted to pay close attention because the agent was a trainee and might be rigid or have difficulty communicating clearly. We expected upgrade likelihood to be higher in the voicebot condition than in either human condition, as human-to-human conversations are inherently more flexible and responsive—even when suboptimal—and thus less cognitively taxing than interactions with voicebots. Demonstrating this boundary condition would also rule out alternative accounts based on social presence (Epley et al., 2007) or verbal compliance (Mariadassou et al., 2023).

7.1 Method

Study 4 was preregistered (https://aspredicted.org/L1Z_YQM). American Prolific workers ($M_{age} = 37.9$, 53.0% women; $N = 115$) were randomly assigned to one of three conditions: voicebot, human agent, or enhanced human agent. Similar to Study 2, participants chose between normal ($3.50, 10-day) or expedited shipping ($9.50, 2-day) for AirPods Pro. Participants in the voicebot condition made their choice via voicebots. Those in the human agent condition completed their order through an audio-only Google Meet call. In the enhanced human condition, participants were additionally informed that the agent was a new trainee who might be rigid or have difficulty communicating clearly. We operated a mini "call center" where trained

RAs, blind to the hypotheses, followed the same script as the voicebot. RAs were instructed not to provide any additional information deviating from the script.

7.2 Results

Results confirmed our predictions. Upgrade likelihood was higher in the voicebot condition (62.9%) than in both the baseline human (39.5%; $\chi^2(1) = 3.99$, $p = .046$) and enhanced human condition (35.7%; $\chi^2(1) = 5.64$, $p = .018$), which did not differ from each other ($\chi^2(1) = .120$, $p = .729$; omnibus $\chi^2(2) = 6.43$, $p = .040$). See Fig. 4.

7.3 Discussion

Study 4 identified an important boundary condition: voicebots increase upgrading even compared with human agents delivering an identical script. Study WA5 ($N = 191$), reported in Web Appendix J, replicated these findings: participants using Amazon Alexa were more likely to upgrade (45.3%) compared to those interacting with a human agent (19.6%; $p = .004$) or webpage (26.3%; $p = .019$). These findings suggest that interactions with voicebots may be unique, despite sharing modality with person-to-person conversations.

8 General discussion

Voice shopping is a rising trend. This research extends the literature on technology-driven consumer behavior by examining how they affect upgrade decisions.

Our research makes several theoretical contributions. First, we contribute to the emerging literature on voicebots (Dellaert et al., 2020; Hildebrand et al., 2020; Melumad, 2023) by examining their influence on upgrade decisions—an important

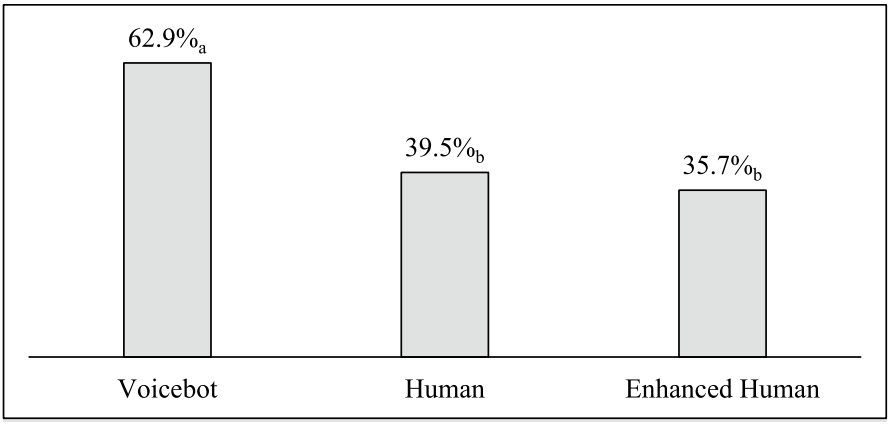


Fig. 4 Choice of expedited shipping in Study 4. *Note:* Percentages with different subscripts indicate significant difference $p < .05$

consumer context that has received little attention. Whereas prior work studied how voicebots shape what consumers say, we examined how they influence what information consumers process and, consequently, what they choose.

Second, we extend research on interaction modalities (Leahy & Sweller, 2011; Mariadassou et al., 2023) by showing that the cognitive demands of auditory interfaces make it harder to process costs, increasing upgrading. Third, we bridge human–machine and human–human communication. Although voicebots and people both use speech, interacting with them may not produce the same outcomes.

Our findings offer practical implications. When using voicebots, emphasizing vivid, benefit-driven descriptions over price comparisons might be more effective. Price advantages tend to get lost in voice, while emotionally resonant benefits carry more weight. That said, some upgrade benefits may still require thoughtful evaluation—possibly limiting this effect. Future research should explore this boundary condition.

8.1 Speaking vs. listening

Since conversations involve both speaking and listening, one might ask which drives the upgrading effect. To test this, we ran a supplementary study with two added conditions: a listen-only condition, where participants heard the voicebot but responded via keyboard or mouse, and a speak-only condition where they saw the information on-screen (as in the chatbot condition in Study 2) but responded aloud into a microphone. The upgrading effect replicated in the listen-only condition but not in the speak-only condition (see Web Appendix K). This suggests that it is the processing of transient, auditory information—compounded by the rigidity of voicebot interactions, such as limited chances for clarification—that drives the effect, rather than speaking aloud.

8.2 Alternative explanations and future research

One possible alternative is that the increased cognitive effort of interacting with voicebots increases involvement. But if voicebots raise involvement as a motivational state, we would expect consumers to be more deliberative and careful—not less. Our data show the opposite: voicebot users process price information less carefully, not more.

Another possibility is that voicebot interactions improve mood, which in turn increases upgrading. While we did not measure mood directly, the pretest reported in the introduction to Study 4 suggests otherwise. Participants described voicebot interactions as more stressful and demanding than other modalities—hardly consistent with a positive mood-based explanation.

A third possibility is that the upgrading effect is driven by novelty. This is a compelling question for future research, especially as voicebots become more familiar. That said, in an unreported study, we found the opposite pattern: participants with more experience using voicebots showed a stronger upgrading effect than those with less experience. This finding aligns with our account—experienced voicebot users

may anticipate more communication breakdowns (Fox & Gambino, 2021)—but runs counter to a novelty-based explanation. Future research should investigate this possibility more directly.

Relatedly, future work should consider how improvements in voicebot technology might moderate these effects. As AI improves and enables smoother, more humanlike conversation, the cognitive load we document may diminish. Whether that reduces the tendency to upgrade remains an open question, but by capturing responses to today's voicebots, our findings provide a useful benchmark for tracking how these effects evolve over time.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11002-025-09806-1.

Funding This research was partially funded by the Spanish Ministry of Science and Innovation, State Research Agency (AEI) 10.13039/501100011033 Grant No. PID2020-119622GA-I00.

Data Availability All experimental data and materials, including the survey and preregistration documents, are publicly available in the Open Science repository (https://osf.io/5n3bj/?view_only=426d0ca1b2304db19e6f3b7f8a71948a).

Declarations

Ethical approval We conducted our studies that involved human subjects with the approval of the corresponding institutional review board (IRB).

Consent to participate Subjects participated in our experiments with informed consent approved by the corresponding IRB.

Conflict of interest The authors declare no competing interests.

References

Bostrom, R. N., & Waldhart, E. S. (1988). Memory models and the measurement of listening. *Communication Education, 37*(1), 1–13.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association.

Clore, G. L., Schwarz, N., & Conway, M. (1994). Affective causes and consequences of social information processing. *Handbook of Social Cognition, Hillsdale* (pp. 323–417). Erlbaum.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review, 78*, 97–119.

Dellaert, B. G., Shu, S. B., Arentze, T. A., et al. (2020). Consumer decisions with artificially intelligent voice assistants. *Marketing Letters, 31*(4), 335–347.

DeStefano, D., & LeFevre, J. A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*(3), 353–386.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864–886.

Fox, J., & Gambino, A. (2021). Relationship development with humanoid social robots: Applying interpersonal theories to human–robot interaction. *Cyberpsychology, Behavior, and Social Networking, 24*(5), 294–299.

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences, 8*(1), 8–11.

Greifeneder, R., Bless, H., & Pham, M. T. (2011). When do people rely on affective and cognitive feelings in judgment? A review. *Personality and Social Psychology Review, 15*(2), 107–141.

Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. *Journal of Business Research, 121*, 364–374.

Holtgraves, T. (2008). Automatic intention recognition in conversation processing. *Journal of Memory and Language, 58*(3), 627–645.

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational Behavior and Human Decision Processes, 67*(3), 247–257.

Hsee, C. K., & Rottenstreich, Y. (2004). Music, pandas, and muggers: On the affective psychology of value. *Journal of Experimental Psychology: General, 133*(1), 23–30.

Jeffress, J. (2018). Using yes/no intents with dialog management. Retrieved from https://developer.amazon.com/en-US/blogs/alexa/alexa-skills-kit/2018/04/using-yes-no-intents-with-dialog-management

Judd, C. M., & Brauer, M. (1995). Repetition and evaluative extremity. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 43–72). Lawrence Erlbaum Associates.

Leahy, W., & Sweller, J. (2011). Cognitive load theory, modality of presentation and the transient information effect. *Applied Cognitive Psychology, 25*(6), 943–951.

Mariadassou, S., Bechler, C., & Levav, J. (2023). The effect of auditory and visual recommendations on choice. *Psychological Science, 34*(1), 47–59.

Melumad, S. (2023). Vocalizing search: How voice technologies alter consumer search processes and satisfaction. *Journal of Consumer Research, 50*(3), 533–553.

Munz, K. P., Morwitz, V. G. (2022). Sound judgment: Evaluability and memory in speech-based product evaluation and choice. *Working Paper* Available at SSRN: https://ssrn.com/abstract=3462714

Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's renaissance. *Annual Review of Psychology, 69*(1), 511–534.

Nowlis, S. M., & Simonson, I. (1997). Attribute–task compatibility as a determinant of consumer preference reversals. *Journal of Marketing Research, 34*(2), 205–218.

Okada, E. M. (2006). Upgrades and new purchases. *Journal of Marketing, 70*(4), 92–102.

Park, S. K., & Sela, A. (2020). Product lineups: The more you search, the less you find. *Journal of Consumer Research, 47*(1), 40–55.

Pham, M. T., Cohen, J. B., Pracejus, J. W., & Hughes, G. D. (2001). Affect monitoring and the primacy of feelings in judgment. *Journal of Consumer Research, 28*(2), 167–188.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*(2), 169–190.

Rubin, D. L., Hafer, T., & Arata, K. (2000). Reading and listening to oral-based versus literate-based discourse. *Communication Education, 49*(2), 121–133.

Schindler, D., Maiberger, T., Koschate-Fischer, N., & Hoyer, W. D. (2024). How speaking versus writing to conversational agents shapes consumers' choice and choice satisfaction. *Journal of the Academy of Marketing Science, 52*(3), 634–652.

Sela, A., & LeBoeuf, R. A. (2017). Comparison neglect in upgrade decisions. *Journal of Marketing Research, 54*(4), 556–571.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research, 26*(3), 278–292.

Simonson, I., & Tversky, A. (1992). Choice in context: Tradeoff contrast and extremeness aversion. *Journal of Marketing Research, 29*(3), 281–295.

Tesser, A., Martin, L., & Mendolia, M. (1995). The impact of thought on attitude extremity and attitude-behavior consistency. In R. E. Petty, & J. A. Krosnick (Eds.), *Attitude Strength* (pp. 73–92). Psychology Press.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(2), 151–175.

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing, 52*(3), 2–22.