

# Can user choice alter experimental findings in human-computer interaction? : similarity attraction versus cognitive dissonance in social responses to synthetic speech

Lee, Kwan Min; Jung, Younbo; Nass, Clifford

2011

Lee, K. M., Jung, Y., & Nass, C. (2011). Can user choice alter experimental findings in human-computer interaction? : similarity attraction versus cognitive dissonance in social responses to synthetic speech. *International journal of human-computer interaction*, 27(4), 307-322.

<https://hdl.handle.net/10356/100909>

<https://doi.org/10.1080/10447318.2011.540473>

---

© 2011 Taylor & Francis Group, LLC. This is the author created version of a work that has been peer reviewed and accepted for publication by *International Journal of Human-Computer Interaction*, Taylor & Francis Group, LLC. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at: [DOI:<http://dx.doi.org/10.1080/10447318.2011.540473> ].

Running Head: USER CHOICE AND COMPUTER-SYNTHESIZED SPEECH

Can User Choice Alter Experimental Findings in Human Computer Interaction?:

Similarity Attraction vs. Cognitive Dissonance in Social Responses to Synthetic Speech

Kwan Min Lee, Associate Professor

Annenberg School for Communication, University of Southern California

Department of Interaction Science, Sungkyunkwan University

Los Angeles, CA 90089

1-323-632-9865, [kwanmin.lee@usc.edu](mailto:kwanmin.lee@usc.edu)

Younbo Jung, Assistant Professor

Wee Kim Wee School of Communication & Information

Nanyang Technological University

31 Nanyang Link, Singapore 637718

65-6316-8968, [ybjung@ntu.edu.sg](mailto:ybjung@ntu.edu.sg)

Clifford Nass, Professor

Department of Communication, Stanford University

Stanford, CA 94305-2050

1-650-723-5499, [nass@stanford.edu](mailto:nass@stanford.edu)

### Abstract

In this study, we investigate the effect of the user choice on social responses to computer-synthesized speech. Three previous findings about social responses to computer-synthesized speech (i.e., social identification, proximate source orientation, and similarity attraction) were tested using the choice paradigm. Social identification and proximate source orientation effects were found even when users had chosen a computer voice at their discretion. In addition, the primacy effect in the user choice prevailed: Participants were more likely to select whatever voice that they heard first between two options. The similarity attraction effect, however, was negated by the cognitive dissonance effect after user choices. We discuss the robustness of social responses, its implications for human-computer interaction, and the importance of the user choice in voice-interface designs.

### Keywords

Speech user interfaces, user choice, TTS (Text-to-Speech), computers are social actors (CASA), similarity attraction effect, proximate source orientation, primacy effect, cognitive dissonance.

## Can User Choice Alter Experimental Findings in Human Computer Interaction?:

### Similarity Attraction vs. Cognitive Dissonance in Social Responses to Synthetic Speech

“You didn’t come here to make the choice. You’ve already made it. You are here to try to understand why you made it” (conversation between Neo and the Oracle in the movie, “Matrix Reloaded”).

The experimental method provides researchers with great controllability over variables, subjects, and the overall research environments that helps establish causal relationships among the variables (Wimmer & Dominick, 2000). In this regard, researchers have used experiments as one of the key research methods for the study of Human-Computer-Interaction (HCI) because identifying human factors (e.g., independent variables such as the size, movement, personality, gender, or anthropomorphism of the objects) that would trigger a particular type of human responses is a central research question in HCI.

However, most previous HCI experiments have a limitation in that an experimenter *assigns* users to a strictly controlled experimental condition in which the users are forced to accept whatever stimuli pre-determined by the experimenter. This is an ontological limitation of socio-psychological research where validity rules over generalizability. In contrast to experimental settings, users in real life are usually provided with the plethora of options that they can choose from. For example, almost every *text-to-speech* (TTS) system provides pre-set speech parameters so that users can choose whatever TTS voices they want to hear. Despite the ubiquity of user choice in HCI, the effect of the user choice in HCI has seldom been investigated because it creates extra noises and random errors and thus invites criticism about the lack of control in the

experiment. As a consequence, the findings from previous studies need to be cautiously interpreted when they are applied to real-life HCI situations in which users freely choose interfaces or devices.

Therefore, empirical re-tests of previous HCI experiments in the context of free-user-choice are important, not only for practical reasons, but also for theoretical reasons. Making choices is an important concept that has been the subject of social psychology and many other disciplines (e.g., mathematics, statistics, economics, political science, and sociology; see Kahneman & Tversky, 1984). For example, there have been empirical studies about the process of making the right choice via unconscious thought (Dijksterhuis, Bos, Nordgren, & van Baaren, 2006), the choice between rewards available at different time points (McClure, Laibson, Loewenstein, & Cohen, 2004), and the consequences of the choice with regards to happiness (Schwartz, Ward, Monterosso, Lyubomirsky, White, & Lehman, 2002; Van Boven, & Gilovich, 2003). In addition, theoretical investigation into the effects of the user choice in HCI is practically important because the free choice is a de-facto norm in interface or device design from the choice of a particular product to subtle personalization of purchased interfaces and devices.

In order to provide both practical guidelines and theoretical insights into the effects of the user choice in HCI, the current study replicates a previous study in a free-choice context. That is, we replicate a previous study in the exactly same way as it was conducted before, except that participants in the current study choose their interface instead of being assigned to one. Because we were able to get an access of all experimental stimuli and questionnaires, we chose to replicate an experiment originally reported in Nass and Lee (2001) that examined three social responses to synthesized voices. The original study is one of the flagship *Computers Are Social*

*Actors* (CASA) studies that have been quite influential in the design of social interface over the last decade (e.g., Detenber & Reeves, 1996; Isbister & Nass, 2000; Lee & Nass, 2004; Nass & Lee, 2001; Nass & Moon, 2000; Nass, Moon, & Carney, 1999; Reeves, Detenber, & Steuer, 1993; Reeves & Nass, 1996; Sundar & Nass, 2001).

### Social Responses to Synthesized Voices

Based on findings from social and personality psychology, Nass and Lee (2001) discovered a variety of social responses to computer-synthesized voices (see also Qiu & Bendasat, 2005 for online consumer trust as social responses to computer-synthesized voices). More specifically, they found three types of social responses to computer-synthesized voices; social identification, proximate source orientation, and similarity attraction.

First of all, people easily recognize the social identity of computer-synthesized voices, such as personality (see Pianesi, Mana, Cappelletti, Lepri, & Zancanaro [2008] for a review of multimodal recognition of personality traits). It should be no surprise that people recognize the gender and age of computer-synthesized voices, because the identification of gender and age is so evolutionarily hard-wired into our brain (see Lee and Jung, 2005 for more detailed discussions about the evolutionary nature of virtual experience). For example, any type of speech would be identified as either male or female in order to determine mating possibility. In a similar vein, the discrimination of age from a voice would also be automatic in order to recognize potential threats from fully-grown adults. The perception of social identities such as personality from a computer voice, however, is clearly challenging in that the perception of personality from a voice is arguably acquired in later stages of life and, for that reason, does not seem to be obviously programmed in our brain. The study by Nass and Lee (2001) discovered that our social responses to computer voices are so profound that people can identify even the personality of a computer

voice. They reported very clear evidence that people do recognize personalities in computer voices.

Secondly, people attribute the personality of a remote source (e.g., the reviewer of a text) to the personality of a proximate source (e.g. the computer voice that narrated the text). For instance, people perceive the personality of the reviewer extroverted when the descriptions are narrated by the extroverted computer voice. This proximate source orientation phenomenon happens even when people are precisely aware of the clear dissociation between the computer voice and the reviewer (Nass & Lee, 2001). This finding indicates how mindlessly users respond to a social cue of computer-synthesized voices even though they know that the computer voices have nothing to do with the reviewer or the content.

Finally, people show strong similarity attraction tendencies to computer-synthesized voices. The similarity attraction posits that individuals are more attracted to other people who are similar to themselves or who show similar personalities regardless of their awareness of the personalities (Byrne & Griffitt, 1969; Byrne, Griffitt, & Stefaniak, 1967). The areas of perceived similarity could range from physical appearance and personality traits to opinions, background or lifestyle (Cialdini, 2001). Once the perceived similarity is established, people tend to show social responses of likeness, positive evaluation, and/or compliance towards their communication partners. For example, a salesperson always asks potential customers questions about their backgrounds, favorite sports, or hobbies, hoping to find any similarity between the salesperson and the customers. In the Nass & Lee study (2001), participants liked and trusted a voice exhibiting a personality similar to their own more than one exhibiting a dissimilar personality, and they evaluated the same content more positively when a computer voice narrating the content exhibited the personality similar to their own. Moreover, participants liked and trusted

the same reviewer more when the reviewer's writing was narrated by a computer voice manifesting the personality similar to their own.

Nonetheless, one critical limitation of the above-mentioned findings is that all of them were discovered from strictly controlled experiments in which users were randomly assigned to one of the manipulated conditions without being provided with alternative choices. Because of this ontological limitation, these findings may not hold up in the field where people freely make a choice among alternatives. The current study replicates the previous study by Nass & Lee (2001) in order to answer the following research question:

*RQ1. Will users show the three types of social responses (social identification, proximate source orientation, and similarity attraction) to computer voices even when they choose a computer voice at their own discretion?*

#### After Choice: Cognitive Dissonance

Our everyday life is full of making choices from what to eat to when to sleep. Once choices are made, knowingly or unknowingly, consequences follow. According to Schwartz (2004), the mere action of a voluntary selection itself is psychologically powerful enough to influence the evaluation of the choice. For example, a study by Schraw, Flowerday, and Reisetter (1998) demonstrates that making voluntary choices increases the affective engagement of people who have made the choices. Specifically, participants who were given choice options showed more favorable attitudes regarding their participation, greater perceived control over the experiment, and more interests in the story (Schraw et al., 1998; see also Flowerday & Schraw, 2003). In a similar vein, studies indicate that learners with choice options (e.g., self-controlled learning environments) are more satisfied with their learning and performance than learners without choice options (Hannafin & Sullivan, 1996; Yang & Chin, 1997).



One of the most influential theories for individuals' post-choice states is cognitive dissonance (Festinger, 1957). There is a natural tendency for people to maintain consistency between their belief and their behavior after making choices because good personal consistency is highly valued in our society. People even assure themselves that their choice is right "because of the need to be consistent within their system of beliefs" (Cialdini, 2001: p. 85). To put it differently, when an individual experiences the presence of inconsistency between what the individual believes and what the individual does, there tends to be psychological discomfort, cognitive dissonance, in the person's mind (Festinger, 1957). In the existence of dissonance, people try not only to reduce the dissonance but also to actively avoid situations that would likely increase the dissonance. That is, people tend to change their behaviors or thoughts in order to achieve consonance when they face cognitive dissonance. According to Festinger (1957), the simplest and easiest way in which people can eliminate the dissonance is to change the action of the choice or to revoke psychological elements involving the dissonance.

In real life, it is often hard to undo the behavioral part of the user choice because it is cumbersome or simply impossible to undo. As a consequence, people tend to reinforce their minds to like whatever choices that they have made in order to eliminate the dissonance. Changing one's mind after making choices usually results in the increased confidence in the decision and the difficulty of reversing the choice (Festinger, 1957). Having said that, incorporating the choice paradigm into HCI experiments is likely to influence people's social responses to computer-synthesized voices and, especially, people's evaluation of computer voices (see Oulasvirta, Hukkinen, & Schwartz [2009] for effects of choice in a different HCI context). Since the current study is exactly the same as the Experiment 1 of Nass & Lee (2001) except for the free-choice manipulation, it provides a unique opportunity to test the effects of the

user choice in HCI. Therefore, we will try to answer the following research question by combining the data from the current and previous study.

*RQ2: Does having a choice matter in interface and content evaluations?*

## Method

### *Experimental Design*

This study was executed in the context of a book-buying website used in the previous study by Nass & Lee (2001). The only difference between the current and previous study is that the participants in this experiment went to a preliminary website first and heard both extrovert and introvert computer voices. Then, the participants in the current study chose whatever voice they wanted to hear in the main website. This differs from the previous study, where the participants were randomly assigned to either an extroverted or introverted computer voice without the opportunity for making a choice.

The main website was designed the exactly same way as the one used in the previous study by Nass & Lee (2001). More specifically, each web page contained an identical visual interface based on the interface of book descriptions at Amazon.com, including the titles, the reviewers (in text) and the pictures of five books. Instead of having the book descriptions in text, there were links to audio (.wav) files; clicking on the link would play the review. After hearing the reviews, the participants answered questions listed on the same web page (see Nass & Lee, 2001 for detail). A total of 94 college students enrolled in a large introductory course at a private university in the West Coast of the United States participated in the experiment.

### *Procedure*

In the previous study done by Nass & Lee (2001), all participants (N=72) were either clearly introverted (N=36) or clearly extroverted (N=36) to fully investigate the similarity

attraction effect. In a similar vein, we conducted the exact same personality tests (a short form of the Myers-Briggs Type Indicator and Wiggins personality tests) to identify both personality types. Based on the personality scores, we collected 64 responses (responses from 16 introverts who chose the introverted voice; responses from 16 introverts who chose the extroverted voice; responses from 16 extroverts who chose the introverted voice; and responses from 16 extroverts who chose the extroverted voice) from the total of 94 responses. Thirty participants who did not show clear personality types were not included in our final analysis. This procedure enabled us to conduct the same 2 (computer voice personality: extrovert vs. introvert) by 2 (participant personality: extrovert vs. introvert) full factorial *analysis of variance* (ANOVA) done in the previous study, as well as a combined analysis for the three way interaction effects (computer personality vs. participant personality vs. choice).

### *Measures*

We used the exactly same measures as the ones used in the previous study by Nass & Lee (2001), except for the quality of the review and users' buying intention that are irrelevant for our analysis. All measures were based on items from the web-based questionnaires with an independent, ten-point Likert scale for each question.

Extrovertedness-introvertedness was measured for both the TTS voice (Cronbach's  $\alpha = .85$ ) and the reviewer (Cronbach's  $\alpha = .87$ ) by using an index of Wiggins (1979) 10 personality adjective: cheerful, enthusiastic, extroverted, introverted (reverse coded), inward (reverse coded), jovial, outgoing, perky, shy (reverse coded) and vivacious. The higher the score, the more extroverted the TTS voice or the reviewer is.

Five questions about the liking of the voice were asked using a combination of independent 10-point scales and 10-point semantic differential scales: "How much did you enjoy

hearing the computer voice?,” “How likely would you be to have the voice read you other descriptions?,” and the following adjectives: enjoyable, likable, satisfying, and enjoy listening (Cronbach’s  $\alpha = .80$ ).

The liking of the reviewer was measured by an index composed of three adjectives: enjoyable, likable, and satisfying (Cronbach’s  $\alpha = .87$ ).

The credibility of the voice was measured by an index composed of three adjectives: credible, reliable, and trustworthy (Cronbach’s  $\alpha = .88$ ).

Finally, the credibility of the reviewer was measured by Wheelless and Grotz’s (1977) trust scale (Cronbach’s  $\alpha = .90$ ).

## Results

A full-factorial analysis of variance (ANOVA) was used to test participants’ social responses to computer-synthesized voices, with the computer-voice personality and the participant personality as the between-subjects factors. In addition, a combined analysis was used to test the effects of the user choice in HCI.

Consistent with the previous finding of social identification, participants identified the personality cues in a computer voice even when they voluntarily chose the voice. Specifically, the extroverted computer voice was perceived as being more extroverted ( $M = 4.66$ ) than the introverted computer voice ( $M = 3.30$ ),  $F(1, 60) = 21.22$ ,  $p < .001$ ,  $\eta^2 = .26$ , regardless of the participant personality (See Table 1 for the mean of each experimental condition).

The results also showed the effect of proximate source orientation. The perceived personality of the voice influenced the perception of the reviewer’s personality, even when participants had chosen a computer voice that they would hear. Specifically, the reviewer was perceived as being more extroverted when the descriptions were narrated by the extroverted

computer voice ( $M = 5.11$ ) than by the introverted computer voice ( $M = 4.34$ ),  $F(1, 60) = 6.18$ ,  $p < .05$ ,  $\eta^2 = .09$ .

Contrary to the previous finding of the similarity attraction effect, we could not find any types of similarity attraction with the four dependent variables measured by the liking and credibility of the voice and the reviewer respectively. In order to investigate the non-significant finding of the similarity attraction effect further, we conducted an additional analysis with the data. A total of 64 participants were tested in a logistic regression analysis. The results showed that the order of voice presentation influenced participants' choices of the computer voice, whereas participants' own personalities did not influence their choices (see Table 2). The results suggest that a user is approximately 2.95 times as likely to choose the extroverted voice when they hear the extroverted voice first than when they hear the introverted voice first.

Finally, we conducted a combined full-factorial analysis of variance (ANOVA), with the choice (choice vs. no choice), the computer-voice personality (extrovert vs. introvert) and the participant personality (extrovert vs. introvert) as the between-subjects factors (see Table 3).

We found a main effect of the user choice on the attractiveness of the computer voice. Participants liked the computer voice more when they had a choice ( $M = 3.05$ ) than when they had no choice ( $M = 2.35$ ),  $F(1, 120) = 9.47$ ,  $p < .01$ ,  $\eta^2 = .07$ , after controlling for the personality of the participant and the computer voice. An interesting pattern of a three-way interaction effect was also found,  $F(1, 120) = 9.47$ ,  $p < .01$ ,  $\eta^2 = .07$ . The results of further analysis showed a significant, simple two-way interaction between the computer-voice personality and participant personality for liking of the voice in the no-choice condition,  $F(1, 120) = 7.84$ ,  $p < .01$ , but a non-significant, simple two-way interaction between the computer-voice personality and participant personality for liking of the voice in the choice condition  $F(1, 120) = 0.03$ , *n.s.* (see

Figure 1). The pattern of the significant cross-over interaction in the no-choice condition revealed that introverts preferred the introverted voice and extroverts preferred the extroverted voice.

Similarly, we found main effects of the user choice and participant personality on the credibility of the computer voice. Participants evaluated the credibility of the computer voice more positively when they had a choice ( $M = 5.90$ ) than when they had no choice ( $M = 5.08$ ),  $F(1, 120) = 5.04$ ,  $p < .05$ ,  $\eta^2 = .04$ , after controlling for the personality of the computer voice and the participant. In general, introverts evaluated the credibility of the computer voice more positively ( $M = 5.93$ ) than extroverts ( $M = 5.04$ ),  $F(1, 120) = 5.98$ ,  $p < .05$ ,  $\eta^2 = .05$ . However, there was no significant interaction effect among the three independent variables.

The same  $2 \times 2 \times 2$  ANOVA analysis from above was conducted on the attractiveness and credibility of the reviewer. The main effect of the user choice on liking of the reviewer was significant. In particular, participants liked the reviewer more when they had a choice ( $M = 4.89$ ) than when they had no choice ( $M = 4.23$ ),  $F(1, 120) = 3.80$ ,  $p = .05$ ,  $\eta^2 = .03$ , after controlling for the personality of the computer voice and the participant. The results also showed a significant three-way interaction effect,  $F(1, 120) = 7.92$ ,  $p < .01$ ,  $\eta^2 = .06$ . Similar to the previous results for the voice attractiveness, there was a significant, simple two-way interaction between the computer-voice personality and participant personality for liking of the reviewer in the no-choice condition,  $F(1, 120) = 27.44$ ,  $p < .01$ , but a non-significant, simple two-way interaction between the computer-voice personality and participant personality in the choice condition,  $F(1, 120) = 1.50$ , *n.s.* (see Figure 2). The pattern of the significant cross-over interaction in the no-choice condition indicated that introverts preferred the introvert reviewer and extroverts preferred the extrovert reviewer.

Finally, there was a significant main effect of the user choice on the credibility of the reviewer. Contrary to the previous results, participants evaluated the credibility of the reviewer more positively when they had no choice ( $M = 4.70$ ) than when they had a choice ( $M = 3.88$ ),  $F(1, 120) = 76.71, p < .001, \eta^2 = .39$ , after controlling for the other two independent variables. In addition, a further analysis of the significant three way interaction effect showed a significant, simple two-way interaction between the computer-voice personality and participant personality in the no-choice condition,  $F(1, 120) = 14.14, p < .01$ , and a non-significant, simple two-way interaction between computer-voice personality and participant personality in the choice condition  $F(1, 120) = 0.23, n.s.$  (see Figure 3). The pattern showed that introverts evaluated the credibility of the introvert reviewer more positively, whereas extroverts evaluated the credibility of the extrovert reviewer more positively.

### Discussion

This study provides strong evidence that people socially respond to computer-synthesized voices even after they choose a particular computer voice. It also replicates the previous findings of social identification of computer voices and the proximate-source orientation from the study by Nass & Lee (2001). The results effectively eliminate two alternative explanations to the previous proximate-source orientation findings.

First, it evidently eliminates the argument that social responses to computer voices are nothing more than social responses to programmers and designers behind the computers. All participants in this experiment clearly recognized that (1) the computer synthesized voices were generated by a machine without being related to the reviews or the reviewers; and (2) no one other than themselves is responsible for a specific type of voice that they have heard because of their voluntary choice. In the current study participants chose a type of computer voice (i.e.,

extrovert or introvert) narrating the same reviews after hearing both options. Therefore, there were no hidden designers or programmers that they would imagine and socially respond to.

Second, the results also eliminates the criticism that social responses to computer voices are natural in that people can assume a computer voice as a distorted version of pre-recorded human speech. This argument criticizes the proximate-source orientation finding in that people might assume that a computer voice is actually the distorted version of the genuine voice of the source. Or more simply, people might not tell the difference between a human voice and a computer voice. From this viewpoint, proximate-source orientation comes from people's lack of knowledge about the connection between a computer voice and a reviewer of the review. All participants in this experiment, however, clearly knew the disassociation between the computer voices and the review or the reviewer. In addition, the results evidently eliminate the usual criticism against the experimental method that results are from participants' intention to please the experimenter. Since participants in this experiment chose an experimental condition by themselves, they could not guess what would be the best way to please the experimenter.

We believe that the act of making choices nullifies the similarity-attraction effect. The results from the logistic regression analysis and a combined analysis provide the empirical evidence of (1) a strong primacy effect on user choice: participants selected whatever voice that they heard first; and (2) a cognitive dissonance effect: participants showed more positive social responses to whatever computer voice that they have chosen. First of all, the primacy effect posits that the sequence of presentation order influences people's preferences (Zajonc, 1980). More specifically, people prefer the first one to which they are exposed if there is no discrimination of quality among alternatives. For example, studies have demonstrated the primacy effect on people's memory on personality impressions (Luchins & Luchins, 1985) and



claimed receipt and leadership of trade publications (Whipple & McManamon, 1992). Thus, the primacy effect could explain why participants chose the voice that they heard first, regardless of the voice personality. This finding implies that industry practitioners may be able to provide customized services to a user (e.g., matching personalities between the voice interface and the user) by controlling for the presentation order of options even when the user can make a choice at the customer's own discretion. The synergetic combination of customized services and the free-choice interface may result in creating more positive experiences on the user side.

The major reason for the non-significant finding of the similarity attraction principle might be that once people choose a particular voice, people evaluate the chosen voice and the reviewer more positively, in order to minimize any potential cognitive dissonance. The results indicated the powerful effects of the user choice on their post-choice evaluation measured by liking of the computer voice, credibility of the computer voice, and liking of the reviewer. However, it is worth noting that people evaluate the credibility of the assigned reviewer more positively, compared to the credibility of the chosen reviewer. A possible explanation is that the criteria for the participants' choice might not be related to the dimensions of credibility. As a result, the participants might be unsure of the credibility of the reviewer although they simply liked the reviewer because it was their choice. In addition, credibility is closely related to expert and authority that are often presented to people in the top-down way. When the participants from the previous study were assigned to a certain reviewer, they might have attributed credibility to the reviewer because the participants did not have any control over the selection process, thus took the credibility granted for. On the other hand, when the participants from the current study chose a certain reviewer by themselves, they might have been uncertain about their own choice with regards to the reviewer credibility (e.g., I like my choice but am unsure of whether I have

selected a right one). Taken together, people like the voice and the reviewer regardless of their own personality and the personality of the voice when they make a choice. However, the powerful effects of the user choice may not be expanded to the reviewer's credibility when people make a voluntary choice without considering credibility as a major criterion for their choice. The potential reason for positive voice credibility in the choice condition is the positive relationship between liking and credibility (see Nass & Lee, 2001). It is interesting to see that increased liking of the computer voice due to the voluntary user choice could influence the liking of the reviewer and the credibility of the computer voice positively, but could not influence the credibility of the reviewer significantly (i.e., an attenuated proximate source orientation effect).

### Implications

Designers of TTS systems should be keenly aware that even a pathetic TTS voice can have a personality and can affect the perception of remote sources. Voice casting, thus, should be carefully determined before being implemented into the system. The finding that the similarity attraction effect is weakened by cognitive dissonance when users make a choice could be a double-edged sword for business marketers. It demonstrates that letting users choose from multiple options would be a wise decision in the interface design. Designers, however, lose the opportunity to utilize the similarity attraction effect when they give choices to users. One possible way to compromise these edges is to control for the presentational order of choices in the way to utilize customized services such as providing perceived similarities between the interface and the user.

The booming market for cell phone "ring tones" demonstrates the attractiveness of incorporating the choice paradigm in practice. Imagine, for example, that all cars were equipped with a voice interface. Should every car of a particular manufacturer speak the same voice or

should each driver select the unique voice of his or her own car? Similarly, would people enjoy and appreciate the interface more if they are able to choose the telephone voice that provides customer support? The key point is that choices are psychologically powerful as shown in this study. In general, because of cognitive dissonance, people will like the voices they select and will try to avoid information about the desirability of other alternative choices (Festinger, 1957). Although there is some evidence for buyer's remorse that people regret their choices soon after making them (Schwartz, 2004), people tend to change what they believe (i.e., dislike of their choices) because they usually cannot undo the behavior of the choice.

Taken together, there is no evidence that people are successful at selecting the voices they will like in the long run; however, the mere act of selection can be a positive. This result indicates a powerful impact of the mere act of user choice in HCI, which is in line with findings in a recent study in mobile advertising (Gao, Rau, & Salvendy, 2009). Gao and colleagues found that participants perceived mobile advertisements more interactive when more response options for mobile advertisements were provided, which, in turn, resulted in more positive attitudes toward the advertisements. In addition to the choice itself, the feeling that the user hears a voice that virtually no other users will hear may create a perception of a unique interpersonal bond between the user and the voice, especially when there is an interaction rather than just listening to the voice. Then, how should the choice be managed? The best approach is to have a new voice (i.e., a voice other than the voice the user is currently using) proposes the options for selecting a different voice. Because people are polite to voices even if the voices are computer generated, people might find it hard to "reject" a voice with which they have been working (see Nass et al., 1999).

Another critical point is that too many options can be overwhelming. In a classic series of studies performed in both field and laboratory settings, Iyengar and Lepper (2000) showed that people were more likely to purchase gourmet jams and to undertake an optional class essay assignment when only 6 choices were offered as compared to 24 or 30 choices. Furthermore, participants felt better about their choice and performed better when their selections were limited. Unfortunately, there are no clear guidelines or threshold as to exactly how many options of voices, or options of any product or service, are optimal. Assuming the voices are quite different, it is possible that the optimal number of options is approximately seven because seven is approximately the number of voices can be kept in short-term memory (Miller, 1956). If the voices can be categorized (e.g., by gender), users will likely exhibit a greater tolerance and desire for more options. Of course, all of the options should remain consistent with the brand.

Currently, the goal of cutting-edge voice interface designers is to allow users to create a voice for the interface: Better algorithms will soon make this a reality. Some people might even decide to have their own voice for the interface; research shows that hearing own voice in the voice interface is not disturbing (Nass et al., 1998). Other people may choose to have a voice of family members, relatives, friends, or famous celebrities. The danger of allowing this level of customization is that the product or service may become invisible because of the presence of the voice, thereby undermining branding.

As briefly discussed, the implications of this study are broad. We expect to see more and more of new interfaces that are designed to evoke more natural responses from users implicitly. The findings of social responses to computer-synthesized voices and powerful influence of user choice in this study should be wisely considered and incorporated in the HCI design, especially in the future design of voice interfaces.

## References

- Byrne, D., & Griffitt, W. (1969). Similarity and awareness of similarity of personality characteristic determinants of attraction. *Journal of Experimental Research in Personality*, 3, 179-186.
- Byrne, D., Griffitt, W., & Stefaniak, D. (1967). Attraction and similarity of personality characteristics. *Journal of Personality and Social Psychology*, 5, 82-90.
- Cialdini, R. (2001). *Influence: Science and practice* (4th ed.). Boston: Allyn & Bacon.
- Detenber, B., & Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. *Journal of Communication*, 46(3), 66-84.
- Dijksterhuis, A., Bos, M. W., Nordgren, L. F., & van Baaren, R. B. (2006). On making the right choice: The deliberation-without-attention effect. *Science*, 311, 1005-1007.
- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford Univ. Press.
- Flowerday, T., & Schraw, G. (2003). Effect of choice on cognitive and affective engagement. *Journal of Educational Research*, 96(4), 207-215.
- Gao, Q., Rau, P. P., & Salvendy, G. (2009). Perception of interactivity: Affects of four key variables in mobile advertising. *International Journal of Human-Computer Interaction*, 25(6), 479-505.
- Hannafin, R. D., & Sullivan, H. J. (1996). Preferences and learner control over amount of instruction. *Journal of Educational Psychology*, 88, 162-173.
- Isbister, K. & Nass, C. (2000). Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, 53, 251-267.

- Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, 79 (6), 995-1006.
- Kahneman, D. & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39, 341-350.
- Lee, K. M., & Jung, Y. (2005). Evolutionary nature of virtual experience. *Journal of Cultural and Evolutionary Psychology*, 3(2), 161-180.
- Lee, K. M., & Nass, C. (2004). The multiple source effect and synthesized speech: Doubly disembodied language as a conceptual framework. *Human Communication Research*, 30 (2), 182-207.
- Luchins, A. S., & Luchins, E. H. (1985). Influences of memorizing information on impressions of a person. *Journal of Psychology*, 119 (4), 369-383.
- McClure, S. M., Laibson, D. I., Lowenstein, G., & Cohen, J. D. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science*, 306, 503-507.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Nass, C. & Lee, K. M. (2001). Does Computer-Generated Speech Manifest Personality? Experimental Tests of recognition, Similarity-Attraction, and Consistency-Attraction. *Journal of Experimental Psychology: Applied*, 7 (3), 171-181.
- Nass, C. & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56 (1), 81-103.

- Nass, C., Moon, Y., & Carney, P. (1999). Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology*, 29 (5), 1093-1110.
- Oulasvirta, A., Hukkinen, J., & Schwartz, B. (2009). When more is less: the paradox of choice in search engine use, *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 516-523.
- Pianesi, F., Mana, N., Cappelletti, A., Lepri, B., & Zancanaro, M. (2008). Multimodal recognition of personality traits in social interactions. *Proceedings of the 10th International Conference on Multimodal Interfaces*, 53-60.
- Qiu, L., & Bendasat, I. (2005). Online consumer trust and live help interfaces: The effects of text-to-speech voice and three-dimensional avatars. *International Journal of Human-Computer Interaction*, 19(1), 75-94.
- Reeves, B., Detenber, B., & Steuer, J. (1993, May). *News televisions: The effects of big pictures and big sound on viewer responses to the screen*. Paper presented to the Information Systems Division of the International Communication Association, Washington, DC, U.S.A.
- Reeves, B., & Nass, C. (1996). *The media equation: How people respond to computers, television, and new media like real people and places*. Cambridge University Press.
- Schraw, G., Flowerday, T., & Reisetter, M. F. (1998). The role of choice in reader engagement. *Journal of Educational Psychology*, 90(4), 705-714.
- Schwartz, B. (2004). *The paradox of choice: Why more is less*. Hopewell, NJ: Ecco.

- Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. R. (2002). Maximizing versus satisfying: Happiness is a matter of choice. *Journal of Personality and Social Psychology*, 83(5), 1178-1197.
- Sundar, S., & Nass, C. (2001). Conceptualizing sources in online news. *Journal of Communication*, 51(1), 52-72.
- Van Boven, L., Gilovich, T. (2003). To do or to have? That is the question. *Journal of Personality and Social Psychology*, 85(6), 1193-1202.
- Wheless, L. R. & Grotz, J. (1977). The measurement of trust and its relationship to self-disclosure. *Human Communication Research*, 3 (3), 250-257.
- Whipple, T. W. & McManamon, M. K. (1992). Primacy order effects in the measurement of trade magazine receipt and readership. *Journal of Advertising Research*, 32 (5), 24-29.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37 (3), 395-412.
- Wimmer, R. D., & Dominick, J. R. (2000). *Mass Media Research: An introduction* (6<sup>th</sup> ed.). Belmont, CA: Wadsworth.
- Yang, Y. C., & Chin, W. K. (1997). Motivational analysis of the effects of type of instructional control on learning from computer-based instruction. *Journal of Educational Technology Systems*, 25, 25-35.
- Zajonc, R. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175.





Table 1

*ANOVA results when users choose voice interfaces*

Dependent variables	Means and standard deviations				F values and effect sizes		
	Introvert subject		Extrovert subject		Main effects		Interaction effects
	Introvert voice	Extrovert voice	Introvert voice	Extrovert voice	Participant (P) personality	Voice (V) personality	P x V
Voice extrovertedness	3.36 (0.72)	5.39 (1.91)	3.23 ( 0.74)	3.93 (0.94)	7.31** $\eta^2 = .11$	21.22** $\eta^2 = .26$	5.13* $\eta^2 = .08$
Liking of the voice	2.98 (0.57)	3.49 ( 1.59)	2.66 ( 1.73)	3.06 (1.53)	1.07 $\eta^2 = .02$	1.64 $\eta^2 = .03$	.03 $\eta^2 = .00$
Credibility of the voice	6.94 (1.34)	5.77 (1.88)	5.38 (2.48)	5.50 (2.60)	3.00 <sup>+</sup> $\eta^2 = .05$	1.00 $\eta^2 = .02$	1.38 $\eta^2 = .02$
Reviewer extrovertedness	4.14 (0.98)	5.64 (1.20)	4.53 (1.45)	4.57 (1.29)	1.22 $\eta^2 = .02$	6.18* $\eta^2 = .09$	5.55* $\eta^2 = .09$
Liking of the reviewer	4.92 (1.46)	5.52 (1.80)	4.83 (2.38)	4.27 (1.98)	1.91 $\eta^2 = .03$	0.00 $\eta^2 = .00$	1.46 $\eta^2 = .02$
Credibility of the reviewer	3.86 ( 0.19)	3.86 ( 0.26)	3.84 ( .16)	3.97 (0.22)	0.56 $\eta^2 = .01$	1.37 $\eta^2 = .02$	1.43 $\eta^2 = .02$

Note. <sup>+</sup>p<.10, two-tailed. \*p<.05, two-tailed. \*\*p<.01, two-tailed.

Table 2

*Logistic Regression: Probability of choosing extrovert voice*

Predictors	Logit Coefficients ( $\beta$ )	Exp ( $\beta$ )
Hearing extrovert voice first	1.08*	2.95
User personality	-0.21	0.89

*Note.* \* $p < .05$ , two-tailed. \*\* $p < .01$ , two-tailed.

Table 3

*ANOVA results from the combined analysis*

Dependent variables	Means and standard deviations				Means and standard deviations				F values and effect sizes			
	Without Choice				With Choice							
	Introvert subject		Extrovert subject		Introvert subject		Extrovert subject		Main effects		Interaction effects	
	Introvert voice	Extrovert voice	Introvert voice	Extrovert voice	Introvert voice	Extrovert voice	Introvert voice	Extrovert voice	Choice (C)	Participant (P) personality	Voice (V) personality	C x P x V
Liking of the voice	2.62 (1.19)	2.11 (0.88)	1.69 (1.00)	2.98 (1.34)	2.98 (0.57)	3.49 (1.59)	2.66 (1.73)	3.06 (1.53)	9.47** $\eta^2 = .07$	0.80 $\eta^2 = .01$	3.49 <sup>+</sup> $\eta^2 = .03$	4.43* $\eta^2 = .04$
Credibility of the voice	6.35 (1.54)	4.67 (2.39)	4.01 (1.62)	5.29 (2.24)	6.94 (1.34)	5.77 (1.88)	5.38 (2.48)	5.50 (2.60)	5.04* $\eta^2 = .04$	5.98* $\eta^2 = .05$	1.00 $\eta^2 = .01$	1.32 $\eta^2 = .01$
Liking of the reviewer	4.77 (2.02)	4.33 (2.18)	2.82 (1.32)	5.00 (1.85)	4.92 (1.46)	5.52 (1.80)	4.83 (2.38)	4.27 (1.98)	3.79* $\eta^2 = .03$	3.79* $\eta^2 = .03$	1.76 $\eta^2 = .01$	7.92** $\eta^2 = .06$
Credibility of the reviewer	4.94 (0.55)	4.52 (0.95)	4.39 (0.73)	4.97 (0.58)	3.86 (0.19)	3.86 (0.26)	3.84 (.16)	3.97 (0.22)	76.71** $\eta^2 = .39$	0.00 $\eta^2 = .00$	0.56 $\eta^2 = .01$	5.39* $\eta^2 = .04$

Note. <sup>+</sup>p<.10, two-tailed. \*p<.05, two-tailed. \*\*p<.01, two-tailed.

### Figure Captions

Figure 1. Liking of the voice: The simple interaction between subject personality and voice personality at the different levels of choice (no choice vs. choice)

Figure 2. Liking of the reviewer: The simple interaction between subject personality and voice personality at the different levels of choice (no choice vs. choice)

Figure 3. Credibility of the reviewer: The simple interaction between subject personality and voice personality at the different levels of choice (no choice vs. choice)

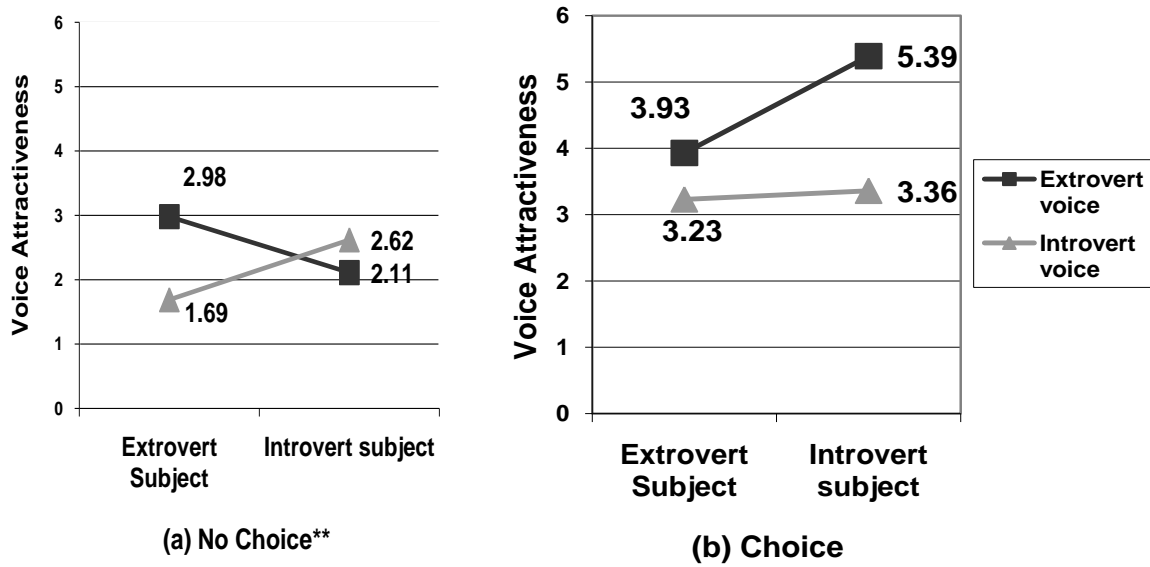


Figure 1. Liking of the voice: The simple interaction between subject personality and voice personality at the different levels of choice (no choice vs. choice).

Note. (a) Voice attractiveness as a function of subject personality and voice personality, controlled for the choice condition. (b) Voice attractiveness as a function of subject personality and voice personality, controlled for the no-choice condition.

\*\* $p < .01$ , two-tailed.

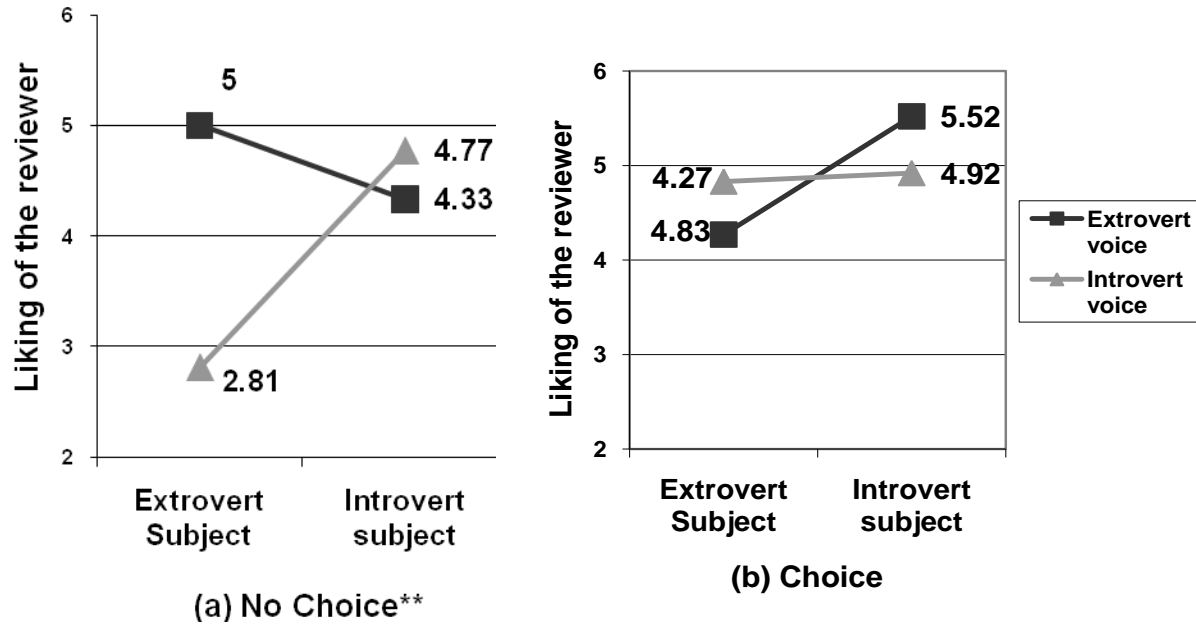


Figure 2. Liking of the reviewer: The simple interaction between subject personality and voice personality at the different levels of choice (no choice vs. choice).

*Note.* (a) Liking of the reviewer as a function of subject personality and voice personality, controlled for the choice condition. (b) Liking of the reviewer as a function of subject personality and voice personality, controlled for the no-choice condition.

\*\* $p < .01$ , two-tailed.

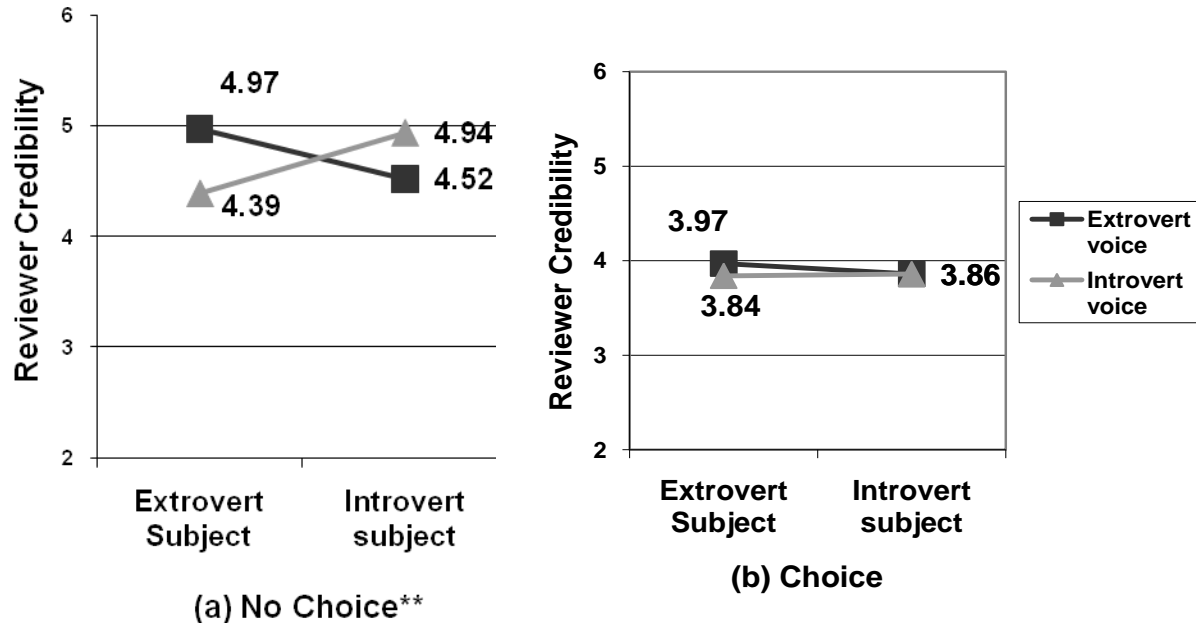


Figure 3. Credibility of the reviewer: The simple interaction between subject personality and voice personality at the different levels of choice (no choice vs. choice).

Note. (a) Credibility of the reviewer as a function of subject personality and voice personality, controlled for the choice condition. (b) Credibility of the reviewer as a function of subject personality and voice personality, controlled for the no-choice condition.

\*\* $p < .01$ , two-tailed.