

Preliminary Comparison of Immediate and Delayed Confirmation Strategies as a Function of Number of Corrections

TR 29.3663
June 30, 2003

James R. Lewis

IBM Pervasive Computing
Boca Raton, Florida

Abstract

Guidelines for speech user interfaces generally advise against the use of immediate confirmation of user input and promote the use of delayed confirmation. A usability test conducted with a prototype of a speech recognition interactive voice response (IVR) system for a cellular telephone company indicated a potentially serious usability problem with a simple delayed confirmation strategy. In the simple delayed confirmation strategy, callers had to review a fairly large amount of correct input multiple times when asked to make two corrections. This observation led to the recommendation for a better design for delayed confirmation.

ITIRC Keywords

Voice systems
Speech systems
Interactive Voice Response (IVR) systems
Speech user interface
Immediate confirmation
Delayed confirmation
Confirmation strategy

Contents

INTRODUCTION	1
METHOD.....	3
PARTICIPANTS.....	3
APPARATUS AND MATERIALS.....	3
PROCEDURE.....	3
EXPERIMENTAL DESIGN	3
RESULTS	5
ABSTRACTED TRANSCRIPTS OF TURNS REQUIRED TO COMPLETE TASKS.....	5
ASQ.....	5
TIME-ON-TASK	6
RECOMMENDATIONS	7
DISCUSSION.....	9
REFERENCES	11
APPENDIX A. AFTER SCENARIO QUESTIONNAIRE (ASQ).....	13
APPENDIX B. ABSTRACTED TRANSCRIPTS OF TURNS REQUIRED TO COMPLETE TASKS	15
TASK 1: PAYMENT BY ELECTRONIC CHECK, IMMEDIATE CONFIRMATION, NO CORRECTIONS.....	15
TASK 2: PAYMENT BY ELECTRONIC CHECK, DELAYED CONFIRMATION, NO CORRECTIONS.....	16
TASK 3: PAYMENT BY CREDIT CARD, IMMEDIATE CONFIRMATION, TWO CORRECTIONS.....	17
TASK 4: PAYMENT BY CREDIT CARD, DELAYED CONFIRMATION, TWO CORRECTIONS	18

Introduction

The purpose of this study was to investigate two confirmation strategies for data input with a speech recognition interactive voice response (IVR) system: immediate and delayed. Guidelines for speech user interfaces (Balentine & Morgan, 2002; IBM, 2001) generally advise against the use of immediate confirmation of user input and promote the use of delayed confirmation because delayed confirmation requires fewer dialog turns – as long as there is no need to correct errors.

The following sample dialogs for the collection of a date and time illustrate six confirmation situations: immediate confirmation with no corrections (4 turns, or system-human dialog pairs), immediate confirmation with one correction (6 turns), immediate confirmation with two corrections (8 turns), delayed confirmation with no corrections (3 turns), delayed confirmation with one correction (6 turns), and delayed confirmation with two corrections (9 turns).

Figure 1. Immediate Confirmation with No Errors (Sys = System; Hum = Human)

Sys: Date?
Hum: May 3rd.
Sys: Was that May 3rd?
Hum: Yes.
Sys: Time?
Hum: 2:00 PM.
Sys: Was that 2:00 PM?
Hum: Yes.

Figure 2. Immediate Confirmation with One Error (Sys = System; Hum = Human)

Sys: Date?
Hum: May 3rd.
Sys: Was that May 3rd?
Hum: Yes.
Sys: Time?
Hum: 2:00 PM.
Sys: Was that 2:00 AM?
Hum: No.
Sys: Time?
Hum: 2:00 PM.
Sys: Was that 2:00 PM?
Hum: Yes.

Figure 3. Immediate Confirmation with Two Errors (Sys = System; Hum = Human)

Sys: Date?
Hum: May 3rd.
Sys: Was that May 1st?
Hum: No.
Sys: Date?
Hum: May 3rd.
Sys: Was that May 3rd?
Hum: Yes.
Sys: Time?
Hum: 2:00 PM.
Sys: Was that 3:00 PM?
Hum: No.
Sys: Time?
Hum: 2:00 PM.
Sys: Was that 2:00 PM?
Hum: Yes.

Figure 3. Delayed Confirmation with No Errors (Sys = System; Hum = Human)

Sys: Date?
Hum: May 3rd.
Sys: Time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 3rd at 2:00 PM?
Hum: Yes.

Figure 5. Delayed Confirmation with One Error (Sys = System; Hum = Human)

Sys: Date?
Hum: May 3rd.
Sys: Time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 3rd at 2:00 AM?
Hum: No.
Sys: What needs to change? Select date or time.
Hum: Time.
Sys: New time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 3rd at 2:00 PM?
Hum: Yes.

Figure 6. Delayed Confirmation with Two Errors (Sys = System; Hum = Human)

Sys: Date?
Hum: May 3rd.
Sys: Time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 1st at 2:00 AM?
Hum: No.
Sys: What needs to change? Select date or time.
Hum: Date.
Sys: New date?
Hum: May 3rd.
Sys: Let's review. Was that May 3rd at 2:00 AM?
Hum: No.
Sys: What needs to change? Select date or time.
Hum: Time.
Sys: New time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 3rd at 2:00 PM?
Hum: Yes.

Note that the average number of turns for each strategy across the different numbers of corrections for this simple input example is the same at six turns (4, 6, and 8 turns for immediate; 3, 6, and 9 turns for delayed). The spread in the number of turns is greater for delayed than for immediate confirmation. These examples show that the advantage of delayed over immediate confirmation is evident only when there are no errors to correct, at least for this simple delayed confirmation strategy.

Method

Participants

Three IBM employees participated in this study. Two of the participants were male and one was female.

Apparatus and Materials

The participants interacted with a bill-paying prototype application running in the IBM¹ VoiceXML Toolkit (set for energy-based barge-in) on an IBM ThinkPad² A22m laptop computer, using an Andrea³ NC61 headset microphone for speech input. A Canon⁴ ZR 45 digital video camera captured the user/system interactions in the test sessions.

Participants completed four tasks (with tasks presented in the same order for all participants). After each task, participants completed an After Scenario Questionnaire (ASQ – Lewis, 1995).

The ASQ contains three questions, each scored from 1-7, with lower scores indicating greater satisfaction and ease of use (see Appendix A). The overall score is the arithmetic mean of the three item scores.

Procedure

As the experimenter presented tasks to a participant, the participant read the details for each task, asked for any necessary clarifications, and indicated to the experimenter when he or she was ready to begin the task. Participants completed the ASQ following each task. After the experimental session, the times on task were captured from the session videos. See Table 1 for a brief description of each task, and see Appendix B for abstracted transcripts for each task.

Table 1. Brief Task Descriptions

Task #	Task Name	Task Description
1	Immediate No Errors	Pay a bill with an electronic check; don't correct any errors.
2	Delayed No Errors	Pay a bill with an electronic check; don't correct any errors.
3	Immediate Two Errors	Pay a bill with a credit card; make two corrections.
4	Delayed Two Errors	Pay a bill with a credit card; make two corrections.

Experimental Design

The primary independent variable of interest in this study was the comparison of data for immediate and delayed confirmation. Due to the small scale of the study (only three participants), the independent variables of less interest (payment task and number of errors) were aliased, with the task of payment by electronic check bound to the condition of no errors and the task of payment by credit card bound to the condition of correcting two errors. This means that the only reasonable inferential statistics are *t*-tests comparing the results within each of the two combinations of payment task and number of errors.

¹ IBM is a registered trademark of International Business Machines Corp.

² ThinkPad is a registered trademark of International Business Machines Corp.

³ Andrea, Andrea Electronics and its products, logos and services are either trademarks or registered trademarks of Andrea Electronics Corporation or Lamar Signal Processing, Ltd.

⁴ Canon is a trademark of Canon Inc.

Results

Abstracted Transcripts of Turns Required to Complete Tasks

Appendix B contains abstracted transcripts of the turns required to complete the experimental tasks. Task 1 (payment by electronic check, immediate confirmation, no corrections) required 15 turns. Task 2 (payment by electronic check, delayed confirmation, no corrections) required 11 turns. Task 3 (payment by credit card, immediate confirmation, two corrections) required 24 turns. Task 4 (payment by credit card, delayed confirmation, two corrections) required 23 turns.

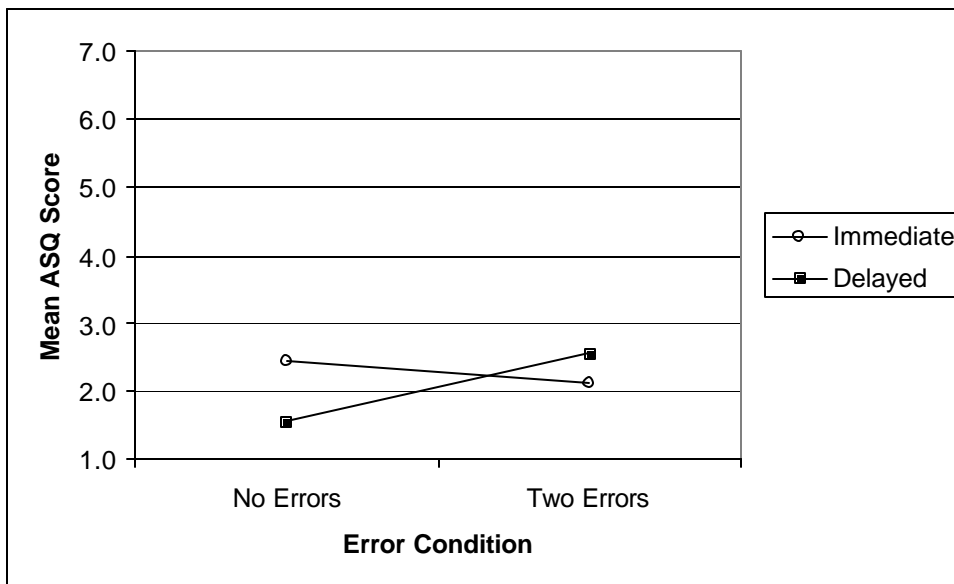
ASQ

Table 2 and Figure 5 illustrate the patterns of ASQ scores for the combinations of confirmation style and number of corrections. Neither of the planned ASQ comparisons indicated statistical significance (for check payment/no errors: $t(2) = 2.00, p = .18$; for credit card payment/two errors: $t(2) = 1.00, p = .42$).

Table 2. ASQ Scores as a Function of Confirmation Style and Number of Corrections

	No Errors	Two Errors
Immediate	2.4	2.1
Delayed	1.6	2.6

Figure 5. Mean ASQ Ratings for Each Task



Time-on-Task

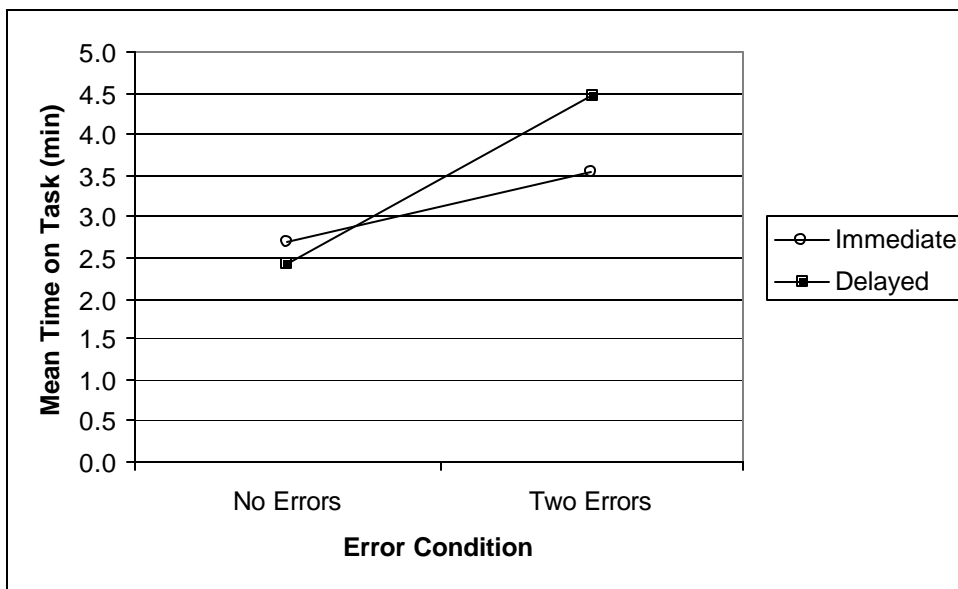
Table 3 and Figure 6 illustrate the patterns of times-on-task for the combinations of confirmation style and number of corrections. Due to the difference in the number of turns, the time required to complete payment with a credit card (including the two corrections) was expected to be longer than payment using an electronic check with no corrections. As mentioned previously, the only planned comparisons are within the two task/error combinations.

For payment by check with no corrections, the statistically significant difference in time ($t(2) = 16.00$, $p = .004$) was relatively small (about 15 seconds), but very consistent. The time difference for payment by credit card with two corrections was also statistically significant ($t(2) = 9.04$, $p = .01$). When there were no corrections, delayed confirmation was faster than immediate confirmation, but when there were two corrections, immediate confirmation was the faster strategy by about a minute.

Table 3. Time-on-Task (in minutes) as a Function of Confirmation Style and Number of Corrections

	No Errors	Two Errors
Immediate	2.7	3.5
Delayed	2.4	4.5

Figure 6. Mean Times-on-Task for Each Task



Recommendations

Recommendation 1: Until we develop a better delayed confirmation strategy, we should use immediate confirmation for turns that have a moderate to high likelihood of recognition error

For example, turns that require the input of alphanumeric data or lengthy numeric strings should use immediate confirmation.

Recommendation 2: We need to develop and test better designs for delayed confirmation.

For example, consider the following two delayed confirmation strategies:

Figure 7. Delayed Confirmation with Two Errors using Immediate Correction

Sys: Date?
Hum: May 3rd.
Sys: Time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 1st at 2:00 AM?
Hum: No.
Sys: What needs to change? Select date or time.
Hum: Date.
Sys: New date?
Hum: May 3rd.
Sys: Was that May 3rd?
Hum: Yes.
Sys: Anything else need to change?
Hum: Time.
Sys: New time?
Hum: 2:00 PM.
Sys: Was that 2:00 PM?
Hum: Yes.

Figure 8. Delayed Confirmation with Two Errors using Batch Collection and Correction of Changes

Sys: Date?
Hum: May 3rd.
Sys: Time?
Hum: 2:00 PM.
Sys: Let's review. Was that May 1st at 2:00 AM?
Hum: No.
Sys: What needs to change? Select date or time.
Hum: Date.
Sys: We'll make that change in a moment. Change anything else?
Hum: Yes, the time.
Sys: New date?
Hum: May 3rd.
Sys: Was that May 3rd?
Hum: Yes.
Sys: New time?
Hum: 2:00 PM.
Sys: Was that 2:00 PM?
Hum: Yes.

Discussion

Although the sample size for this study was small, the results indicated a potentially serious usability problem with a simple delayed confirmation strategy in the presence of two or more recognition errors. Because the use of immediate confirmation has its own well-known user acceptance issues (Balentine & Morgan, 2001), it is important to develop and test improved methods for delayed confirmation. This paper includes descriptions of two potentially better designs.

Note that both of these designs still require nine turns to complete in the given examples, but do not require users to listen repeatedly to the full confirmation message or to a complete menu of items to change for each correction, solving the major usability problem observed for the simple delayed confirmation strategy. Note also that the magnitude of the usability problem associated with repeatedly playing the full confirmation message increases as a function of the number of items in the confirmation message (see Appendix B). Before deploying applications using these strategies, though, it will be necessary to prototype and test them to understand their relative benefits and potentially unknown drawbacks.

For example, it isn't clear which approach will have the least demand on working memory during task completion. The Immediate Correction design (Figure 7) requires users to make the first correction before selecting the second item to change. Performing the first correction action could disrupt the maintenance of additional items to change in working memory. The Batch Collection design (Figure 8) lets users select all items to change before beginning the correction step, but does not conform to the correction steps that users seem to expect (to make a change immediately after identifying the element to change). This unexpected turn in the interaction could have a disruptive effect on the maintenance of additional items to change in working memory.

A future experiment should investigate these two confirmation strategies, using an experimental design that does not confound the important independent variables of confirmation strategy, order of use of strategy, and the pairing of task and strategy.

References

- Balentine, B., & Morgan, D. M. (2001). *How to build a speech recognition application: A style guide for telephony dialogs* (2nd ed.). San Ramon, CA: Enterprise Integration Group.
- International Business Machines, Corp. (2001). *IBM VoiceXML Programmer's Guide*. Author.
- Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction*, 7, 57-78.

Appendix A. After Scenario Questionnaire (ASQ)

TELL THE TEST MONITOR WHEN YOU FINISH THIS TASK. Then complete the following short questionnaire.

For each of the statements below, circle the rating of your choice.

1. Overall, I am satisfied with the ease of completing this task.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

2. Overall, I am satisfied with the amount of time it took to complete this task.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

3. Overall, I am satisfied with the support information (on-line help, messages, documentation) when completing this task.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

Appendix B. Abstracted Transcripts of Turns Required to Complete Tasks

Task 1: Payment by Electronic Check, Immediate Confirmation, No Corrections

Sys: *Welcome to Maxtel's automated pay-by-phone service. You can say Repeat or Help at any time. Say or enter your account number or ten-digit cellular phone number.*

Hum: 5 6 1 4 9 5 9 0 2 2

Sys: *Was that 5 6 1 4 9 5 9 0 2 2?*

Hum: Yes.

Sys: *Billing zip code?*

Hum: 3 3 4 4 3

Sys: *Was that 3 3 4 4 3?*

Hum: Yes.

Sys: *Your most recent payment was \$57.16 on May 27th. That payment was made in the past 72 hours. Your account is current, with a balance of \$67.34. Say Pay It All, Cancel Current Transaction, or state a specific amount to pay.*

Hum: Pay it all.

Sys: *Pay \$67.34?*

Hum: Yes.

Sys: *Pay with credit card or check?*

Hum: Check.

Sys: *To pay by electronic check, please have your check number, checking account number, and ABA routing number available. Say or enter your check number.*

Hum: 3 8 9 0

Sys: *Was that 3 8 9 0?*

Hum: Yes.

Sys: *The ABA routing number is a nine-digit number located at the bottom of your check. Say or enter the routing number.*

Hum: 1 1 1 2 2 3 4 5 6

Sys: *Was that 1 1 1 2 2 3 4 5 6?*

Hum: Yes.

Sys: *Say or enter your checking account number.*

Hum: 9 7 6 3 2 8

Sys: *Was that 9 7 6 3 2 8?*

Hum: Yes.

Sys: *Ready to make electronic check payment of \$67.34. Say Process Transaction or Cancel.*

Hum: Process transaction.

Sys: *Transaction complete. The confirmation number is 1 1 1 3 6 8. Did you know that you can set up a schedule for paying your bill automatically? Visit our web site for details. To continue, say Repeat, Pay Another Bill, or Exit.*

Hum: Exit.

Task 2: Payment by Electronic Check, Delayed Confirmation, No Corrections

Sys: *Welcome to Maxtel's automated pay-by-phone service. You can say Repeat or Help at any time. Say or enter your account number or ten-digit cellular phone number.*

Hum: 5 6 1 4 9 5 9 0 2 2

Sys: *Billing zip code?*

Hum: 3 3 4 4 3

Sys: *Your most recent payment was \$57.16 on May 27th. That payment was made in the past 72 hours. Your account is current, with a balance of \$67.34. Say Pay It All, Cancel Current Transaction, or state a specific amount to pay.*

Hum: *Pay it all.*

Sys: *Let's review. Pay \$67.34 to 5 6 1 4 9 5 9 0 2 2 with billing zip 3 3 4 4 3?*

Hum: *Yes.*

Sys: *Pay with credit card or check?*

Hum: *Check.*

Sys: *To pay by electronic check, please have your check number, checking account number, and ABA routing number available. Say or enter your check number.*

Hum: 3 8 9 0

Sys: *The ABA routing number is a nine-digit number located at the bottom of your check. Say or enter the routing number.*

Hum: 1 1 1 2 2 3 4 5 6

Sys: *Say or enter your checking account number.*

Hum: 9 7 6 3 2 8

Sys: *Let's review. Pay with check 3 8 9 0 using routing number 1 1 1 2 2 3 4 5 6, and checking account number 9 7 6 3 2 8?*

Hum: *Yes.*

Sys: *Ready to make payment. Say Process Transaction or Cancel.*

Hum: *Process transaction.*

Sys: *Transaction complete. The confirmation number is 1 1 1 3 6 8. Did you know that you can set up a schedule for paying your bill automatically? Visit our web site for details. To continue, say Repeat, Pay Another Bill, or Exit.*

Hum: *Exit.*

Task 3: Payment by Credit Card, Immediate Confirmation, Two Corrections

Sys: *Welcome to Maxtel's automated pay-by-phone service. You can say Repeat or Help at any time. Say or enter your account number or ten-digit cellular phone number.*

Hum: 5 6 1 4 9 5 9 0 2 2

Sys: *Was that 5 6 1 4 9 5 9 0 2 2?*

Hum: Yes.

Sys: *Billing zip code?*

Hum: 3 3 4 4 3

Sys: *Was that 3 3 4 4 3?*

Hum: Yes.

Sys: *Your most recent payment was \$57.16 on May 27th. That payment was made in the past 72 hours. Your account is current, with a balance of \$67.34. Say Pay It All, Cancel Current Transaction, or state a specific amount to pay.*

Hum: Pay it all.

Sys: *Pay \$67.34?*

Hum: Yes.

Sys: *Pay with credit card or check?*

Hum: Credit card.

Sys: *Which card?*

Hum: Visa.

Sys: *For security and accuracy, we will need your credit card's number, expiration date, billing zip code, and the security code located on the back of the card. Say or enter the credit card number.*

Hum: 3 3 6 0 5 3 9 9 0 7 1 1 7 8 7 6

Sys: *Was that 3 3 6 0 5 3 9 9 0 7 1 1 7 8 7 6?*

Hum: Yes.

Sys: *Expiration date?*

Hum: 0 7 0 3

Sys: *Was that 0 7 0 3?*

Hum: Yes.

Sys: *Credit card billing zip code?*

Hum: 3 3 3 4 9

Sys: *Was that 3 3 3 4 8?*

Hum: No.

Sys: *Credit card billing zip code?*

Hum: 3 3 3 4 9

Sys: *Was that 3 3 3 4 9?*

Hum: Yes.

Sys: *Check the back of your credit card. Do you have at least three digits to the right of the credit card number located on the signature panel?*

Hum: Yes.

Sys: *Are the last three digits legible?*

Hum: Yes.

Sys: *Say or enter the last three digits.*

Hum: 2 2 7.

Sys: *Was that 6 7 8?*

Hum: No.

Sys: *Say or enter the last three digits.*

Hum: 2 2 7.

Sys: *Was that 2 2 7?*

Hum: Yes.

Sys: *Ready to make Visa payment of \$67.34. Say Process Transaction or Cancel.*

Hum: Process transaction.

Sys: *Transaction complete. The confirmation number is 1 1 1 3 6 8. Did you know that you can set up a schedule for paying your bill automatically? Visit our web site for details. To continue, say Repeat, Pay Another Bill, or Exit.*

Hum: Exit.

Task 4: Payment by Credit Card, Delayed Confirmation, Two Corrections

Sys: *Welcome to Maxtel's automated pay-by-phone service. You can say Repeat or Help at any time. Say or enter your account number or ten-digit cellular phone number.*

Hum: 5 6 1 4 9 5 9 0 2 2

Sys: *Billing zip code?*

Hum: 3 3 4 4 3

Sys: *Your most recent payment was \$57.16 on May 27th. That payment was made in the past 72 hours. Your account is current, with a balance of \$67.34. Say Pay It All, Cancel Current Transaction, or state a specific amount to pay.*

Hum: Pay it all.

Sys: *Let's review. Pay \$67.34 to 5 6 1 4 9 5 9 0 2 2, with billing zip 3 3 4 4 3?*

Hum: Yes.

Sys: *Pay with credit card or check?*

Hum: Credit card.

Sys: *Which card?*

Hum: Visa.

Sys: *For security and accuracy, we will need your credit card's number, expiration date, billing zip code, and the security code located on the back of the card. Say or enter the credit card number.*

Hum: 3 3 6 0 5 3 9 9 0 7 1 1 7 8 7 6

Sys: *Expiration date?*

Hum: 0 7 0 3

Sys: *Credit card billing zip code?*

Hum: 3 3 3 4 9

Sys: *Check the back of your credit card. Do you have at least three digits to the right of the credit card number located on the signature panel?*

Hum: Yes.

Sys: *Are the last three digits legible?*

Hum: Yes.

Sys: *Say or enter the last three digits.*

Hum: 2 2 7.

Sys: *Let's review. Charge to Visa account number 3 3 6 0 5 3 9 9 0 7 1 1 7 8 7 6, expiring 07 03, with security code 6 7 8 and billing zip 3 3 3 4 8?*

Hum: No.

Sys: *Say Change Card Type, Change Card Account Number, Change Expiration Date, Change Security Code, or Change Card Billing Zip.*

Hum: Change Card Billing Zip.

Sys: *Credit card billing zip code?*

Hum: 3 3 3 4 9

Sys: *Let's review. Charge to Visa account number 3 3 6 0 5 3 9 9 0 7 1 1 7 8 7 6, expiring 07 03, with security code 6 7 8 and billing zip 3 3 3 4 9?*

Hum: No.

Sys: *Say Change Card Type, Change Card Account Number, Change Expiration Date, Change Security Code, or Change Card Billing Zip.*

Hum: Change Security Code.

Sys: *Check the back of your credit card. Do you have at least three digits to the right of the credit card number located on the signature panel?*

Hum: Yes.

Sys: *Are the last three digits legible?*

Hum: Yes.

Sys: *Say or enter the last three digits.*

Hum: 2 2 7.

Sys: *Let's review. Charge to Visa account number 3 3 6 0 5 3 9 9 0 7 1 1 7 8 7 6, expiring 07 03, with security code 2 2 7 and billing zip 3 3 3 4 9?*

Hum: Yes.

Sys: *Ready to make payment. Say Process Transaction or Cancel.*

Hum: Process transaction.

Sys: Transaction complete. The confirmation number is 1 1 1 3 6 8. Did you know that you can set up a schedule for paying your bill automatically? Visit our web site for details. To continue, say Repeat, Pay Another Bill, or Exit.

Hum: Exit.