

Moving on from IVR

Lynn Berry and Dominique Estival

Syrinx Speech Systems

lberry@syrinx.com.au, estival@syrinx.com.au

1. Introduction

In commercial spoken dialogue applications, consideration must be given to the client commissioning the application and the actual users of the application. Typically, the dialogue designer must strike a balance between the requirements of the technology and the requirements of the organisation putting in place the application, and create a directed speech interaction which is acceptable to regular clients.¹ The dialogue designer is confronted not only with the task of writing the application dialogue, but also with particular views customers may already have regarding the application and how it should be structured. These views tend to be influenced by the many IVR (Interactive Voice Response) touch-tone systems already used by numerous organisations. In this paper, leaving aside the larger issues concerning dialogue design in general, we address the specific issue of fielding a commercial dialogue application at a customer site. We report on our experience in developing a banking and stock market telephone-based transaction system for a large financial institution, and we discuss the issue of dealing with preconceptions when transferring from one technology to another, in particular from a touch-tone to a voice interface [1].

2. Dialogue Analysis

In the course of developing the spoken dialogue application for Commonwealth Securities,² we conducted an analysis of a corpus of recorded human-human interactions between the clients and the agents in the stock trading service of Commonwealth Securities. The service offered by the bank allows registered clients to ring and request stock prices, trade stocks, confirm and cancel trade orders. Stock trading involves negotiation about stock prices whether for buying or selling.

Recording these calls is a legal requirement for the bank, and we randomly selected for analysis one day's worth of calls (approximately 6 hrs, with a duration of about 5 min per call). The corpus was analysed for dialogue flow,

vocabulary, syntactic structure, and pragmatic information (for instance whether any information is taken for granted in the context). The information obtained from the analysis was then used to build the lexicon, and to construct the dialogue and the grammars for the automated telephone system to be implemented for that service. Two main characteristics of human-human dialogue became apparent.

The first is that these dialogues are domain specific, and the interaction between the user and the agent is structured. Users know what to expect, i.e. what information is required and when. The agent is trained to obtain information quickly and effectively, and the user is 'trained' by this process.³

The second characteristic is that there is close cooperation between the interlocutors, involving constant checking between the participants. This occurs for example when using stock codes containing confusable letters, e.g. "BHP" and "BHV", and when checking the price in dollars, or the amount of shares to be traded.

3. Dialogue Design

The aim of the dialogue designer is to create a call flow corresponding both to clients' expectations and customers' requirements.

In general, most people are more familiar with IVR employing DTMF (i.e. touch tone) than with speech recognition systems. This familiarity influences how customers expect their application should be constructed. We have found that one of the major concerns of customers is how their clients will respond to a spoken language system, and a new interface [2,3].

In adjusting to a new medium, e.g. speech-driven applications, people understandably tend to compare it with a more familiar system. In this case, a comparison may be made with IVR systems, where users hear prompts such as (1).

(1) press 1 for account balance, press 2 for transfers. In their concern about their clients' reaction to a spoken dialogue system and a new interface (i.e. speech as opposed to DTMF), customers tend to conceive menu-styled dialogue

¹"Customer" refers to the organisation which commissions the application (e.g. a bank) while "clients" refer to the end users of that application (the bank customers).

²We appreciate the assistance given to us by the staff at Commonwealth Securities, and in particular by John O'Donnell.

³This structure and the order it imposes on the flow of the dialogue make this type of dialogue very suitable for a system where information can be organised according to the different tasks involved, and where a form need not be completed for each task.

scripts. For instance, in a spoken dialogue system, (1) may be interpreted as a prompt such as (2):

(2) for account balance say account balance, for transfers say transfer.

For applications which may require a complex menu hierarchy, a scenario such as the one in (2) would not be effective, or acceptable. One of the issues is thus how to create prompts to allow for a more naturally flowing dialogue [4,6,7], when the system specifications are given in the form of an IVR call flow diagram.

When designing a new spoken dialogue for an existing application with a complicated IVR menu structure, another issue is to convey to novice users which services are available. This is one of the primary concerns of the customer and one of their system requirements. However, the customer wants to include as much information as possible in the prompts, and a challenge for dialogue designers is to allow for information to be communicated in a comprehensive but brief manner. Too much information in all the prompts slows down the interaction; it puts frequent users, who are familiar with the application and do not need this information, at a disadvantage. Frequent users confronted with a verbose and inflexible dialogue may become irritated and might prefer using an alternative mode of interaction (e.g. web, or human).

Prior to touch tone, telephonists were the voice of the organisation; they almost never gave a list of the services available, but instead guided callers to their desired destinations. We want to emulate the role of the traditional telephonist, without encouraging lengthy negotiations over a user request.

4. Implementation

The spoken dialogue application for Commonwealth Securities was introduced in two stages. This enabled the customer to trial the system and receive feedback from users. A limited directed speech system, which handled calls for stock prices, was first introduced to selected bank clients. The dialogue consisted of basic steps, such as providing context sensitive help, confirming information at certain points in the dialogue, and providing an exit to a human agent for additional services or in case of recognition or understanding difficulties. In addition to these features, the second system allows freer input and the dialogue handles trading on the stock market, i.e. buying and selling shares. This system also prompts for missing information and resolves conflicts.

In both systems, prompts cover information to navigate the system and they are designed to be as specific as possible. This is one way to resolve misunderstandings between system and users [5,8,9].

5. Conclusion

The implementation of a commercial system at a customer site has highlighted a number of issues in dialogue design. Some of these were unexpected and are due to prior

familiarity with IVR touch tone systems, giving rise to assumptions about the way speech recognition systems operate. Developing automated telephone spoken applications for customers who are familiar with human-human interactions and IVR rather than with spoken language processing presents particular challenges to dialogue designers. As a consequence, we now have a better understanding of customer expectations, and this has allowed a more constructive interaction with the customer during dialogue design.

This experience with our customer is a particular instance of a more general issue, that is how people deal with changes in technology. For spoken language technology, we have to be aware that it can be compared with another one. In addition, customers may not share dialogue designers' familiarity with spoken dialogue systems. Thus, dialogue designers must be prepared to tackle customer pre-conceptions and to manage customer requirements. In particular, if these requirements come in the form of an IVR dialogue, this will need to be carefully restructured into a dialogue more appropriate for the technology.

6. References

- [1] Balentine, Bruce (1999) How to Build a Speech Recognition Application, EIG, San Ramon.
- [2] Billi, Roberto, Giuseppe Castagneri and Morena Danieli (1997), 'Field Trial evaluations of two different information inquiry systems', *Speech Communication* 23:83-93.
- [3] Davies, K et al (1999) The IBM Conversational Telephone System for Financial Applications, Esca.
- [4] Gorin, AL, GRiccardi, JH Wright (1997) 'How may I help you?', *Speech Communication* 23: 113-127.
- [5] Hambleton, Myra (Feb/Mar 2000) 'Directing the Dialogue: the art of IVR', *Speech Technology On Line* www.speechtechmag.com/st23/hambleton.htm.
- [6] Larsen, Lars Bo (1995) 'Development and Evaluation of a Spoken Dialogue for a Telephone Based Transaction System', in proceedings Eurospeech '95 vol 3, p:1973-6.
- [7] Rudnicky, A.I. et al, (1999) 'Creating Natural Dialogs in the Carnegie Mellon Communicator System' in proceedings Eurospeech '99.
- [8] Yankelovich, Nicole (1996) 'How Do Users Know What to Say?' *ACM Interactions* Vol. 3 No 6 November/December, Sun Microsystems.
- [9] Yankelovich, Nicole, Gina-Anne Levow and Matt Marx, 'Designing Speech Acts: Issues in Speech User Interfaces', (1995) Proceedings of CHI '95, Conference on Human Factors in Computing Systems, Denver, CO, May 7-11, 1995.