ELSEVIER

System-initiated digressive proposals in automated human–computer telephone dialogues: the use of contrasting politeness strategies

J. Wilkie[a,*], M.A. Jack[a], P.J. Littlewood[b]

[a]*Centre for Communication Interface Research, School of Engineering and Electronics, The University of Edinburgh, Edinburgh EH9 3JL, UK*
[b]*Lloyds TSB Bank Plc, Canons House, Canons Way, Bristol BS99 7LB, UK*

Abstract

System-initiated digressive proposals may be used to introduce new and unexpected information into automated telephone services. These digressions may be viewed as particularly pronounced forms of unsolicited interruptions as they contain information not directly related to the caller's intended activity. In human–human conversation, interruptions are considered to be speech acts which intrinsically threaten both the positive and negative face wants of the addressee and conversants adopt specific verbal strategies to mitigate the negative impact of their interruptions. A question therefore arises whether the introduction of face-redressive expressions, based on human–human conversational strategies, into the design of system-initiated proposals in automated services can mitigate the negative impact of the interruptions. A usability experiment was conducted to examine the effectiveness of three contrasting politeness strategies for system-initiated digressions in a mass-market telephone banking dialogue using speech recognition technology. Participants ($N = 111$) experienced these proposals while using the automated service to perform banking tasks. Results indicated

---

*Corresponding author. Tel.: +44-131-6502801; fax: +44-131-6502784.
E-mail address: jenny.wilkie@ccir.ed.ac.uk (J. Wilkie).

that all these system-initiated digressions—irrespective of politeness strategy employed—had a negative impact on the user attitudes towards the service. This paper reports these results and explores participants' perceptions of the politeness styles and registers employed in the system-initiated proposals.

1. Introduction

Speech recognition technology is increasingly used in the mass-market domain of self-service telephone applications. Compared to their push-button counterparts, applications which use spoken language input offer users a more natural and flexible way of interacting with a computer-based system. However, the system messages and the turn-taking in these speech operated applications often still resemble those found in push-button operated services in that they follow a rigid prompt-response sequence where the input options are presented to users in the form of vocal menus and explicit instructions about what to say. The dialogue between the human user and the automated service in such applications typically follows a pre-defined script involving a fixed turn-taking structure (the computer prompts then the user responds) and valid user responses are restricted by the capabilities of the speech recognition grammar. Users of these mass-market applications can expect a controlled and predictable interaction with the computer in a dialogue that does not change between phone calls.

Mass-market automated services are primarily designed to handle task-driven conversations within a narrow topic domain, such as flight information, banking account transactions or cinema bookings. The user of such services typically expects the interaction to be restricted to the chosen topic and task at hand and that the computer will cooperate fully to complete the goal of the call. Fixed turn-taking, goal-driven, prompt-response interaction has become the conventional way of designing automated self-service telephone applications. It is not common practice for an automated service to initiate an interruption or launch into new topics, a fact which may explain why such dialogue behaviour remains largely unexplored in the current literature for human–computer spoken interaction. The possibility of deploying system-initiated digressions in human–computer conversation raises new and interesting dialogue engineering issues regarding the design, usability and acceptability of such applications.

The research described in this paper explores how system-initiated digressive proposals may be used to disseminate unsolicited financial information in a speech-driven automated telephone banking service. These proposals work by interrupting the user and suspending the regular dialogue turn-taking for the duration of the informational message. The key issue examined in this research is how politeness strategies (considered an important factor in the choice of vocabulary in

human–human dialogue interruptions) may be employed to influence the impact of such system-initiated digressive proposals on user attitudes.

2. Motivation

Whilst the capability of speech recognition technology is continuously improving, the need still exists for the system messages (or prompts) to be designed to control for the range of user inputs that can be accepted by the computer (Bernsen et al., 1997). Much of the dialogue design for mass-market automated telephone services is centered around making the interaction fit the technology at hand, relying on explicit instructions and error recovery strategies in order to guide users as to what to say, how to say it and when to speak. In mass-market applications, menus in the form of explicit list selections are usually employed as a method for informing users (especially novice users) about the range of services available to them. Once such automated services have been designed, implemented and launched, changes to the dialogue turn-taking, menus or prompts can be costly, can result in customer objections and are rarely made. As a result, less frequently demanded information and transactional services are usually excluded from the dialogue and there is no straight-forward way of adding or deleting service options from menus once the application is up and running.

There are a number of reasons for looking beyond conventional menu-based dialogue design to explore alternative and more flexible means of offering users access to services through an automated application. For example, an enterprise may want to introduce new informational or transactional services that may not be considered in the initial application design under normal circumstances, such as access to services which are infrequently requested, short-term offers or product promotions, but at the same time avoid adding these as options to menus which may become unnecessarily long and complex. Furthermore, successful take-up of an automated service may result in the enterprise losing opportunities to advise customers about relevant products or services. This may involve the use of a "logical link", such as when a specific transaction on a customer's account is used to trigger advice on a relevant service or product.

One solution for adding new options in the service is to introduce a short system-initiated informational message within the dialogue structure with the intention of disseminating new information relevant to a particular customer at a specific point in time during their use of the automated service. This system-initiated message could simply consist of a brief prompt which may or may not be followed by a short dialogue (e.g. requesting a yes/no response) enabling the user to pursue or decline the offer immediately. The system-initiated message interrupts the regular turn-taking of the dialogue and, in doing so, impedes the human user from continuing with the flow of the call as anticipated. These messages may therefore be viewed as a particularly pronounced form of system-initiated digression since they are in effect unsolicited, unexpected and not directly related to the current topic or the prime goal of the call.

Successful strategies in this area could have important positive commercial implications.

3. Background

Dialogue engineering for speech-driven mass-market applications is mainly concerned with development issues relating to the technology at hand, such as whether to use voice recordings or text-to-speech for system prompts; whether to allow callers to barge-in during system prompts; whether to use open or closed prompt styles (Hone and Baber, 1999); whether to use isolated word recognition or allow for more fluent speech; and whether to allow universal commands such as "cancel" or "exit". The design principles for automated dialogues and research into voice interactive services described in the recent literature (Balentine and Morgan, 2001; Gardner-Bonneu, 2001) offer broad coverage of design aspects. However, they offer little in terms of guidelines on how to implement system-initiated interruptions and digressions in speech-driven applications: the area addressed in this paper. Studies of digression (often referred to as "out-of-turn interaction" or "unsolicited reporting" (Allen et al., 1999)) that may be found in human–computer spoken interaction research have focused mainly on providing models for handling *user*-initiated digressions (Haller, 1994; Narayanan et al., 2000; Ramakrishnan et al., 2002) which occur when the user supplies extra or out-of-turn information in response to system prompts. This new or extra information supplied by the user is however normally related to the overall goal of their participation in the conversation.

Previous research by the authors (Wilkie et al., 2002) identified two key dialogue engineering issues for the design of system-initiated digressions in human–computer dialogues: the *location* of the proposal in the application call-flow; and the dialogue *turn-taking strategy* employed for delivering the proposal information. In order to address these design issues, two experiments were devised in which system-initiated proposals were introduced in the call-flow of a speech-driven automated telephone banking service (Wilkie et al., 2002). These proposals informed users that they could apply for an overdraft facility on their account through the automated service by saying "overdraft" at the menu of services (the main menu). The wording employed in the proposal was low-key, short and terse in order to be consistent with the register employed in the rest of the automated service. In order to soften the impact of the interruption, phrases such as "*You might like to know that...*" were used in the opening statement of the proposal.

In order to assess the impact of the digressions in that research, participants' attitudes towards the telephone banking service were measured before and after they were subjected to this additional overdraft message. Participants were not forewarned about the pending interruption, nor did they receive any experiment priming to create a potential interest in the overdraft product. Results from that research revealed that participants' attitudes towards the usability of the service remained unaffected by the delivery of the overdraft proposal during their phone call. Additionally, measurements of participants' attitudes towards the digressive

dialogue itself suggested that there were no overall strong indications that one particular location or turn-taking strategy was more favoured than the other. These findings suggest that automated telephone banking dialogues can be successfully augmented using system-initiated digressive proposals. The results from that research serve as a point of departure for this current investigation into other aspects of system-initiated digressions in automated dialogues, where the role of politeness strategies and the effect of the stylistic manner employed when interrupting are investigated.

4. Politeness in human–human interaction

Politeness in human communication has received much attention in the field of pragmatics and sociolinguistics over the past two decades and has mainly been focused on how communicative strategies are employed in order to promote and maintain social harmony[1] in human–human interaction. One of the most influential theories of politeness is that developed by Brown and Levinson (1987). Their politeness theory is based on the notion that each individual has positive and negative "face wants" and that these are ascribed by all (rational) interactants to themselves and to one another in any social interactive situation. Brown and Levinson define the two face wants as (1987, p. 61):

Negative Face: the desire to be un-impeded in one's actions, the basic claim to territories, personal preserves, rights to non-distraction—i.e. freedom of action and freedom from imposition.
Positive Face: the desire (in some respects) to be approved of, the positive consistent self-image or "personality" claimed by interactants (crucially including the desire that this self-image be appreciated and approved of).

Any utterance or action in a communicative situation can be seen as potentially threatening to the positive or negative face of either of the interactants and, consequently, expressions of politeness are normally used as mitigations aimed at redressing this threat. Although Brown and Levinson give examples of how the speaker's *own* positive and negative face wants may be at risk,[2] much of their politeness theory is primarily focused on the explicit strategies used by a speaker to avoid damaging the addressee's face wants (Chen, 2001). The speaker uses politeness expressions to indicate that no face threat is intended or desired and to convey that the addressee's face wants are recognized and approved of by the speaker.

Brown and Levinson calculate the relative "seriousness" of a face-threatening act based on three "social dimensions" (1987, p. 74). These are: the relative power of the addressee over the speaker; the social distance between the speaker and the

---

[1]For an account of impoliteness strategies see Culpeper (1996).

[2]For example, expressing thanks or making an excuse are damaging to the speaker's negative face. Admissions of guilt or non-control of emotions (laughter or tears) are examples of damage to a speaker's positive face.

addressee; and the ranking of the imposition involved in doing the face-threatening act. Brown and Levinson point out that each of these dimensions is context-sensitive, meaning that the relationship between two individuals (such as the relative power of a manager over an employee) may be inverted under certain circumstances. Depending on the seriousness and social setting for the face-threatening act, a number of options are presented to the speaker on how to redress a potential face threat. First of all, the speaker has the option of not performing the act at all and could therefore theoretically avoid damaging the face of the addressee altogether. However, if the speaker decides to go ahead with the face-threatening act, Brown and Levinson identify a taxonomy of politeness which includes four principal categories of expression strategies: (1) doing the act without redressive action (baldly), (2) using positive face-redress, (3) using negative face-redress, or (4) doing the act off-record. The ''off-record'' strategy attempts to minimize the face threat by creating uncertainty as to the existence of the face-threatening act itself, e.g. by using ambiguous or vague expressions, or by using hints such as ''it's cold in here'' (implying ''shut the window'').

To carry out an act ''baldly'', without redress, involves doing it in the most direct, clear, unambiguous and concise way possible (Brown and Levinson, 1987, p. 69). The speaker may use the bald strategy when there is no fear of retribution by the addressee (e.g. in the interest of urgency or efficiency, e.g. ''watch out!''); where the danger to the addressee's face is very small (such as in proposals and requests); or where the speaker is considerably superior in power to the addressee.

Face-redressive politeness strategies are used when there is a perceived potential threat in an utterance to either the positive or negative (or both) face wants of the addressee. Utterances that are considered threatening to the *negative* face wants of the addressee will include: ordering the addressee to do something, making an offer which may incur debt for the addressee and expressions of strong emotions towards the addressee. Negative face-redressive strategies are characterized by formality and distancing. It is such forms of ''negative politeness'' that are conventionally associated with politeness in everyday language, such as ''excuse me'' and ''thank you'', as these relate to the imposition itself. *Positive* face-redress, on the other hand, widens the sphere of politeness to include the appreciation of the addressee's wants in general or to the expression of similarity between speaker's and addressee's wants. Threats to the addressee's positive face wants are caused by, e.g. bringing bad news about the addressee, expressing disapproval or raising emotionally divisive topics. The positive face-redress strategy is characterized by ''intimate'' language behaviour and makes reference to a close interdependent social relationship between the interactants. For example, the speaker might use in-group identity markers (hey buddy) or show intensified interest in the addressee's wants (your hair looks *great*).

Some face-threatening acts, such as interruptions,[3] are considered to be intrinsically threatening to *both* the negative and positive face wants of the addressee

---

[3]Other face-threatening acts considered to intrinsically threaten both the negative and positive face wants of the addressee are complaints, threats, strong expressions of emotions and requests for personal information.

(1987, p. 67). By Brown and Levinson's definition, an interruption constitutes a threat to the negative face wants of the addressee because it infringes to some degree on the addressee's right to non-distraction and desire to be un-impeded in their actions. Interruptions also pose a threat to the addressee's positive face wants by implying that the person who interrupts ignores or does not care about the addressee's feelings and wants.

5. Relevance to human–computer interaction

Politeness is undoubtedly an important aspect of *human–human* conversation, but little prior work has been undertaken to investigate how relevant it is to *human–computer* dialogues. What are the conversational rules or social dimensions that govern the use of politeness registers in dialogues where one of the interactants is a computer? Can existing politeness theories be expanded to encompass human–computer interaction? If so, what politeness strategies should the computer (in the capacity of the speaker) be endowed with and how are the resulting politeness expressions received by the human user?

People's interactions with computers (and other media) are fundamentally social (Nass et al., 1994; Reeves and Nass, 1996; Nass and Moon, 2000). This view is founded on the notion that the human brain has evolved to respond and relate socially to human-like entities in our surroundings and that this innate reaction is almost impossible to overcome—even in situations where humans interact with a supposedly non-social entity such as the computer. This propensity for humans to relate socially to media has been explored in a series of controlled experiments (Reeves and Nass, 1996; Nass and Moon, 2000). The results showed that users applied gender stereotypes to computers; they identified with computer agents sharing their ethnicity; and they were more attracted to agent characteristics (submissive/dominant) that were similar to their own personality. The authors also concluded that users apply "over-learned" social rules to computers, such as politeness: experimental results showed that participants gave a significantly more positive evaluation of a computer's performance when questioned directly by the computer itself compared to when questioned by a different computer or through pen and paper questionnaires. This would indicate that politeness is an important factor in human–computer interaction. However, the work on politeness in human–computer interaction carried out by Nass and colleagues has been centered around how humans behave politely towards computers, rather than investigating how humans respond to a computer that tries to portray polite behaviour.

The experimental results obtained by Nass and colleagues strongly suggest that human users have a subconscious tendency to apply deeply rooted social rules to interactions with computers in the same way as they do when interacting with other fellow humans. These social rules seem to relate to our innate disposition and cultural upbringing. But how do users react to a computer that blatantly attempts to exploit these social rules? Fogg and Nass (1997) explored the effects of employing computer-initiated flattery when giving feedback to users in a text-based guessing

game application. Experimental results showed that flattering feedback (compared to the generic feedback condition) had a positive effect on a number of aspects of the interaction. For example, the flattery increased participants' feelings of power; made them more positive towards their own and the computer's performance; and made them enjoy the interaction more.

Colón et al. (2001) studied the use of politeness in interruptions in a graphical library search engine interface. These interruptions involved on-screen error text messages (resulting from either system errors or user errors) that were presented with or without politeness (courtesy). The messages were deployed in the library application and evaluated in a controlled experiment. The two main findings from the experiment were: firstly, the interruption performed by the computer interface had a detrimental effect upon the user perception of the interaction with the computer (the participants judged the interaction as being less friendly, less motivating and less beneficial). Secondly, it was found that politeness strategies had no effect on minimizing participants' negative reaction towards the interruption.

The idea of treating the computer as a social entity and endowing it with emotive qualities such as politeness may be considered to be controversial given the fact that the computer does not have any real understanding about the effect its behaviour may have on its dialogue partners. Some user interface designers are opposed to the idea of anthropomorphizing computers and stress that users should be discouraged from thinking that computers may have human-like abilities (Shneiderman, 1988, 1993, 1998). This position derives from the point of view that human relationships are rarely a good model for designing effective human–computer user interfaces and that the primary goal for interface design should be predictable and controllable interaction (Shneiderman, 2000). McFarlane (1998), in his work on interruptions of the visual display in human–computer interaction, concludes simply that politeness is an irrelevant topic for the design of user interfaces as computers do not have "face" and people do not have face-wants relative to their computers. MacFarlane therefore suggests that the "bald" strategy is adequate for these purposes and should be employed.

Much of the research effort into the social aspects of human–computer interaction has been focused on the visual screen interface, which is operated by keyboard and mouse. The human–computer interaction that takes place through speech over a unimodal telephone channel is different from the visual interface and possibly even more sensitive to linguistic and social effects. The use of language in a user interface (and the use of speech in particular) is considered one of the most likely characteristics of technology that prompt a social response (Nass in Anderson, 2000, p. 95). Automated telephone services rely on speech output and the characteristics of the voice (such as the pitch, register and tone) carry sensitive information about personality and identity of the speaker. For example, Boyce (2000) compared a number of contrasting voice personalities which ranged from "from butler to hip youth" in a voicemail system and found that users reacted differently to these extremes. Some participants "loved" the butler personality whereas others found "him" annoying; the voice personalities that exhibited least extreme speaker characteristics caused fewer negative reactions from users (but also

resulted in fewer really strong positive reactions). Furthermore, the social interaction appears to be enforced further by the use of speech recognition technology in that it is not uncommon for users of speech-driven telephone applications to answer politely "yes please" or "no thank you" in response to system prompts.

6. The politeness experiment

6.1. Introduction

To explore issues in politeness with automated telephone dialogues, a controlled usability experiment was conducted in which participants ($N = 111$) experienced system-initiated digressions while they performed banking tasks using an automated telephone banking service. The system-initiated digressive proposals explored in this research explicitly stated in the opening phrase that the proposal constituted an interruption. This forthright method is likely to be perceived by users as more intrusive compared to the more low-key "*you might like to know*" opening phrases used in previous research (Wilkie et al., 2002, summarized in Section 3), but may however better serve to alert users to the ensuing information by capturing their attention. The purpose of making deliberate digressive interruptions in the current experiment was to explore if politeness strategies for human–human interaction (as defined by Brown and Levinson, 1987) could be employed to mitigate the adverse effects of these dialogue intrusions.

The experiment had four conditions based on the prompt register applied in the proposal: (1) Positive face-redress, (2) Negative face-redress (3) Bald (no face-redress) and (4) A no-proposal control condition. Participant attitudes towards the proposals were assessed, both in terms of the relative politeness of the proposal strategy in the context of the automated banking service (main experiment) and, secondly, the absolute politeness (Leech, 1983) associated with the face-threatening act, independent of dialogue context. The absolute politeness was established by allowing participants to listen to each individual proposal over computer speakers after they had completed the main experiment.

It is anticipated that user attitudes to system-initiated digressions will vary according to the relevance of the information to the user's specific situation. Determining what is, or is not, relevant to an individual caller is a complex matter involving modelling of the caller's intentions, wants, needs and goals: most of which are in the mind of the user and not accessible to the computer system. The research reported here does not address issues relevant to defining the business criteria or user models for deciding whether or not to make a proposal to a particular caller on a particular occasion; rather, it assumes that the decision to deploy the digression has already been taken. This approach is comparable with real life situations in which a (human) call centre agent reviews a customer's accounts and decides to approach the customer with a product offer, which may or may not be related to the original purpose of the phone call. The agent perceives a potential need for the product but

has little insight into the customer's general financial situation or needs, external to the details at hand.

Results from previous experiments (Wilkie et al., 2002) have already asserted that system-initiated digressions can successfully be deployed in the automated service without relying on complex user models. In the current experiment, the information contained in the system-initiated digression was chosen on the grounds of being financially beneficial and applicable to the majority of callers: i.e. the new "On-line Saver" account offers a higher interest rate than the accounts that the customer currently holds.

6.2. Participants

Participants were recruited from the general public in Edinburgh. Although some participants had used an automated telephone service for their personal banking, no previous experience of automated telephone banking was required in order to take part in the experiment. In total, 111 complete participant data sets were attained and used in the statistical analysis (Table 1).

6.3. Experiment procedure

Participants were told that they would use an automated telephone banking service to perform some banking transactions. For ethical and data protection reasons, no personal data were used at any point in the experiment. Participants were presented with a sheet of paper containing their fictitious persona to be used throughout the experiment: a membership number, a 6-digit personal telephone identification number (TIN) and details of "their" two (fictitious) accounts (a current account and a savings account). Prior to the first call to the automated service, participants were given a task sheet instructing them to find out and make a written note of the balance of "their" current account. Between phone calls participants were asked to imagine that "a few days had gone by" and that they were then to call the service to check their balance again. In total, participants made five phone calls to the automated service (the No-proposal control group made only three phone calls). The experiment proceeded in a number of clearly defined stages which are detailed below.

6.3.1. Two phone calls (No-proposal)
Each participant was asked to make two phone calls to the automated service (without any savings proposals being made at this stage). This procedure allowed all

Table 1
Participant demographics

	Age 18–35	Age 36–49	Age 50 +	Total
Female	28 (25.3%)	18 (16.2%)	17 (15.3%)	63 (56.8%)
Male	23 (20.7%)	15 (13.5%)	10 (9.0%)	48 (43.2%)
Total	51 (46.0%)	33 (29.7%)	27 (24.3%)	111 (100%)

of the participants to become familiar with the service functionality and their persona details. Following these two phone calls, the participants completed a questionnaire (referred to here as USAB1) to establish their attitude towards the usability of the service for later comparison after having experienced the proposal in their third phone call. The "USAB" questionnaire contents are detailed in Section 6.6.2.

6.3.2. Third phone call (with system-initiated proposal)

In the third phone call to the service, three-quarters of the participants experienced one of three randomly selected contrasting system-initiated digressions (Positive, Negative or Bald) while carrying out their banking enquiry (the No-proposal control group simply used the same banking service they had experienced in the previous two calls). By design, in order to avoid pre-empting participant reactions to the digression, no mention of savings proposals had been made up to this point in the experiment. Following this third phone call, all participants completed a second service usability questionnaire (USAB2). An additional questionnaire (PROP1) was also administered to participants who had experienced a proposal delivery during their phone call. The "PROP" questionnaire (detailed in Section 6.6.3) was targeted at user attitudes towards the interrupting digression itself (as moderated by the politeness strategy).

6.3.3. Two phone calls (additional proposals)

Participants (excluding those in the No-proposal group) were asked to make two additional phone calls to the service. These phone calls allowed participants to experience the remaining two face-redressive strategies in a controlled randomized order. Participant attitudes were assessed following each of these phone calls (PROP2 and PROP3). For these final two phone calls, only the questionnaire concerning attitudes towards the proposal itself (PROP) was used.

6.3.4. Listening session

After all of their phone calls to the service had been completed, participants[4] in the No-proposal control group listened to each of the three savings proposals over computer speakers. This was done in order to obtain a measure (manipulation check) of the absolute politeness of the proposals when abstracted from the context of the telephone banking dialogue. The results of the listening session data analysis revealed that the contrasting registers used in the proposals carried significantly discernable information regarding politeness attributes and that these findings were in-line with Brown and Levinson's theory. In summary: (1) the Negative face-redressive strategy was perceived to be most polite, apologetic and respectful and the

---

[4]All 111 participants took part in this session, however, only the data from the No-proposal control group ($N = 25$) were used in the analysis so as to avoid participants' responses being influenced by their experience of the proposal delivery in the context of the automated service.

speaker using the strategy was perceived to be most tactful, professional and caring; (2) the Positive face-redressive strategy was perceived to be least polite, formal, to the point and respectful and the speaker using the strategy was perceived to be least tactful and professional; and (3) the Bald strategy was perceived to be the most unapologetic, formal and to the point. For further details about these findings, see Appendix A.

6.3.5. Exit interview

A structured interview was then conducted containing questions relating to participants' perceptions and preferences of the politeness strategies used. Finally, participant details such as age and familiarity with automated telephone banking services were recorded. Participants then received an honorarium payment of £20.

6.4. The automated banking dialogue

The automated telephone banking service used in this research was modelled on an existing real-world application which provides customers with access to personal account information (e.g. balance information or recent transactions) and enables them to perform a number of banking transactions such as funds transfers. The service enables users to employ spoken natural language input by allowing for some extraneous speech (*Can I have…, please*) and the possibility of giving multiple pieces of information at the main menu (e.g. "*I'd like the balance of my current account, please*").

The application dialogue is outlined in the simplified flow-chart in Fig. 1, showing the system prompts and user responses for the identification and verification stage, followed by a balance enquiry.

The prompt style in the dialogue is terse, with politeness expressions limited to "please", "thank you" and "I'm sorry". Speech input is promoted throughout the service dialogue; each dialogue stage features a three error-level recovery where push-button options are mentioned in the third level (error recovery) prompt. The banking dialogue was implemented using commercially available speech recognition software. Prompts were recorded using a recording artist (female) who has a Southern British English accent.

6.5. Design of the digressive dialogue

The resulting three proposal variants have the following basic design criteria in common: they start out with an explicit interruption (mitigated by contrasting politeness strategies); they point out the financial benefits to the customer; they give details of restrictions that apply (that transfers to and from On-line Saver accounts can only be done via telephone or Internet banking); and, finally, they allow interested customers to pursue the offer immediately by engaging them in a "yes/no" (follow-on) dialogue. If the customer answers "no" at this point the service continues
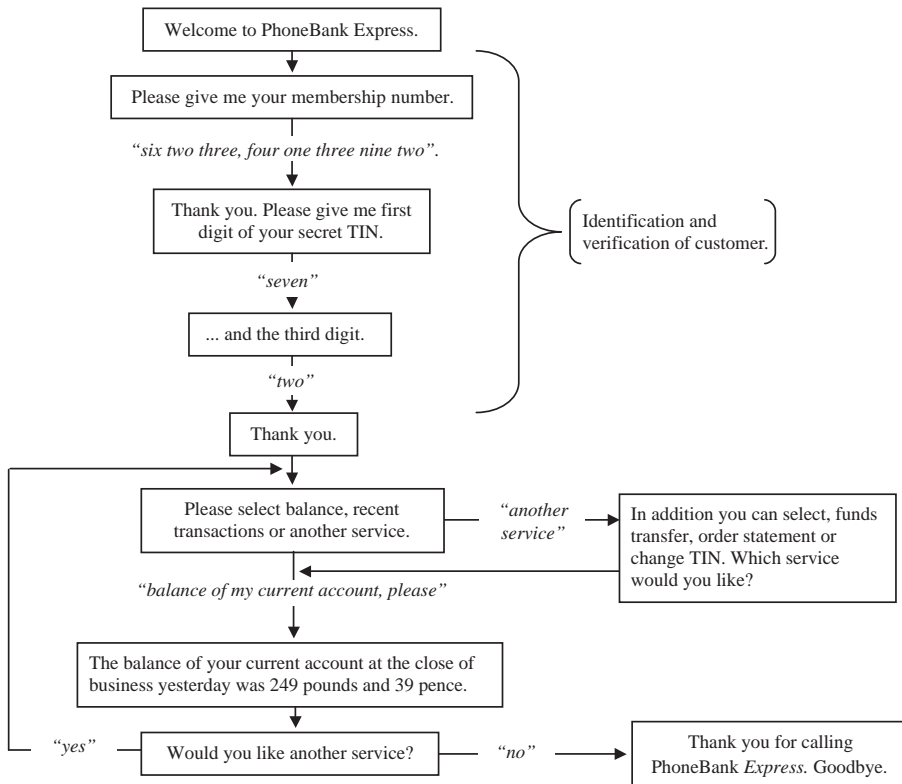
Fig. 1. Overview of the banking dialogue. System messages appear in boxes, and user responses are italicised.

the dialogue with "Would you like another service?". Participants who answer "yes" to the proposal hear the following message (note that the application procedure was simplified for experimental purposes):

"Thank you, your new On-line Saver account will be available from tomorrow."

For the purpose of the experiment, the system-initiated proposals were deployed immediately after a caller had been uniquely identified (after obtaining the two secret TIN digits, just before the prompt with the menu options: "*Please select balance...*" in Fig. 1). In real-world use, this location would ensure that only eligible customers were offered the savings proposal and that the proposal could be monitored such that it would be offered only once to each customer. The wordings for each of the three contrasting styles of proposals are detailed in the following sections.

6.5.1. Positive face-redress register

Brown and Levinson's theory states that threats to the addressee's positive face (through an interruption) are mitigated by using expressions of solidarity, informality and familiarity. Examples of positive face-redress are, exaggerating the interest in the addressee; sympathizing with the addressee; and avoiding disagreement. In the current experiment, the Positive face-redress was realized by the following linguistic devices (Brown and Levinson, 1987, pp. 101–129):

Being optimistic: "*I know you won't mind…*"
Informality: "*…cutting in…*"
Intensifying interest with the addressee: "*…special information for you…*"
Exaggerating approval with addressee: "*…make your growing savings grow even more*".
Presupposing common ground: "*we all want the best return possible…*"
Showing concern for the addressee's wants: "*with your interests in mind, I suggest…*"
Offering and promising: "*…an On-line Saver account that will give you better interest…*"
Giving or asking for reasons: "*why not set one up today!?*"

Proposal with Positive face-redress—(prompt recording 30 s long)

''I know you won't mind me cutting in with some special information for you about how to make your growing savings grow even more. We all want the best return possible from our savings. With your interests in mind, I suggest you open an ''On-line Saver account'' that will give you better interest than the accounts you've got just now. You can transfer money to and from an On-line Saver account through telephone or Internet banking. Why not set one up today! Do you want me to do that for you now?''

6.5.2. Negative face-redress register

Negative face-redress involves expressions of restraint, formality and distancing, such as being conventionally indirect, giving deference and apologizing. In the current experiment the negative face-redress was realized by the following linguistic contents (Brown and Levinson, 1987, pp. 129–211):

Apologizing: "*I'm very sorry to interrupt…*"
Stating the face-threatening act as a general rule: "*it is the bank's policy to notify…*"
Impersonalizing speaker and addressee: "*…notify customers how to…*"
Being indirect: "*we wish to inform you…*"
Giving deference: "*…as a valued customer…*"

Being pessimistic: "*you may therefore want to consider…*"
Going on record as not indebting addressee: "*we would be happy to…*"

Proposal with Negative face-redress—(prompt recording 31 s long)

"I'm very sorry to interrupt, but it is the bank's policy to notify customers about how to improve their savings returns. We wish to inform you, as a valued customer, that an "On-line Saver account" offers better interest than the accounts you hold at present. You may therefore want to consider opening an account of this type. Transfers to and from On-line Saver accounts are made through telephone or Internet banking. We would be happy to set up an On-line Saver account for you today. Would you like us to do that now?"

6.5.3. Bald register (No face-redress)

Undertaking a speech act without positive or negative face-redress is described by Brown and Levinson (1987) as performing the act "baldly". In contrast to the registers used to mitigate positive and negative face threats, the primary concern in the *Bald register* is to be direct and concise. The Bald register is applied under circumstances where the face threat can be ignored, in the interest of urgency and efficiency. The speaker might, e.g. feel that the information is so important or interesting to the addressee such that there is no need for a more convoluted expression. Alternatively, the speaker might be unconcerned about any imposition on behalf of the addressee. The Bald proposal in the experiment was stripped of any kinds of face-redress and started with: "*I'm interrupting to inform you about…*".

Bald proposal (No face-redress)—(prompt recording 18 s long)

"I'm interrupting to inform you about how to improve your savings returns. The "On-line Saver account" offers better interest than the accounts you have at present. You can transfer money to and from an On-line Saver account through telephone or Internet banking. Do you want to set up an On-line saver account now?"

6.6. Usability evaluation

6.6.1. Aim

The experimental research had two principal aims: (1) to establish whether the presence of a digressive interruption (as moderated by politeness strategy) influenced

participant attitudes towards the usability of the automated service; and (2) to evaluate the effects of contrasting politeness strategies on participant attitudes towards the interruption itself (in the context of the automated service dialogue).

6.6.2. Measurement of overall service usability

The design of the usability questionnaire (referred to in this paper as USAB) followed standard practice (Likert, 1932) by using an equal number of negative and positive statements presented to the respondent in a randomized order. In this way the danger that the overall usability score could reflect the respondent's tendency to agree rather than disagree with the questionnaire statements (an effect known as "*response acquiescence set*") is removed. Respondents mark their opinion for each statement by ticking the appropriate box along 7-point Likert scales that range from "Strongly Agree" (1) to "Strongly Disagree" (7). Following reversal of the polarity of positive questionnaire statements, in this paper a score of 7 consistently indicates a strong positive attitude and 1 a strong negative attitude.

Previous research has identified key attributes required for evaluating the usability of automated telephone interfaces and for assessing the contributions to usability made by each of the attributes (Love et al., 1992) by means of written questionnaires. The usability questionnaire used in this research consisted of 20 statements that address a range of issues pertaining to human–computer telephone interaction: *cognitive issues* (level of concentration and degree of confusion), *the fluency and transparency of the service* (knowledge about what is expected, ease of use, degree of complication), *system performance* (reliability of service, efficiency of service, amount of improvement service is felt to require) and *system voice* (clarity of the voice, politeness of the service, friendliness of the service).

All participants ($N = 111$) completed the service usability questionnaire following two "practice" phone calls (USAB1) and then again after their first exposure to the dialogue which included the system-initiated proposal (USAB2). Comparisons of the mean scores from these two questionnaires were used to establish the impact of the system-initiated proposal on participant attitudes towards the usability of the service.

6.6.3. Measurement of attitudes towards digressive proposals in the dialogue

Participants' reactions towards the system-initiated digressions were evaluated using two different approaches. Firstly, the manipulation check (detailed in Appendix A) allowed control group participants to experience each proposal in isolation over computer speakers, in essence removing the proposals from the context of the dialogue and focusing the analysis on the qualitative aspects inherent in the contrasting politeness strategies employed.

The second assessment approach involved capturing participant reactions towards the proposal interruption itself (as moderated by politeness strategy) in the context of the automated service dialogue. For this purpose, a supplementary set of questionnaire statements (referred to as PROP here) was added to the USAB usability questionnaire. The proposal attributes included in the PROP questionnaire were: *relative disruptiveness* (whether the proposal was annoying, intrusive,

distracting and interrupted the call too much), *face-redressive characteristics* (polite, friendly, formal, apologetic, patronizing, manipulative, caring for individual needs), *durational aspects* (length and long-windedness), *information quality* (helpfulness, efficiency, relevance of contents, appropriateness to context), *cognition* (ease of understanding) and *trust* (confidence in service, willingness to pursue the offer through the service). Participants ($N = 86$, excluding the No-proposal control group who did not experience a proposal in the context of the automated service) responded to these questionnaire statements (PROP1-3) following each exposure to a system-initiated proposal (which occurred in phone calls three, four and five during the experiment). The mean scores for these questionnaire items enabled direct comparisons of participants' attitudes towards the proposals based on the three contrasting politeness strategies employed.

7. Results

The results analysis was performed in two separate stages. Firstly, an assessment was made of the impact of the system-initiated proposal on participants' attitudes towards the automated service dialogue (USAB). This was achieved by comparing how participants ($N = 111$) rated the overall service usability before (USAB1) and after (USAB2) experiencing the first proposal delivery, which occurred in the third call to the service. The No-proposal (control) group used exactly the same service on all their three phone calls and did not experience a proposal in any of these calls.

The second analysis stage investigated participants' ($N = 86$) perceptions towards the system-initiated proposals specifically (PROP) and explored how attitudes towards the interruptions were affected by employing contrasting politeness strategies. The analysis compared participants' attitudes towards the contrasting proposal strategies based on their response data (PROP1) following the very first exposure to a proposal (in call three) and then, in a separate analysis, by pooling the mean scores (PROP1-3) following exposure to all three proposal variants (calls three, four and five).

In the analysis of the questionnaires, scales with participant responses were adjusted for polarity to ensure that all mean scores below 4 indicate a negative response to the statement, whereas values above 4 indicate a positive response.

7.1. Usability attitudes towards the automated service dialogue

The dependent measures used in these analyses were the mean responses to the questionnaire statements on service usability completed after the two familiarization calls (USAB1) and after the first exposure to the product proposal in the third call (USAB2).

The mean usability scores prior to and following the introduction of the unsolicited proposal for each individual experimental condition are shown in Table 2. Within each proposal group, repeated measures ANOVAs were carried out with age (3 levels) and gender as between-subject variables. Results showed that

Table 2
Mean usability scores based on proposal condition

Proposal condition	N	Mean usability before proposal	Mean usability after proposal	Statistical results
No proposal[a]	25	5.71	5.72	*df=1, F=.14, p=.710*
Positive face redress	29	5.56	4.94	*df=1, F=10.27, p=.004*
Negative face redress	28	5.79	5.10	*df=1, F=24.25, p=.000*
Bald (no face redress)	29	5.72	4.83	*df=1, F=28.06, p=.000*

[a]Control group.

there was no significant change in attitude for the No-proposal group before and after call three, but for each of the three proposal conditions there is a noticeable drop in the attitude towards the service following the introduction of a proposal delivery. The change in attitude for each of the three proposal groups in Table 2 was highly significant ($p < .01$).

A univariate ANOVA was carried out based on the mean score differences between the two questionnaires (USAB2-USAB1). The between-subject variables used in the first analysis were age (3 levels), gender and presence/absence of proposal (2 levels). Results showed that, when compared to the No-proposal control group, the overall drop in attitude for participants who experienced the presence of a proposal was significant [$df = 1$, $F = 15.05$, $p = .000$]. Analysis of mean score differences for individual questionnaire items revealed that proposal group participants found the service significantly ($p < .01$) more *frustrating* and *less enjoyable to use*, making it *less efficient* and *more in need of improvement*. Participants felt *more under stress*, *less in control* when using the service and they were *less happy about using the service again*. At a lower level of significance ($p < .05$), participants found the service with the proposal *more confusing*, *more complicated* and *less easy to use*. They felt *more flustered* when using the service, they had to *concentrate harder* and *knew less what to do*.

A further univariate ANOVA analysis was based solely on the mean score differences between the three proposal groups, with age (three levels), gender and proposal strategy (three levels) as the between-subject variables. The results showed that there was no significant difference overall between the three proposals [$df = 2$, $F = 1.85$, $p = .165$]. Further analyses revealed that only two questionnaire items were statistically significant ($p < .05$): *the service was easy to use* and *the service was reliable*.[5]

---

[5]Where the results of statistical tests such as *t*-tests or ANOVAs show only one or two significant differences in a set of 20 questionnaire items, these should not be relied on since it is a statistical fact that when a number of such tests are carried out there is a high probability that at least one at the 95% level will be a false positive.

Table 3
PROP questionnaire mean scores, based on proposal condition and call number

Proposal condition	PROP1 mean first proposal	PROP2 mean second proposal	PROP3 mean third proposal	Overall mean
Positive	4.02	3.90	3.57	3.83
Negative	3.81	4.20	4.14	4.06
Bald	3.72	4.50	4.45	4.22
Total mean	3.85	4.20	4.05	

In summary, the presence of a proposal in the dialogue had a significantly negative impact on service usability overall, while there were no significant differences in mean score differences between the three proposal groups. These results suggest that it was the presence—rather than the politeness strategy—of the proposal that had the major impact on attitudes toward the service usability in this experiment.

7.2. Attitudes towards the digressive proposals

Two analyses were carried out on the PROP questionnaire items, which were specifically aimed at capturing user attitudes towards the proposal interruption and the politeness strategy employed. Firstly, a univariate ANOVA was carried out on the responses after the participant's first exposure to the proposal dialogue (PROP1). Secondly, repeated measures ANOVA was carried out on the data after exposure to all three contrasting proposals (PROP1-3), where responses had been pooled according to the style of the proposal (Positive, Negative and Bald). The No-proposal control group did not experience any proposals and are therefore not included in these analyses. Mean scores for the PROP questionnaires are presented in Table 3.

7.2.1. Analysis based on first exposure to a proposal

The between-subject analysis was performed on participant responses (PROP1) to their first exposure to the proposal dialogue, which occurred during their third phone call to the service. This simulates the reactions from customers who encounter the (unsolicited) proposal for the first time during automated telephone banking. The results from the analysis show that there were no significant differences between proposal strategy groups overall. There were some differences in attitude for individual items in the questionnaire however, based on the sample size used in this experiment, these were not strong enough to produce statistically significant result.[6] Thus, it can be concluded that there were no differences in the way participants responded to the contrasting face-redressive strategies employed in the proposals based on first proposal exposure.

---

[6]In fact, the questionnaire item regarding *"the proposal interrupted the call too much"* showed a weakly significant difference. This result, on a single item in a 24-item questionnaire, could easily be due to chance.

   The mean scores show that the general attitude towards any of the three proposals was negative; more than half of the questionnaire item scores fell below (or nearing) the neutral point 4 on the 7-point scale for all proposal strategies. Items that were aimed at eliciting the face-redressive characteristics of the proposal register received generally positive responses. In particular, the items relating to the *politeness* and *friendliness* achieved mean scores on, or above, 5 on the scale. In addition, with scores above neutral, participants did not seem to think that either proposal was *too apologetic*, *too formal* or *patronizing*. Questionnaire items that resulted in markedly negative responses (scores below 3) related to the disruptiveness of the proposal in the call: the proposal was perceived to be *distracting, intrusive, too long, annoying* and believed to *interrupt the call*.

   In summary, results based on a participant's first exposure to the proposal indicate that there were no significant differences between politeness strategies employed, with regards to the sample size used in this analysis.

7.2.2. Analysis of pooled response data

   Participant responses for all three questionnaires (PROP1-3) were pooled according to politeness strategy. The pooled-data approach has two main advantages: it increases sample size and enables the use of within-subject comparisons (which in turn reduces the unsystematic variability in the design and provides greater power to detect effects). The main disadvantage with the pooled-data approach is that it includes data from the second and third proposal calls where the proposal content no longer is new or unexpected (creating a learning effect). As a consequence, if participant responses to the first exposure are significantly different compared with subsequent exposures, then two different conditions—"proposal novice" and "proposal-aware" participant groups—are mixed in the results.

   A repeated-measures ANOVA was carried out on the pooled data with one within-subjects variable (proposal strategy) and three between-subject variables: age (3 levels), gender and proposal order (6 levels, based on all possible controlled permutations of exposure). The analysis of the overall difference between participant mean scores (column labelled "Overall Mean" in Table 3) revealed an overall statistically significant difference for proposal strategy [$df = 2$, $F = 4.629$, $p = .012$]. Within-subject contrasts showed that this difference lay between the Positive face-redress proposal and the Bald strategy [$df = 2$, $F = 11.432$, $p = .001$].

   The analysis also showed a significant interaction between proposal strategy and order group [$df = 10$, $F = 2.528$, $p = .009$], indicating that participant attitudes towards the proposals were confounded with one of the following: (a) order effect due to call number; (b) exposure to preceding proposals, or; (c) an interaction between call number order effect and the current proposal wording. Following this finding, the data were adjusted to compensate for the effect due to call number by subtracting the overall questionnaire mean for each call number (e.g. 3.85 for PROP1, Table 3) from individual participant mean scores within that proposal call. The new mean scores were then used in a re-run of the repeated-measures ANOVA. Results showed that that the significant effect of proposal strategy remained unchanged, but that the interaction between proposal strategy and order of exposure
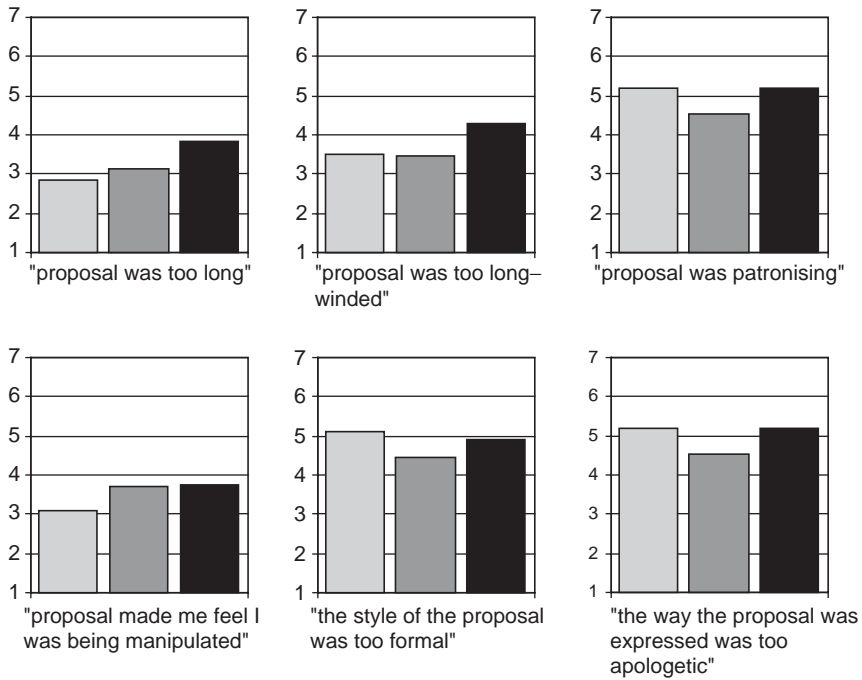
Fig. 2. Mean responses by condition. ☐ Positive; ▨ Negative; ■ Bald.

became much weaker and was no longer significant [$df = 8$, $F = 1.059$, $p = .240$].[7] These results support the theory that there are two simple effects present: a simple effect of proposal strategy regardless of previous exposure and a simple effect of call number regardless of strategy involved. To conclude, this suggests that there is a genuine effect of proposal strategy, applying in both the ''novice'' and the ''proposal-aware'' conditions, but it is not conclusive since these results were not reflected in the analysis of the data from the first proposal call.

Further ANOVAs, performed on individual statements in the unadjusted pooled data scores, revealed a number of attributes with highly significant differences between the three proposal styles, as illustrated graphically in Figs. 2 and 3 (mean scores and results from the statistical analysis are shown in Table 4). Note that higher mean scores indicate a more positive and supportive attitude towards the concepts conveyed by the Likert statements in the questionnaire. For example, the first of the charts in Fig. 2 reveals that the Bald proposal generated a more positive response regarding the proposal length compared to the Positive and Negative proposal strategies.

In addition to being favoured in terms of its shorter *length*, the Bald proposal was also perceived by participants to be significantly *less long-winded* compared to the

---

[7]The *F*-value and degrees of freedom here have been adjusted for the fact that the means used for compensation were estimated from the data.
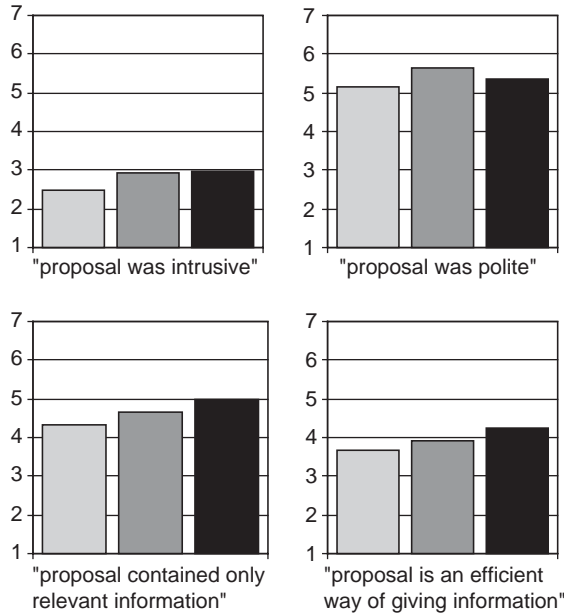
Fig. 3. Mean responses by condition. ☐ Positive; ▨ Negative; ■ Bald.

Positive and Negative face-redressive proposal strategies. This suggests that attributes such as the length and wordiness of a proposal have a strong impact on user attitudes towards system-initiated digressions.

Fig. 2 also highlights participants' reactions towards the face-redressive characteristics employed in the contrasting politeness strategies. The Positive face-redress proposal which relied on an informal and intimate register was found by participants to be *more manipulative* than the Negative face-redress and the Bald strategy. The Positive face-redress proposal was also perceived to be significantly more *patronizing* than the Bald strategy. The Negative face-redress was found to be significantly more *formal* and *too apologetic* when compared to the Bald strategy and the Positive face-redress.

Participant responses highlighted in Fig. 3 give some further indications to participants' perceived differences of the proposal strategies. In terms of the relative *intrusiveness* there was a significant difference in attitude between the Positive face-redressive proposal and the Bald strategy, the Positive face-redress being perceived to be *more intrusive*. The Negative face-redressive proposal was rated most *polite* of the three, with the difference between the Positive and Negative proposal strategies approaching highly significant ($p = .012$). Noticeably, all three proposal strategies received strong positive scores (>5) in terms of perceived politeness. The comparatively high mean score for the Bald proposal (lacking face-redress) suggests that the perceived politeness of a proposal strategy is determined relative to the context in which it occurs and not only as a consequence of using expressions which are commonly associated with politeness, such as "I'm sorry" and "thank you".

Table 4
Within-subject contrasts of the three proposal conditions for questionnaire items that showed a statistically significant main effect of proposal condition [$df = 2$, $*p<.05$; $**p<.01$; $***p<.001$]

Questionnaire Item	Positive face means	Negative face means	Bald strategy means	Positive vs. Negative face redress $F =$	Negative face redress vs. bald strategy $F =$	Positive face redress vs. bald strategy $F =$
The style of the proposal was too formal	5.12	4.45	4.88	13.69**	5.98*	1.61
The proposal was too long	2.85	3.13	3.83	1.19	8.64**	18.00***
The proposal made me feel I was being manipulated	3.09	3.73	3.76	5.50*	.20	13.15**
The proposal was an efficient way of giving information about the On-line Saver account	3.66	3.93	4.23	.60	2.63	6.89*
I found the proposal intrusive	2.48	2.94	2.99	3.29	.83	11.98**
The proposal was polite	5.15	5.65	5.33	6.74*	4.32*	.44
The proposal contained only relevant information	4.34	4.65	4.99	3.93	1.83	7.58**
The proposal was very long-winded	3.49	3.48	4.28	.12	8.39**	7.02*
I found the proposal patronising	3.55	4.19	4.55	3.08	5.01	14.84***
The way the proposal was expressed was too apologetic	5.19	4.53	5.21	11.20**	10.23**	.63

The Bald proposal was considered to contain the most amount of *relevant information* of the three proposals, but this was only the difference between the Positive and Bald proposals that showed statistically significant results. In terms of *efficiency*, there was no preferred strategy among the proposals. Not even the Bald strategy (which is aimed to be short and terse to promote efficiency) was rated strongly positively and it was only slightly more favoured than the Positive face-redressive proposal.

In summary, the analysis of the pooled responses highlighted differences in participants' perception of the contrasting politeness styles and registers employed in the proposals. The Bald proposal strategy was perceived to be significantly shorter and less long-winded than the Positive and Negative face-redressive strategies. In line with Brown and Levinson's theory, the wording in the Negative face-redressive strategy was perceived to be more formal, more polite and more apologetic. The Positive face-redressive strategy was rated as the most manipulative of the three proposals and it was considered significantly more patronizing and intrusive than the Bald strategy.

7.3. Task completion

In each call, participants were asked to telephone the service and find out the balance of their current account and then to take a note of the amount on their task sheet. In the third call to the service, participants experienced the product proposal and then had to accept or reject the proposal to set up an (On-line Saver) savings account straight away. Following this, the automated service then asked participants "Would you like another service?" and participants were required to answer "yes" in order to proceed with their account balance enquiry. Successful balance task completion rates for the two (practice) phone calls (1 and 2) and the proposal phone call (3) are shown in Table 5.

The lowest task completion (76%) occurred in the participant group which experienced the Bald style of proposal. When participants in this group were asked if they would like another service, seven out of 22 individuals answered (wrongly) "no" and their call was transferred from the service.

7.4. Interview comments

At the end of the experiment session an opportunity was taken to investigate each participant's reactions to a number of issues raised by their experience of the product

Table 5
Task completion rate: participants who manage to obtain the account balance in a call

Proposal condition	First phone call	Second phone call	Third phone call
No proposal	26 (100%)	25 (96%)	25 (96%)
Positive	28 (97%)	29 (100%)	26 (90%)
Negative	28 (100%)	28 (100%)	28 (100%)
Bald	29 (100%)	29 (100%)	22 (76%)

proposal in the service, including direct comparisons of the three different politeness styles of proposals. This involved a structured interview in which the question order and wording remained the same for each participant. Most of the questions in the interview required the participant to select a proposal of their choice,[8] with the option for participants to volunteer additional comments. Participants were encouraged (but never required) to give more detailed reasons for their responses. The purpose of the interview was to allow participants to express more freely their thoughts about the wording and style of the proposals.

When asked about *which of the three proposals they preferred*, the majority of participants (54%) chose the Bald strategy stating it was shorter and more to the point than the other two. This group also commented that they perceived the Bald style of the proposal as "less patronizing", "less intrusive", "less formal", "more honest" and "more professional". The Negative face-redress proposal received 29% of participants' votes for preferred choice, claiming that they preferred it because it was "polite" and "apologetic" and referred to specific appealing expressions used in the proposal such as "sorry to interrupt", "bank's policy" and "happy to set up". There was no strong consensus in the comments for participants who said they preferred the Positive face-redress (11%). Examples of comments were that the positive style proposal was: "more positive", "more caring", "more polite" and "not so apologetic". Interestingly, when examining only the responses from the control-group participants (who had not experienced the proposal during their use of the automated service), 50% of participants preferred the *Negative* face-redress proposal whereas 38% were in favour of the Bald strategy and 12% the Positive face-redress.

When asked which of the three proposals they perceived as the *most polite way to address the caller*, the Negative face-redress strategy generated a majority (66%) of participants' votes. Consistent with Brown and Levinson's theory, most of the participants who selected the Negative face-redress as the most polite regarded the apology in the opening statement of the proposal as the primary reason for their choice. Here, 18% of participants chose the Bald strategy as the most polite way to address the caller, mainly commenting on that they thought it was "less patronizing" and "less apologetic". The remaining 13% of participants who chose the Positive face-redress proposal did so as they thought it was the most polite way to address the caller. Their comments were: "more familiar", "more natural" and "not as blunt as the Bald strategy nor as apologetic as the Negative face-redress".

In addition, participants were asked which of the proposals was the *least polite way to address the caller*. The Positive face-redress was selected by 48% of participants as the least polite way, mainly commenting on the opening statement in the proposal "I know you won't mind" which many perceived as presumptuous. Participants who chose the Bald proposal (35%) as the least polite of the three said they found it "abrupt" and that they did not like the opening statement (I'm interrupting to inform you...). Only 5% of participants thought that the Negative face-redress was the least polite way to address the caller. In this case, half of the

---

[8]Some participants selected more than one proposal. In order to simplify the discussion in this section, only participant responses where one proposal was selected were included in the analysis.

comments regarded the statement about "bank's policy" as indicative of less concern about the customer's finances.

Finally, participants were asked which of the proposals they found to be *the friendliest*. In this case, the Negative face-redress proposal received the majority (49%) of participants' votes with reasons that it was "apologetic", "personalized", "more human sounding" and that they liked the phrase "valued customer". Participants who thought the Positive face-redress proposal was friendliest (29%) said they found it to be "more informal" and "more personalized". The remaining 15% who selected the Bald strategy did so mainly because they thought it was better in comparison with either of the face-redressive proposals.

In summary, participants comments indicated a preference for the Bald strategy to be employed for system-initiated digressive dialogues. However, a significant proportion of participants favoured some kind of politeness strategies indicating that users were aware of the importance of mitigating face strategies in human–computer interaction.

8. Discussion

This paper has described an experiment in which participants ($N=111$) experienced a digressive proposal offering a new product to the caller as part of the interactive dialogue of an automated telephone banking service. The opening phrase in the proposal explicitly stated that the proposal constituted an interruption. Three contrasting politeness strategies (Positive, Negative and Bald), derived from established face-redress theories in human–human communication, were employed in order to mitigate the adverse effects of these dialogue intrusions. Participants' attitudes towards these three proposals were explored, both in terms of their impact on perceived usability of the banking service and the perception of the interrupting digression itself (as moderated by politeness strategy).

The experiment data presented in the paper reveal that the usability of the spoken telephone banking service is reduced with the introduction of these digressive interruptions in the dialogue. Participants' initial mean usability score of 5.69 (7-point response scale) fell to a mean score 4.96 after they had experienced their first proposal. This significant reduction in usability was observed for each of the three politeness strategies explored—Positive face-redress, Negative face-redress and Bald register. Participants found the service with such proposals "more frustrating to use", "less enjoyable", "less efficient" and "more in need of improvement". The proposals also placed more cognitive strain on the participants rendering the interaction "more confusing", the service "more complicated" and "less easy to use". Interestingly, the results show that the types of apology and politeness used in the Negative face-redress strategy (which are typically associated with politeness etiquette) were not effective. The use of "I'm very sorry to interrupt…" in the Negative face-redress was no better received than the phrase "I'm interrupting…" in the Bald strategy.

Despite employing contrasting polarities of face-redressive strategies, there were no overall significant differences between the three proposals, based on participants' first exposure. Two real-life issues were considered in detail in the design of the experiment that may have modified participants' perception of the proposal, making the dissimilarities between the contrasting strategies more prominent. Firstly, since the digressive proposal forms only a brief part of a larger automated banking dialogue, the impact of the contrasting proposal strategies may have been more strongly differentiated had participants been forewarned about the pending interruption; this may have encouraged participants to pay more careful attention to the contents and wording of the message. In a real-life scenario, however, it is difficult to envisage how and when such warnings might be delivered to customers and, in consequence, the un-primed (worst-case) approach was adopted in the experiment. Secondly, the scenarios in the experiment might have been extended to involve a secondary task implicitly instructing participants to maximize their savings returns, thereby making them more positively disposed towards the product introduced in the digression offer. However, if such digressions were introduced in a real-world automated telephone banking service, possibly based on some assessment of a customer's individual need for the product, there is no guarantee that the customer would actually share the enterprise's perceived need for that information. Hence in the experiment design a totally un-targeted (worst-case) approach was adopted. The un-primed, un-targeted scenario approach had been used successfully in previous dialogue digression experiments by the authors (Wilkie et al., 2002).

The analysis of the pooled response data (after participants had experienced each of the three proposals) revealed that there were significant differences overall between the Bald and the Positive face-redress strategies (in favour of the Bald strategy). The results provide some guidance on the design issues involved in attempts to add such digressions to automated telephone dialogues by eliciting participants' preferences for the wordings of such proposals. The Bald strategy received significantly more positive responses in terms of being shorter, less long-winded and contained more relevant information. The Positive face-redress, on the other hand, was found to be significantly more manipulative, patronizing and intrusive. In the post-experiment listening tests, support of the Bald proposal strategy was strengthened: 54% of participants expressed a preference for the Bald strategy with the main arguments that it was shorter and more to the point than the other designs. Interestingly, however, 50% of participants in the (No-proposal) control group chose the Negative face-redress as their preferred proposal strategy. When heard in isolation, the Negative face-redress might seem the most appropriate design choice when approaching a customer; in the context of the automated telephone banking service, however, the Negative face-redress approach was shown to be judged as lengthy, long-winded and was perceived to be too apologetic and formal.

Much of the research of anthropomorphic computer behaviour in human–computer interaction to date has primarily focused on the visual user interface; the impact of social phenomena, such as politeness, in the audio-only interface have yet

to be fully explored. The current research contributes to the debate on anthropomorphism in computer systems by exploring the issue of endowing a speech-only human–computer dialogue with specific forms of politeness. In contrast to the visual user interface, the audio-only interface is incapable of displaying multiple pieces of information simultaneously; the system will dominate the dialogue for as long as it takes to deliver a spoken message and the user is not offered the opportunity to rapidly scan information that seems irrelevant. It follows that choice of appropriate wording, duration and speaker characteristics are pivotal in the design of audio interfaces—as demonstrated in this paper. These issues raised here lend themselves to further research in order to obtain a deeper understanding of pronounced forms of speaker characteristics, linguistic behaviour and user expectations unique to such audio-only computer interfaces.

For a given communicative situation between humans, it has been shown that the choice of politeness strategy depends on the mutual expectations about the power relationship and social distance between the interactants, coupled with the degree of imposition involved in making the face-threatening act in that communicative context. When a speaker is overly polite, unexpectedly unfriendly or irrational, or strays from the topic in a human–human conversation, the addressee will draw conclusions about the reasons why the speaker does not behave as expected. This may, e.g. involve re-evaluating the assumptions about their social relationship with the consequence that politeness (or its absence) in a dialogue can serve to modify the social distance or power relationship between interactants. The negative reactions towards the face-redressive strategies employed in the proposals may be attributed to the fact that these were not perceived as being fully integrated with users' assumptions about the relationship with the service, formed by the speaker characteristics presented in the rest of the banking dialogue. Whilst applications, such as the automated banking service explored in this research, are primarily viewed as tools, with repeated use there is the potential that customers will develop aspects of rapport with the service. Endowing audio-only interfaces with personas, which consistently exhibit Negative or Positive face-redressive behaviour as presented here, may thereby serve to enhance this human–computer relationship.

Acknowledgements

Appendix A

In order to establish the absolute politeness in the proposals (i.e. attitudes towards the politeness strategies when removed from the context of the telephone dialogue), an additional session was included at the end of the experiment in which control-group participants listened to each proposal over computer speakers. The aim of the
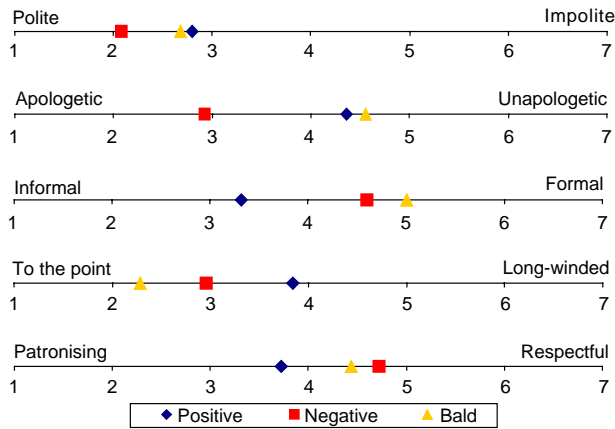
Fig. 4. Control-group participant mean responses. These questionnaire items were introduced to participants with the phrase "*Thinking about the proposal I've just heard, it was…*".
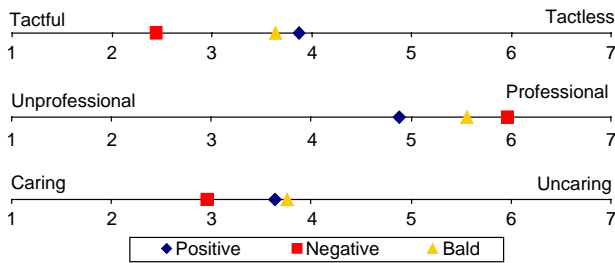

Fig. 5. Control-group participant mean responses. This section was introduced to participants with the phrase "*I associate the choice of wording in the proposal with someone who is…*".

listening session was two-fold: (1) explore the participant's perception of the register and speaker characteristics employed in the contrasting proposals and (2), whether the contrasting face-redress strategies would produce effects consistent with Brown and Levinson's theories.

Immediately after hearing a proposal, participants completed a questionnaire featuring descriptive antonym pairs (such as polite vs. impolite) are presented at either end of a 7-point (semantic differential) scale (Osgood et al., 1957). The first set[9] of antonyms concerned the style and register used in each of the contrasting proposals and were introduced to the respondents with the sentence: "*Thinking about the proposal I've just heard, it was…*". Respondents marked their opinions by ticking the appropriate box along the scale. The second set[10] of antonyms were aimed at assessing some of the social characteristics and personality of the speaker. Response

---

[9] The statements were: polite/impolite; informal/formal; to the point/long-winded; forthright/diplomatic; sincere/insincere; respectful/patronizing; personalized/impersonal; apologetic/unapologetic.

[10] These items were: tactful/tactless; timid/self-confident; sociable/unsociable; reliable/unreliable; caring/uncaring; professional/unprofessional.

Table 6

Statistical analysis of the absolute politeness in the contrasting proposals in the listening session [$df = 2$, $*p < .05$; $**p < .01$; $***p < .001$]

	Positive vs. Negative face redress $F =$	Negative face redress vs. Bald strategy $F =$	Positive face redress vs. Bald strategy $F =$
Polite–impolite	3.47	5.82*	.38
Apologetic–unapologetic	8.42**	11.91**	3.94
Informal–formal	8.95**	2.18	29.07***
To the point–long-winded	21.39***	.84	10.40**
Patronising–respectful	31.34***	2.24	2.84
Tactful–tactless	20.48***	13.83***	.07
Unprofessional–professional	19.84***	2.71	3.46
Caring–uncaring	4.71*	7.65*	1.39

data was first pooled according to the politeness strategy employed (Positive, Negative and Bald) and a three-level repeated-measures ANOVA was then performed based on individual questionnaire attributes. Figs. 4 and 5 summarize the results of the absolute politeness check (further details of the statistical significance are included in Table 6).

The tendency to view the Negative strategy as polite, apologetic and tactful is consistent with Brown and Levinson's theory which states that listeners commonly associate expressions of Negative face-redress with the everyday use of the term politeness—it is "the stuff that fills etiquette books". In contrast, the Positive face-redress is realized through more intimate linguistic output strategies where the aim is to show appreciation and care of the addressee's wants in general, or express the similarity between the speaker and addressee's wants.

The Positive face-redress was perceived to be significantly more *long-winded* than the Negative face-redress ($p < .01$) and the Bald strategy ($p < .01$). This is interesting as it indicates that it is not primarily the duration (the Positive and Negative face-redress proposals were of the same length, ± 1 s) but the choice of wording that contribute to the addressee's perception of long-windedness in the proposal.

References

Allen, J.F., Guinn, C.I., Horvitz, E., 1999. Mixed-initiative interaction. IEEE Intelligent Systems 14 (5), 14–23.

Anderson, R.I. (Ed.), 2000. Conversations with Clement Mok and Jakob Nielsen, and with Bill Buxton and Clifford Nass. Interactions 7(1), 46–80.

Balentine, B., Morgan, D.P., 2001. How to Build a Speech Recognition Application, second ed. EIG Press, ISBN: 0-9671287-2-3.

Bernsen, N.O., Dybkjær, H., Dybkjær, L., 1997. Elements of speech interaction. In: Dybkjær, L. (Ed.), Proceedings of the Third Spoken Language Dialogue and Discourse Workshop, Vienna, September 1997.

Boyce, S.J., 2000. Natural spoken dialogue systems for telephony applications. Communications of the ACM 43 (9).

Brown, P., Levinson, S.C., 1987. Politeness: Some Universals in Language Use. Cambridge University Press, Cambridge.

Chen, R., 2001. Self-politeness: a proposal. Journal of Pragmatics 33, 87–106.

Colón, J.X.E., Pérez-Quiñones, M.A., Ferreira, R., 2001. Effects of face-threatening acts in human–computer dialogues. Proceedings of HFES'01 (Human Factors and Ergonomics Society), pp. 657–662.

Culpeper, J., 1996. Towards an anatomy of impoliteness. Journal of Pragmatics 25, 349–367.

Fogg, B.J., Nass, C., 1997. Silicon sycophants: the effects of computers that flatter. International Journal of Human–computer Studies 46, 551–561.

Gardner-Bonneu, D., 2001. Human Factors and Voice Interactive Systems. Kluwer Academic Publishers, Dordrecht, 1999, Second Printing 2001, ISBN 0-7923-8467-9.

Haller, S., 1994. Recognizing digressive questions. In: Proceedings of AAAI94, Fall Symposium 1994.

Hone, K.S., Baber, C., 1999. Modelling the effects of constraint upon speech-based human–computer interaction. International Journal of Human–Computer Studies 50, 85–107.

Leech, G.N., 1983. Principles of Pragmatics. Longman Group, Ltd., New York.

Likert, R., 1932. A technique for the measurement of attitudes. Archives of Psychology 140.

Love, S., Dutton, R.T., Foster, J.C., Jack, M.A., Nairn, I.A., Vergeynst, N.A., Stentiford, F.W.M., 1992. Towards a usability measure for automated telephone services. Proceedings of the Institute of Acoustics Speech and Hearing Workshop 14 (6), 553–559.

McFarlane, D., 1998. Interruption of people in human–computer interaction. Ph.D. Thesis, The George Washington University, August 1998.

Narayanan, S., Di Fabbrizio, G., Kamm, C., Hubbell, J., Buntschuh, B., Ruscitti, P., Wright, J., 2000. Effects of dialog initiative and multi-modal presentation strategies on large directory information access. In: Proceedings of the International Conference on Spoken Language Processing, ICSLP 2000, Beijing, China, pp. 636–639.

Nass, C., Moon, Y., 2000. Machines and mindlessness: social responses to computers. Journal of Social Issues 56 (1), 81–103.

Nass, C., Steuer, J., Tauber, E.R., 1994. Computers as social actors. Proceedings of the CHI '94, Conference of the ACM/SIGCHI, Boston, MA, April 1994.

Osgood, C.E., Suci, G.J., Tannenbaum, P.H., 1957. The Measurement of Meaning. University of Illinois Press, Champaign, IL.

Ramakrishnan, N., Capra, R., Pérez-Quiñones, M.A., 2002. Mixed-initiative interaction = mixed computation. In: Thieman, P. (Ed.), Proceedings of the ACM SIGPLAN Workshop on Partial evaluation and Semantic-Based Program Manipulation (PEPM'02), January 2002, pp. 119–130.

Reeves, B., Nass, C., 1996. The Media Equation. Cambridge University Press, Cambridge.

Shneiderman, B., 1988. A nonanthropomorphic style guide: overcoming the Humpty Dumpty syndrome. The Computing Teacher 9–10.

Shneiderman, B., 1993. Beyond intelligent machines: just do it!. IEEE Software 10 (1), 100–103.

Shneiderman, B., 1998. Designing the User Interface: Strategies for Effective Human–Computer Interaction, third ed. Addison Wesley Longman, Inc.

Shneiderman, B., 2000. The limits of speech recognition. Communications of the ACM 43 (9).

Wilkie, J., Jack, M.A., Littlewood, P., 2002. Design of system-initiated digressive proposals for automated banking dialogues. Proceedings of the International Conference on Spoken Language Processing, ICSLP 2002, pp. 1493–1496.