Operetta dialogue engineering life-cycle

DISC partner: MIP
Authors: Laila Dybkjær and Niels Ole Bernsen

Overall design goal(s): *(What is the general purpose(s) of the design process?)*

To develop a commercial call answering and routing system with voice mail. The benefits of the system should be no lost calls, calls are answered immediately, the system may off-load an overloaded switchboard operator thereby allowing the switchboard operator to more instantly handle non-routine calls and other tasks, users can just speak a name and do not need to remember extension numbers.

Hardware constraints: *(Were there any a priori constraints on the hardware to be used in the design process?)*

Historically, the SR platform specified was a Pentium 486 under UNIX with Dialogic telephony card.

Software constraints: *(Were there any a priori constraints on the software to be used in the design process?)*

The recogniser software was written in C so this was the specified language. **The telephony stuff was off-the-shelf, everything else was home made.**

Customer constraints: *(Which constraints does the customer (if any) impose on the system/component? Note that customer constraints may overlap with some of the other constraints. In that case, they should only be inserted once, i.e. under one type of constraint.)*

The system is not intended for one particular customer. The customer may be any organisation with a switchboard.

Other constraints: *(Were there any other constraints on the design process (e.g. on cost, manpower, purchase price, development time, standards conformation).)*

The product had to be competitive with other DTMF-based systems.

Design ideas:

Basically the voice routing idea was the key idea.

Designer preferences: *(Did the designers impose any constraints on the design which were not dictated from elsewhere (e.g. programming language preferences, development methodology)?)*

The software which provides the run-time engine (CAGE) had its constraints.

Design process type: *(What is the nature of the design process (exploratory research, product development, redesign, other (explain))?)*

Product development.

Development process type: *(How was the system/component developed (e.g. through Wizard of Oz, using existing development methodology x, y, z).)*

Operetta developed from a technology basis, i.e. here is this interesting SI recognition, what can we do with it?. WOZ testing of dialogue was not used.

Requirements and design specification documentation: *(Is one or both of these specifications documented?)*

Yes, but not available. DW can filter information from them, **see DISC web page**.

Development process representation: *(Has the development process itself been explicitly represented in some way? How (e.g. bits and pieces of it can be found in scientific papers, the entire process was carefully documented in semi-formal notation, most of the process has been systematically represented in reports).)*

No, but DW thinks it is fairly representative of a new technology application. Operetta has developed in a piecemeal way.

Realism criteria: *(Will the system/component meet real user needs, will it meet them better, in some sense to be explained, than known alternatives, is the system/component "just" meant for exploring possibilities, or what (explain)?)*

The system is meant to meet organisations' needs of somebody answering the phone immediately (no queue) and no matter at which time people call. The system may also be configured to turn itself on and off to fit in with the customer's business practices. The activation times can be set up and changed by the system administrator. The idea is not to get rid of the switchboard operator but to reduce the workload on him/her.

Functionality criteria: *(Which functionalities should the system/component have (this entry expands the overall design goals, e.g. "allow users to do tasks X and Y", "include barge-in", "real-time"). Note that this entry is more general than, but may partially overlap with, the "grid" properties.)*

The system must allow people calling the organisation to be connected to the person to whom they wish to talk, to be connected to a human operator if the person is not available, or to leave a message even if the operator is not in; the system runs in real time; DTMF interrupt is possible by typing in the extension number of the person one wants to speak to. The caller may also press * during the greeting which allows him to interrupt and go straight to speaking the name of the person he wants. If a caller asks for a department the system can be configured to ring the number of a hunt or ring group, if they are supported by the switch.

Limitations are that not all switches are supported, the phone book can contain a maximum of 100 names, and the system only supports time break recall.

The DTMF driven menus of Operetta voice mail give the following features:

Change your greeting

Change your password

Message facilities:

- Listen to new messages
- Send messages
- Archive messages
- Reply to messages sent from other Operetta mailboxes
- Forward message with comment

Personal and global distribution lists

Local and remote access to voice mail

Messages are time and date stamped

Operetta can store up to 24.000 messages. It has disk space for a total of 15 hours of messages. There is no allocation of message space per person. The maximum message length is 2 minutes. Mailboxes are protected by a four digit password. Messages are not automatically deleted. Each person usually has his own mailbox. However, mailboxes may be shared but only entirely. It is not possible to share only non-personal voice mail.

Usability criteria: *(What are the aims in terms of usability (e.g. usable with no training, usable with training in Y).)*

No training needed for users calling.

Organisational aspects: *(Will the system/component have to fit into some organisation or other, how (e.g. partially replace the switchboard operator, require backup for difficult or incomprehensible queries)?)*

The system will partially replace the switchboard operator. It may extend the switchboard's opening hours to 24 hours per day and it is able to answer up to 8 calls at the same time thereby reducing/removing waiting time for callers.

Customer(s): *(Who is the customer for the system/component (if any)?)*

Any organisation that wants to extend the switchboard opening hours, reduce waiting time for callers, and rationalise part of the switchboard task.

Users: *(What are the intended users of the system/component (e.g. users speaking High German, walk-up-and-use users, specialised user group X)?)*

We distinguish four user groups:

1. Any person who calls an organisation which has the system installed is a user; walk-up-and-use users; users must speak English.

2. People working in a company which has the Operetta system installed; training needed, cf. Figure 6.

3. Switchboard operator; training needed, cf. Figure 6.

4. System administrator; training needed, cf. Figure 6.

Developers: *(How many people took significant part in the development? Did that cause any significant problems (time delays, loss of information, other (explain))? Characterise each person who took part in terms of novice/intermediate/expert wrt. developing the system/component in question and in terms of relevant background (e.g., novice phonetician, skilled human factors specialist, intermediate electrical engineer).)*

Mostly engineers with a linguist or two. Little HF input during the bulk of its lifetime.

Development time: *(When was the system developed? What was the actual development time for the system/component (estimated in person/months)? Was that more or less than planned? Why?)*

In initial version of the system based on text input has been in operation since September 1992. Planned development: Longer than expected due to feedback from beta-trials and general receptivity. A product looking for a market.

Requirements and design specification evaluation: *(Were the requirements and/or design specifications themselves subjected to evaluation in some way, prior to system/component implementation? If so, how?)*

No. No time. Generally, requirements are captured from marketing surveys (what are competitors doing), from beta-trial results and from marketing hunches. There has been some effort to provide a better requirements capture method for configuring Operetta prior to installation. This used scenarios of use of different Operetta functionality.

Evaluation criteria: *(Which quantitative and qualitative performance measures should the system/component satisfy?)*

No human factors criteria. Recognition targets set.

Evaluation: *(At which stages during design and development was the system/component subjected to testing? How (describe the methodologies used, e.g. glassbox, blackbox, diagnostic, performance, adequacy, acceptance)? What were the results?)*

The real system was evaluated internally at various stages. External beta-trialling provides the bulk of development advice.

Human Factors testing was used in a limited way once the system was on the Vocalis switchboard. This test analysed what people said after the 'say a name' prompt. **The results showed most people said the right thing. The opening prompt was, changed to sound less like an answering machine which caused people to hang up. Testing was done again and showed improvement. The results were not documented.** There has been no analysis of the voicemail dialogue.

Recently, particular yes/no recognition has been studied by listening to recordings from beta-sites. Generally, dialogue changes are evaluated on an expert basis by me and then feedback is gained from beta-sites. Beta-sites in turn provide case by case input often in a closely coupled. This will continue until the system is selling in volume, i.e. until we have got it right.

Beta sites involved companies of between 10 and 70 people.

Feedback from users is in terms of questionnaires and word of mouth. Negative comments relate mostly to recognition performance, speed and accuracy. Dialogue problems are mostly related to the opening prompt which is either too long or not informative enough.

The input prompt is shown in Figure 4. Figure 5 shows what the called party hears. If Operetta thinks it knows who the caller wants but is not sure, it will ask the caller "Was that X?". If the system is totally unsure it will ask the caller to hold, then play the switchboard operator a recording of what the caller said, so that the operator can either transparently transfer him/her to the correct extension or ask the caller who s/he wants to talk to.

Mastery of the development and evaluation process: *(Of which parts of the process did the team have sufficient mastery in advance? Of which parts didn't it have such mastery?)*

Speech technologists, system integrators and algorithm specialist. Product moved out of R&D into operations but there was still a need for speech specialist participation. Team has expanded to include document write, customer care specialist and trainer.

Problems during development and evaluation: *(Were there any major problems during development and evaluation? Describe these (e.g. problems of collaboration in the team, major delays caused by ?, difficulties in satisfying specification requirement X, developer Y left the team, lack of quality of what was delivered by some in the team).)*

Currently switch interegration is a big problem - there are so many. Vocalis is less committed to Operetta so there is less 'engineering' resource available.

Development and evaluation process sketch: *(Please summarise in a couple of pages key points of development and evaluation of the system/component. To be done by the developers.)*

-new technology - what to do with it?
-new product, making a market
-how to sell it - Focus groups organised
-Beta testing
-Hasn't really sold well enough - What now

Component selection/design:

Robustness: *(How robust is the system/component? How has this been measured? What has been done to ensure robustness?)*

Operetta includes high, medium and OOV confidence levels. Error detection dialogue is used for medium confidence, OOV means a swicth to the operator. This may be hidden from the caller, i.e. the operator does the routing manually.

Maintenance: *(How easy is the system to maintain, cost estimates, etc.)*

Easy to update name DB with transcriptions. System admin. interface allows this. Full training course given to system admin. person.

Portability: *(How easily can the system/component be ported? (e.g. OS dependencies, machine dependencies).)*

Move to NT.

Modifications: *(What is required if the system is to be modified)*

Some modifications can be done by the system administrator, others must be done by Vocalis, e.g. new recordings. Operetta uses pre-recorded speech output. Most prompts are predefined. An exception is the greeting message, e.g. "Welcome to the Vocalis automated switch board....". This will be recorded for each customer to be shorter or longer as required. Additional client specific prompts may be recorded for 'out-of-business hours' announcements. This process is time consuming and it is envisaged that once Operetta moves fully out of the beta-phase it will be sold with predefined prompts (apart from the company name).

Integration with better voicemail systems is planned.

Additions, customisation: *Has a customisation of the system been attempted/carried out (e.g. modification of a part of the vocabulary, new domain/task, etc.)? Has there been an attempt to add another language? How easy is it (how much time/effort) to adapt/customise the system to a new task? Is there a strategy for resource updates (e.g. a predefined sequence of update steps to be performed if a new item is added to the lexicon or if a new grammatical description is added to the grammar)? Is there a tool to enforce that the optimal sequence of update steps is followed (e.g. a menu-driven update interface, etc.)? Comment on any peculiarities from the pov. of best practice.*

The Operetta platfrom clearly has other uses, anything which requires name identification. Currently a variant is being used as a directory enquiries system. - obviously a much larger vocabulary. The REWARD project uses an Operetta box with Spanish, Danish, Dutch variants.

Property rights: *Describe the property rights situation for the system/component.*

Documentation:
Operetta System Administrator's Guide: How to turn lines on/off; how to respond to silence; activation times for call routing and for voice mail; how to adjust activation times; how to specify business hours; types of transfer (all calls are routed via Operetta, but some options can affect the way a call is received (blind transfer, checked transfer, call screening, voice mail); how to add new names and extensions, includng speech recogniser updates. The parameters which can be configured by the system administrator are shown in Figure 1.

Parameter	**Default**
Number of rings before giving up on the Operator	8
Operator's exrension number	0
Operator's mailbox number	100
Press 0 for the Operator	on

Action to take when the Operator's phone is engaged	Try another extension
Action to take when the Operator doesn't answer	Try another extension
Alternative extension for the Operator	231
Number of rings before giving up on the night bell	30
Number of rings before giving up	5
Music on hold	on
Call screening	on
Department names	on
Announcing who a call is for	on
Message waiting indication	on
Paging configuration	Full announcement
Phone book's order	By last name
Confirmation of name recognition	off
Ask who do you want to speak to? If confirm fails	off
Say name associated with mailbox	on
Change time and date	Current time and date

Figure 1. A list of the Operetta parameters that can be configured, and their default value.

We have flow charts which show what happens in the following cases:

Extension with no voice mail; call screening on.
Extension with voice mail; call screening on.
Extension with no voice mail; call screening off.
Extension with voice mail; call screening off.
Extension with checked transfer.
Extension with blind transfer.

We have call routing flow charts which describe the call routing depending on how certain parameters (see Figure 2) are configured. And we have a speech routing dialogue flow chart.

Parameter	**Behaviour when ON**	**Behaviour when OFF**	**Default**
CONFIRM NAME RECOG	Always asks "Was that X?" when positive or middle confidence.	Only asks "Was that X?" when middle confidence.	OFF
CONFIRM SECOND CHANCE	After the caller has replied "No" to "Was that X?" asks them to repeat the name of the person they want.	After the caller has replied "No" to "Was that X?" tells the caller they will be transferred by the operator.	OFF
REPEAT NAME RECOG WITH OOV	If the confidence is low (OOV) will ask the caller to repeat the name of the person they want.	If the confidence is low (OOV) will tell the caller to hold and transfer transparently to the operator.	OFF
CALL SCREENING	After Operetta has	Does not ask the caller for	ON

	ascertained the name of the called party, asks the caller for their name.	their name, does not announce name to called party.	

Figure 2. Parameters and behaviour when on/off.

The flow charts cover the combinations (call screening is assumed to be on) shown in Figure 3.

Parameter	ON/OFF							
CONFIRM NAME RECOG	ON	ON	ON	ON	OFF	OFF	OFF	OFF
CONFIRM SECOND CHANCE	ON	OFF	ON	OFF	ON	OFF	ON	OFF
REPEAT NAME RECOG WITH OOV	ON	ON	OFF	OFF	ON	ON	OFF	OFF
CALL SCREENING	ON	ON	ON	ON	ON	ON	ON	ON

Figure 3. Parameter combinations covered by flow charts.

Good morning/afternoon/evening you're speaking to the [company] automatic switchboard. After the tone please clearly say [the name of the department or] the first and last name of the person you want and I will connect you. If you prefer to speak to the operator, please stay on the line.

Figure 4. Standard form of the Operetta greeting. It can be changed to suit the customer's circumstances.

A call for [called party] from [caller].

Figure 5. If the system is configured to ask for the caller's name the called party hears the sentence shown in the figure. Otherwise 'from [caller]' is left out. The words in square brackets are replaced by a recording of the caller's voice. The call announcement allows the called party to make call screening, i.e. to interrupt a call announcement and make it appear to the caller as if the phone was not answered.

Session	Duration	Max attendees
System administrator	2-3 hours	3
Switchboard operator	1_-2 hours	5
User (call routing and voice mail)	1-1_ hours	10
User (call routing)	30-40 minutes	10

Figure 6. Training sessions needed.

Configuration	Description
Front end	All incoming calls go straight to Operetta.

Overflow	Operetta picks up calls when the switchboard operator is busy or does not answer.
Out of hours	Operetta answers calls after the switchboard operator has gone home.

Figure 7. Different configurations of Operetta.