ETSI
TECHNICAL
REPORT

ETR 096

August 1993

Source: ETSI TC-HF

Reference: DTR/HF-1002

ICS: 33.020, 33.040.40

Key words: Human factors, PBI, computer services

Human Factors (HF);
Phone Based Interfaces (PBI)
Human factors guidelines for the design of minimum phone based user interface to computer services

ETSI

European Telecommunications Standards Institute

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis CEDEX - FRANCE
Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE
X.400: c=fr, a=atlas, p=etsi, s=secretariat - **Internet:** secretariat@etsi.fr

Tel.: +33 92 94 42 00 - Fax: +33 93 65 47 16

Contents

Blank page

Foreword

This ETSI Technical Report (ETR) has been produced by the Human Factors (HF) Technical Committee of the European Telecommunications Standards Institute (ETSI).

ETRs are informative documents resulting from ETSI studies which are not appropriate for European Telecommunication Standard (ETS) or Interim European Telecommunication Standard (I-ETS) status. An ETR may be used to publish material which is either of an informative nature, relating to the use or the application of ETSs or I-ETSs, or which is immature and not yet suitable for formal adoption as an ETS or an I-ETS.

The purpose of this ETR is to provide guidelines for services providers, user interface designers, application writers, and human factors engineers, etc., in the development of a Phone Based Interfaces (PBI). The guidelines are intended to facilitate usage of new services by allowing users to transfer knowledge learnt in one application, and apply it in the use of another. What should not be assumed from these guidelines, is a wish to constrain development based on new and evolving technology.

A significant number of the current 400 million telephones will become Dual Tone Multi Frequency (DTMF) based, as new digital exchanges are used in networks. This will provide a basis for customers to access a wide range of services such as:

- voice mail;
- banking;
- information retrieval;
- shopping.

The challenge of application designers is to allow people who currently use these services through a Screen Based Interface (SBI), to achieve a similar degree of access with Phone Based Interface (PBI). The additional advantage is that Phone Based Interfaces (PBI) are still at an early stage of market development. The opportunity exists for customers to be presented with varied services in a coherent framework.

Blank page

1 Scope

This ETR is only concerned with Phone Based Interfaces (PBIs) that assume DTMF input and speech output. Guidelines for the use of speech recognition as a means of user input are not considered as being within the scope of this ETR. The intention of this ETR is to achieve consistency across a number of core functions and facilities within PBIs independent of either invocation by the user or application by different service providers. An example of the use of these guidelines to an application area is given.

The guidelines begin to address the usability of PBIs. To ensure that products and services meet the requirements of a particular market, usability testing and interactive design should be part of the product development process.

It is to be remembered that in order to create a voice service, "expertise in ergonomics is often necessary, just as possessing a medical encyclopaedia does not suffice to cure a sick person" (as stated in "Ergonomic criteria for the evaluation and design of user interfaces" [1]).

2 References

For the purposes of this ETR, the following references apply:

[1] D. L. Scapin in Actes du XXVI Congrès de la SELF, Montreal, Canada, (1990): "Ergonomic criteria for the evaluation and design of user interfaces".

[2] CCITT Recommendation E.183 (1988): "Guiding principles for telephone announcements".

[3] CCITT Recommendations P.80 to P.84 (1988), Vol. 5: "Telephone transmission quality - Subjective opinion tests".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of this ETR, the definitions and explanations of key terms and concepts used apply.

Application: set of files, programmes and processing procedures developed and organised to carry out a task automatically.

Audiofax: data transmission service combining voice service and facsimile.

Audiotex: several definitions exist: audiotex is either the generic term for all services, or simply inquiry access services.

Audiovideotex: data transmission service combining voice service and Videotex.

Beeps: short audio warning signal produced by the system.

Call transmission: ability of the system to make outgoing calls to a user.

Coded speech: speech treated by coding algorithm.

Coding: an operation transforming electrical signals from acoustic transducers into binary codes which can be transmitted or saved in mass storage devices.

Concatenation: assembly mechanism of elementary messages (fragments of sentences, words or syllables) making up a voice message.

Cut-through: system's capability to follow commands without waiting for their introduction or end thereof.

Dial pulsing: the means of signalling generated by telephone rotary dial and some keyboards. Such signalling uses regular momentary interruptions at the sending end of the telephone line.

Dual Tone Multi-Frequency (DTMF): multifrequency codes generated by touch-tone pads.

Earcon (jingle): message, or fragment of message, generally based on repeating a musical theme and designed to trigger recognition.

Ergonomics: adapting working conditions to man: implementing human-oriented scientific knowledge, necessary for developing tools, machine and appliances that can be used by the greatest number of people with maximum comfort, security, and efficiency.

Escape (Key): key allowing access to a second virtual keyboard allotting a different role to the keys.

Incoming calls: ability of system to take user calls.

Interactivity: principle which enables the user and system to exchange commands and messages.

Linear (Dialogue): dialogue structure which, unlike tree-form dialogue, does not offer menu choice to the user. Data and operations are offered in predetermined order.

Navigation: the means of moving between a number of items, or even between a number of application branches.

Pause detection: operation which tells system that the user has not spoken for "a while".

Prosody: modulation of length (sounds and pauses), of register and intonation within the language which make up the diction character.

Speech detection: operation which tells systems that the user has begun to speak.

Speech recognition: technique for detecting and interpreting voice commands from user.

Text-to-speech: technique enabling the machine to restore data, supplied in text form, in voice form.

Tree form (Dialogue): dialogue structure based on consecutive menus.

Voice command: any means of giving system voice commands (see Speech recognition and Speech detection).

Voice messaging: deferred communication system enabling users to receive and transmit audio messages by telephone.

3.2 Abbreviations

For the purposes of this ETR, the following abbreviations apply:

CI	Command Identifier
DTMF	Dual Tone Multi-Frequency
PBI	Phone Based Interface
SBI	Screen Based Interface

4 Principles

The purpose of this Clause is to harmonise voice dialogue using telephone tone keys. Standard procedures are proposed, as well as the assignment, where possible, of each function to a telephone key. Consistency reinforces learning and as a result, makes using voice services easier. Eight general principles are defined here:

Commands Defined functions should be invoked with the same command and, where possible, in different contexts the same command should invoke similar functions.

Commands can be invoked with more than one key press, menu items should be considered as single key entries, where possible.

Help Help should always be available and it should be context sensitive.

User control The user should always be in control, and should not be constrained by the system or service.

NOTE 1:The user should have the ability to "cut through" a prompt and to dial ahead as well as offering BACK STEP function and a return to a main menu function.

Cancelling Enable the user to cancel data entry, or actions that a user cannot recover from.

Delimiter Where data is of unpredictable length, then either a # or time out should be used as a delimiter.

Confirmation Where an action is deemed to be irreversible, always present the user with a system message that asks the user for confirmation and ensure that the correct confirmation message is selected.

Security The log-on procedure is an implementation issue, and is outside of the scope of this ETR.

Feedback The system, or service should confirm user action and input. Where invalid keys or commands are selected feedback should be given to the user.

NOTE 2: The feedback should be on both a key-by-key basis, and when an action has been completed.

5 Human factors guidelines

The following guidelines are presented in alphabetical order, and encompass good human factors practice.

5.1 User context

5.1.1 Commands

Commands and entries are made on a telephone keypad and where commands are accessed as part of a menu structure, then the normal rules of menu structure apply. Using letters mnemonically to indicate a function or command should be avoided as this can rapidly lead to ambiguities or to badly adapted word choices.

Where commands are accessed directly from the keypad, outside of the menu structure, then the following applies.

Basic interaction with the service

These should be provided at all times. Command Identifier (CI) [*] is to be used when there are more than six items in a menu:

- help [*] [0];
- language for further study;
- back step for further study;
- main menu for further study;
- cancel [*] [*].

It is generally recommended that commands associated with PBI are assigned to the keypad such that the left hand column of the keypad indicates regressive steps, the right hand column indicates progressive steps, and the centre column indicates the status of the system.

Specific commands: example of voice mail:

- rewind;
- fast forward;
- pause;
- stop;
- delete;
- confirm;
- playback/repeat;
- edit/continue.

5.1.2 Data entry

Decimal numeric data entry is preferred to alphabetic data entry. However, some PBI applications require alphabetic or alpha-numeric entries.

EXAMPLE 1: The application requires one input to define a key and a subsequent input to define the position of a character on that key. "ETSI" would be coded: 32 81 73 43.

EXAMPLE 2: The application requires a single input to define a key which represents a number of characters. The application can search for appropriate characters by matching each keystroke in turn with previous keystrokes, searching for a matching string in the database. Since the same keystroke may create different valid strings, the application should confirm the selection by voice feedback. "ETSI" would in this case be coded only by: 3874.

A method may be required for entering alpha numeric strings of varying length. It may be necessary to signal to the service the mode in which the user is signalling (i.e. alpha or numeric).

EXAMPLE 3: The application requires the entry of both alphabetic and numeric data. In this case the delimiter key [#] should be used to define the boundary between alphabetic and numeric input (for a customer password, some credit card numbers, etc.).

It is unnecessary to feedback the individual characters that have just been entered to the user. However, confirmation of input should be given where modification of data input or termination of input has occurred.

Entries of unpredictable length need a delimiter. For consistency, the delimiter [#] can be used when there are entries of unpredictable and predictable length in the same application.

Cancellation of the current entry should be enabled by a single command. Offering character-by-character correction is tedious.

For entry of long data strings, it may be useful to break up the request. User inactivity at the beginning of, and during entry, will be considered as a time-out.

5.1.3 Menu

(See cut through, time outs).

Menu options should be presented to the user such that commonly selected options require the shortest available key sequence.

Where possible, menu options should be assigned in a logical order starting on key [1].

At all times the difference between menu options available the dialogue and commands available through the keypad should be distinguished.

Where keys have pre-assigned functions on the keypad, as in an application, then care shall be taken in assigning the remaining additional keys to menu options. In general, no more than six options should be assigned to a menu that have no pre-assigned functions to a keypad. In instances where there is a necessity to have a user access generic function such as help, both implementations outlined above would then use the [*] key to "escape" to an underlying functionality assigned to the keypad.

If a menu is not given, then there should be some means of offering the user a prompt, either explicitly or implicitly.

5.1.4 Prompt messages

The interaction is directed by means of a dialogue using voice messages, and these messages can be identified thanks to their special purpose:

welcome message (e.g. " Automatic Weather Service at your disposal");

menu message enables first choice to be made and are presented in the form (e.g. "menu, entry request etc.");

general help giving all-purpose information about the service, (e.g. "This service is free of charge and always available");

confirmation messages (e.g. "Your reservation has been accepted");

help messages (e.g. "The article referee contains 5 figures, please...);

context prompts menu replayed explaining choices after an idle period (e.g. "For the soccer press 1, for the tennis, press 2";

waiting messages (e.g." Wait a moment while we check the article");

status messages (e.g. "You have received x number of messages");

service message (e.g. "The reservation service is not available");

information messages (e.g. "Weather today will be very sunny and warm");

application specific (e.g. in the voice mail to change voice between gender).

It is not advisable to allow the user to interrupt an important message such as one referring to service unavailability. The user should receive very explicit messages. It is advisable to choose terms which are highly distinctive and unambiguous, which are familiar to users so that everyone will understand them. In particular, a function's title should describe the function exactly.

As soon as users enter a new mode, they should immediately be made aware of this by means of an appropriate message. If this new mode requires action by users, a message should tell them which action is required.

Message structure (from CCITT Recommendation E. 183 [2])

In a menu item, the anticipated result should be indicated before the command.

It is better to say: "To listen to your messages, press key x"

than: "Press key x to listen to your messages".

Better understanding is achieved by using straightforward sentences expressed positively rather than by using negative sentences in the passive voice. However, it is sometimes helpful to use the negative to emphasise a point.

If a choice of several words or expressions is available to convey the same idea, the same word or expression should be used consistently.

5.1.5 Soft switch between pulsing and DTMF dialling

The user may have a telephone that is capable of supporting dual signalling. It is advantageous to offer users a means by which such a switch can occur. It is recommended that the [*] key be used for this purpose.

5.1.6 Welcome message

The welcome phase is very important because it identifies the service and gives the user an initial "impression" of it, and should give the user all the data he needs to identify the service.
It is important, in the welcome message to make the user understand that he is connected to a service.

EXAMPLE: "Hello and welcome to X's voice service...".

5.2 System context

5.2.1 Cut-through

The application should take the user's level of experience into account. The experienced user needs to have commands enabling quicker access to the required data. The opportunity to cut-through each command and rapid system response will satisfy frequent and experienced users.

When the application can identify the user accessing the service, it can offer direct access to a particular part of the dialogue. This direct access can depend on the user's habits or on a user linked parameter (on-going command, incoming message, etc.).

Care should be taken if allowing cut-through on prompt, help or feedback messages.

Pressing a key corresponding to a valid command should cause the current message to stop immediately. This command should be registered and should permit transfer to the next appropriate point in the dialogue. The user should be clearly notified when this is not possible.

The system should react to DTMF cut-through as soon as possible. In some cases, it may be appropriate to allow user to interrupt a help message if they are familiar with the system and if the message is long.

Shortcut commands allow direct access to a part of the dialogue. When these are implemented, users should be provided with a back-up pocket memo to inform them of the commands available as well as giving them an overall picture of the service.

On the whole, global deletions should be avoided. Where such deletions are necessary, such as when an error is made during the input of long strings of data, a command is available. Care should be taken to insure that it is not possible to invoke that command by mistake.

Any DTMF input during an uninterruptable message should be ignored by the system.

5.2.2 Dialogue structure

(See message structure, subclause 5.1.4).

The dialogue structure can be very different according to the services being implemented (information, consultation, transaction, etc.).

Linear or tree-form dialogues will be chosen according to the task. A linear dialogue will have to unfold logically according to the task (natural or customary order), in response to the user input. A tree-form dialogue (choice in menu) will need to be structured coherently, according to the service offered.

In the case of tree-form dialogues, it is desirable that the application of the tree-form levels are structured in a logical manner. Using testing and iterative design are crucial to this activity. Within a tree form dialogue, the number of options presented to users should be minimised and the command key [*] be used to increase the number of options.

The choices proposed on one level do not vary during use, nor when moving from one use to another.

It is advisable that where possible a large number of dialogue levels should be avoided.

5.2.3 Feedback

The user needs to know if their actions have been accepted and what the effects are.

Different tones used in a consistent manner can indicate various states. For example, tones can be used to attract the user's attention, or to indicate that the selected option is invalid. or, in the case of an voice mail, a tone could be use to indicate to a calling user that a message can be left.

Sequence is important when displaying menu choice. For example, the most neutral or the most useful format can be selected.

It is sometimes quicker to give the user an item of information directly rather than propose a menu.

As soon as users have carried out a valid action, they should have it confirmed or hear a message. They should then either hear the new prompt, if this is correct, or hear the information message corresponding to choice.

Certain phrases ("Thank you", etc.) tell the user that his data entry or command has been registered. Over use of these features can lead to slow or irritating dialogues.

5.2.4 Help

5.2.4.1 Error

Any error made by the user, which cannot be corrected by the dialogue, needs to be consistently and immediately pointed out to him. When the user has made a mistake, the next commands should not generate any others, so avoiding chain errors. For example, calling up a non-existent function or one not recognised by the application should be indicated by a help message.

The information supplied by the message should be relevant so that the user can solve the problem (what is wrong, what to do next, etc.). If he repeats the error the message can be rephrased in a different form.

The help message should be short, explicit, non-critical and adapt itself to the type of error in question. It is not advisable to allow the user to interrupt a help message or a service unavailability message.

Characters entered after the error-provoking one should not be taken into account. A beep transmission can tell the user that these are invalid. Also, an application may decide to ignore invalid commands or invalidate them by a single beep.

5.2.4.2 Choice

The user needs to be able to have the available options repeated. Also an experienced user may wish to have a menu of option as the first level in the help menu in an application.

5.2.4.3 Prompts

See also Error (subclause 5.2.4.1) and Time outs (subclause 5.2.7).

These can be used to enable the user to make a choice "e.g. a menu". Consideration should be given as to when prompts should be used. It may be considered applicable to give prompts following a period of user inactivity or a first response to a user request for help. The user should under no circumstances be left alone.

Upon further user request a more detailed context prompt should be available as soon as possible, giving supplementary information and explaining precisely each choice's various possibilities.

When a context prompt has been given, the dialogue should revert to the place where the user encountered the prompt.

Prompt may be a help message to indicate available commands or to indicate user position in the application tree.

5.2.5 Navigability

The system should be structured such that the users should always know what is expected from them, and what the effect of their input will be.

The user should be given the opportunity of backstepping, fast forwarding and moving from one part of the dialogue to another. For example:

- when listening to a message (voice messaging), allow the user to pass directly into the recording branch to make a reply;

- when consulting a timetable, allow the user to go directly into the reservation branch.

Users, and especially occasional users need to always know where they are. This guideline applies to main menu and also to subsequent menus. The main menu is of paramount importance in this matter. It should immediately be recognised as such by the user and never be confused with another.

It is desirable to minimise the number of stages which the user should go through to reach an objective.

The user should be able to interrupt or cancel an on-going transaction, as well as its consequences.

A BACKSTEP function should be provided. A return to main menu function should be provided, but bearing in mind the relatively low number of hierarchical menus, the BACKSTEP key may be tapped several times to access the first-level menu.

It should be noted that the BACKSTEP function can be used repeatedly to return to the main menu.

5.2.6 Response times

The fixed nature of a system's response times is an important feature and should be minimised. Comfort and effectiveness depend on it since the user should be able to predict the system's behaviour. In addition, time-outs should be controlled so that the user is never left in a vacuum.

Where possible, and where the system detects extended response times then a waiting message could be given to the user.

A minimum of 0,5 seconds in all cases, under normal operating conditions, should be implemented. This is to allow use of hand held DTMF signalling devices.

The response time between the user's command and the system response should not exceed 1 second in 95% of cases, given normal operating conditions.

If a response time exceeds 1 second, the user should be informed that the transaction is proceeding by means of a comfort message.

5.2.7 Time outs

In the case of prolonged user inactivity a message, or a default dialogue, should be provided.

When the user does not enter data for a certain period of time (this time depends upon the application), the system should react by prompting the user (between 5 seconds and 20 seconds).

After waiting for a specified time (4 seconds to 8 seconds) for user input, either the last prompt should be repeated, or an in context help message should be presented. This latter item may be further menu options.

When inactivity occurs during the entry of a consecutive digits, after a time-out of 1 second to 4 seconds the user should be reminded how to complete the entry or how to delete it. It may even be necessary for the system to repeat the characters or the digits which have been entered.

The system should be developed so that time-outs may be changed.

5.2.8 Types of users

(See Cut-through, subclause 5.2.1).

The user can also be invited to choose his preferred dialogue by an explicit question during the welcome to service phase.

When the application can identify the user accessing the service more detailed direction messages can be suggested, particularly on initial access to the service.

The user should be guided according to his experience. The application should take different types of users into account. The system should take into account the difference between the "experienced" (or frequent) user who need only be guided and the "beginner" (or occasional) user who needs to be helped with the task in hand. The designation "experienced" or "beginner" not only depends on how practised the user is and how frequently he uses the application, but also on such factors as their familiarity with the task and with using an analogue service.

The experienced user's requirements are covered by concise but complete direction messages while the beginner's are met by the availability of more detailed context prompts (on explicit request or after a period of inactivity).

Annex A (informative): Screen Based Interface versus Phone Based Interface

A.1 Some aspects provided from comparison with Screen Based Interfaces

Usually the use of the telephone is combined with the use of a loudspeaker while the use of a computer is combined with the use of a screen. Here we use a telephone as an interface for the input/ouput of a computer which also supports the dialogue and the contents of the services.

Usually computer services are screen based with visual information. Minimum phone-based computer services can provide only auditory information by use of the telephone earpiece.

Although PBIs are widely accepted, they differ from SBIs in two different ways:

a) the information output is auditory and serial;

b) the user does not have the use of a pointer.

The result of these differences is that SBIs and PBIs have different attributes.

On a SBI, the information displayed is parallel and visual. On the PBI, it is serial and auditory. Therefore, the amount of information displayed is quite different in a same period of time. Input means are restricted to the telephone keypad and eventually spoken words.

External memory elements are typically very poor, and less support for short term memory is provided. Generally, the dialogue can be predictable on the screen and only interactive or reactive on the phone.

The different attributes of SBI and PBI are given in the following table A.1 for comparison. The aim of this table is to show the main differences in input/ouput between a PBI and a SBI. Theses differences affect both system design and operation.

Table A.1: Summary of differences between PBI and SBI

Factor	Screen based	Phone based
Output means	Parallel - Visual - written text - icons - graphics - cursor display	Serial - Auditory - spoken messages - signals - non speech audio
Input means	full keyboard - function keys - cursor control - options with pointing devices	12 keys keypad - spoken words
External memory elements	Past.....Log Present...Cursor/screen Future....Remaining option Semantic.....Hypertext	Content of current message needs to be adapted past/present by special commands.
Dialogue and navigational flow	User needs to find the way through: - instructions; - cues on screen.	Interface leads the way through: - messages; - input requests; - responses.
Short term memory aids	Continual over time on screen.	Degrades over time.
Limits on amount of information in one presentation unit	Screen size Windowing	Language power Short term memory

Annex B (informative): Sound quality

This annex covers:

- general recommendations for making recordings;
- advice, remarks and suggestions relating to these recommendations.

B.1 Sound quality

The sound recordings should be well made. Sound messages are transmitted by means of the public telephone network; therefore any degradation of the recordings accumulates with the limitations intrinsic to network transmission. Sound quality is defined by the following principal factors:

- the sound level which should be equalised and adjusted so that messages may easily be heard anywhere on the public network;

- the band of audio frequencies transmitted should be within the 300 Hz - 3 400 Hz passband to ensure intelligibility;

- the distortion on frequencies transmitted should be kept low to avoid deformation and additional noises which makes listening to messages unpleasant to the ear;

- the speech signal dynamics should be taken into account to ensure against intelligibility loss where the signal is weak;

- the background noise should be restricted to ensure against reduced quality voice messages.

The telephone pass-band particularly affects the quality of musical signals transmitted. Musical signals not unduly affected by restricted bandwidth should be chosen, that is musical signals with frequency spectra mainly concentrated in frequencies lower than 3 400 Hz. Moreover, musical signals whose main harmonies emulate some DTMF signalling tones should be avoided.

One solution is to filter the frequencies used by DTMF tones when the messages are being recorded. Recording in a studio with the help of professional technicians produces quality sound-recordings.

B.2 Audio coding choice

Sound recordings are digitalised, coded and stored in mass storage devices. Using medium and low bit rate compression algorithms reduces the storage capacity required, but also involves a more or less marked reduction in quality. This degradation is accumulated with that caused by the public network.

The sound quality of compression algorithms should be evaluated with the help of subjective quality evaluation tests compliant with CCITT Recommendations P.80 to P.84 [3].

B.3 Message concatenation

Presenting variable data in a message (such as figures, dates, times) requires placing concatenation of the fixed parts of messages ("support" phrases) and the variable parts ("parametered" phrases). The set of "support" and "parametered" utterances needing to be stored for a given application is called "elementary messages". To guarantee that the resulting messages sound natural, particular attention should be paid to the following:

- all the elementary messages should be recorded by the same speaker, if possible during a single recording session (to guarantee voice homogeneity);

- respect for the natural links which occur between certain words;

- respect for the pause length which occurs naturally between certain word groups (voice equivalent to punctuation);

- having elementary messages recorded with different intonations if they are used in different positions within phrases.

For example, to restore a number between 0 and 1 000 correctly by concatenation of figures and some numbers, several recordings of these figures and numbers should be available depending on their position in the resulting utterance (isolated figure, median position, final position, etc. ...).

Using sophisticated sound editors helps applications developers produce quality recordings. These sound editors should in particular facilitate:

- handling a large base of elementary messages;

- displaying the acoustic waveform of elementary messages;

- linking elementary messages, listening to the resulting message and displaying it;

- controlling and adjusting sound levels of elementary messages;

- controlling and adjusting pauses at the beginning and end of elementary messages;

- better than 10 msec to 20 msec concatenation accuracy may be needed for concatenation.

B.4 Speaker choice

Speakers who are used to talking into a microphone, know how to modulate their voices and use the right pitch while adjusting to the content of voice services broadcast. Coding algorithms give different results depending on the speaker's voice. Therefore, it is wise to carry out preliminary coding-decoding tests. Using professional speakers (actors, journalists...) often saves time and meets to these recommendations.

As a general rule it is better not to use voice changes unless they have informative value (e.g. distinguishing between two types of messages) or "structural" value (e.g. in the same type of information message, having a "radio style" exchange between a female and a male speaker). In all cases, juxtaposing similar voices should be avoided.

It is better to have the same voice record all the direction messages from the same service. In the case of a variable item of information (e.g. recording a number), needing concatenation, all the recordings should be made by the same voice. When this is impossible (e.g. user's name recorded by the user himself, in a voice messaging service), a pause should be noted between a message and juxtaposed data.

B.5 Style

The message should be carefully and accurately composed before being read, applying spoken rather than written language. The language used should be up-to-date, faultless and not stilted. Style and diction rhythm should be adjusted to the service, the user, the task in hand and the frequency of use. Sentence and paragraph structure should be respected. Do not hesitate to use a natural pitch, in particular for direction messages.

For example, using humour to point out errors might not be appreciated.

B.6 Text to speech

The recent progress made in text-to-speech means that its use in general public applications can now be considered. There are two text-to-speech operational modes:

a) a real-time mode where the voice message is sent directly to the user from a text file. This mode is essential for services which cannot use prestored voice messages as the content changes with time and is unpredictable (e.g. accessing text messaging, or broadcasting proper nouns). The obvious advantage of the service available through this technique is that it can make up for the deficiencies of synthesised speech used at present (as compared to pre-recorded speech);

b) a deferred-time mode where the voice message is previously formed from the text, in order to be saved on the voice service (the message is subsequently broadcast to the user like any other voice message). At the cost of a reduction in quality, this mode permits the breaking away from using

speakers to record voice prompts. In this case, it is strongly recommended that specific voice editors be employed in association with the synthesis systems in use, thereby enabling some residual synthesis deficiencies to be corrected manually (phonetisation, prosody...).

Synthesis can also be used when developing voice message content before going on to have them definitively recorded by live speakers.

Avoid using text-to-speech and coded speech in the same sentence; when this is unavoidable, a pause should be noted between the pre-recorded speech and the synthesised speech.

B.7 Music

As in all sound media, music can be used to identify, personalise and add "colour". Themes should therefore be carefully chosen and should not muffle content by being too obtrusive. Jingles may be used for punctuation. On the other hand, background and cover sounds, used sparingly, should be mixed and modulated so as not to drown voice messages. In all cases, it is important as a preliminary exercise to listen to the result obtained over the telephone network.

History

Document history	
August 1993	First Edition
February 1996	Converted into Adobe Acrobat Portable Document Format (PDF)