

Speech filing—An office system for principals

by J. D. Gould
S. J. Boies

Business people spend most of their time communicating, or attempting to communicate, with others. We briefly describe our ideas about these communication activities and their resulting problems, and then discuss an experimental tool we developed to help business people solve some of their communication problems. This tool, called the Speech Filing System, allows users to send messages to anybody in the world and receive messages from anybody in the world. The system offers powerful editing, filing, retrieval, and message distribution and control functions, using pushbutton telephones as the terminals.

There is currently much interest in creating new work tools for business people, or so-called office principals, e.g., "white-collar" professionals, managers, executives, salesmen, etc. This interest in principal-support tools is stimulated by increasing white-collar costs, the need for greater principal productivity, potential decreases in secretarial services, and the absence of computerized tools for principals today.

Since 1973, we have been doing research on principal-support tools. This paper first describes the way we structured the problems to be solved and the experimental system we designed to solve them. The main part of the paper then describes what a principal can do with this new system, or tool.

The tool is called the Speech Filing System (SFS). It was developed at IBM Research in Yorktown Heights, New York, from 1973 to 1975. In 1975 principals began to use SFS in their own work, and informal evaluation with users began.^{1,2} At the same time, formal human factors laboratory experiments related to SFS were underway.^{3,4} For the next six years, about 750 IBM principals around the world used SFS in their daily work. This six-year period allowed us to con-

centrate on improving the user interface (including training, documentation, and help facilities) through a series of experiments, field studies, and major changes and the subsequent empirical evaluations of the interface.⁵ In September 1981, IBM announced that the IBM Audio Distribution System (ADS), a direct outgrowth of SFS, would be available as a commercial product. The first customer installation was in February 1982. Other companies, including AT&T, Electronic Communications Systems (ECS), and Wang, have announced audio systems that have some functions in common with ADS.

Many of the audio functions described here are implemented in ADS. A description of the functions contained in ADS is contained in the "subscriber's guide."⁶ We also describe some functions that we have designed or incorporated in research prototypes only, including message composition and transmission of nonaudio messages. The user interfaces for these functions and for ADS were jointly developed to make them compatible.

Whereas this paper provides a functional description of SFS, a related human factors paper⁵ describes the design challenges and how we addressed them by drawing upon the behavioral literature to understand what principals actually do and by the empirical research we carried out through laboratory experiments, simulations, and prototype studies.

©Copyright 1984 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

The problem

Communicating with other people is the main activity of principals, as well as of most other people. Communicating takes a lot of time. Studies show principals spend about 25 percent of their time in *noninteractive* communication, i.e., reading and writing,⁷⁻⁹ and they spend over 40 percent of their time in *interactive* communication, i.e., face-to-face conversations, meetings, and telephone conversations.⁷ If these percentages do not seem intuitively obvious, it may be because principals overestimate the amount of time they spend in noninteractive communication (42 percent estimated versus 26 percent actual) and underestimate the amount of time they spend in interactive communication (28 percent estimated versus 42 percent actual).⁷ Interactive communication is the source of major frustrations for principals, e.g., inability to get the desired person on the telephone, inability to get to the point in a conversation, and inability to be understood.

Noninteractive communications provide a permanent record and have traditional distribution methods (e.g., U. S. Mail, public libraries). But these require very expensive, time-consuming typing and editing. Multiple copies are often required, and distribution is very slow. The increasing use of electronic mail and electronic message systems¹⁰ is still pretty much limited to people using computer systems and terminals which are cost-justified for other purposes.

Interactive communications usually have speed advantages over noninteractive communications. Most interactive communication makes use of voice, and the intonation of speech carries information. But interactive communication has become increasingly difficult, in part as a consequence of having to locate a person in both space and time. Principals are unavailable over half the time because of being on the telephone, out of their offices, or in meetings. Principals make 15 or so telephone calls each day, completing only half of them to the parties they want to talk with.¹¹ They almost never leave a content message with a secretary but rather a message to "call me back." And, of course, when the call-back is made, the same statistics apply. (Incidentally, secretaries spend ten percent of their time handling these telephone messages.⁸) Principals often volunteer that this telephone roulette is their main pet peeve at work. Because people are frequently away from their desks and may be traveling, the caller may not know where to reach the called person. Thus, locating a person at a particular time has become an increasingly difficult task.

The problem, then, is to help principals with their communications needs. The solution lies in understanding these needs and developing a tool that addresses them.

Our solution

Required behavioral characteristics. From our behavioral analyses of principals' communication patterns, we decided that the solution involved many critical characteristics:

1. It must address principals' *communication* problems because communicating is the main activity of principals.
2. It must be a communication *system* that many principals will use.
3. It must be able to be used *directly* by a principal. Direct use is required to avoid the cost and potential errors of a human intermediary specialist and to identify displaceable costs that customers may need to justify the expense of buying the system.
4. It must be able to be used *anywhere* because principals work in many places—their offices, other people's offices, conference rooms.
5. It must be *sufficiently powerful* to be *useful* in addressing the necessary broad range of principals' communication needs.
6. It must communicate with *people*, not things (e.g., machines and traditional computer data bases) because principals communicate mainly with people.
7. It must handle *soft* information (opinions, attitudes, predictions, emotions, etc.) as well as hard information (facts, numbers, propositions, etc.) because principals almost always take into account soft information in making decisions.
8. It must handle information with a *short* lifetime as well as information with a *long* lifetime. Much office information has a short, but critical, lifetime, often of less duration than it typically takes to update most data bases, e.g., "the meeting has been postponed for 15 minutes" or "we have to go now."
9. It must *automatically* update itself.
10. It must be *easy to learn* and *easy to use*. This is the *sine qua non*. Principals say they do not want to read documentation or spend much time being trained. Some critics find it hard to reconcile this position with the fact that principals have spent many years in school learning from books and going to classes. The key insight has to do with the difference between learning to *do*

something, and learning *about* something. People learn to *do* things by *doing* them (walking, riding a bike, playing a game, speaking a foreign language, aerobics, typing, swimming, giving talks, cooking, successfully interacting with people, etc.), not by reading about how to do them.

These criteria led to severe challenges at the beginning of our work in the early 1970s. Office principals did not use computer typewriter or video terminals or computer systems, and there was a general belief that they would not. Computer systems communicated with things, not people, and contained hard,

SFS contains the best features of both noninteractive and interactive communication systems.

not soft, information. Computer terminals were used only by specialists. Data bases were updated periodically, and only by specialists. Computer systems required extensive training to learn. It was hard then, just as it is now almost ten years later, to predict the rate at which these facts will change. However, we believed then—and now—that the manner in which they will change depends upon providing systems that are more useful and easy to learn and use.

Summary of the system. SFS contains the best features of both noninteractive and interactive communication systems. Figure 1 shows a conceptual schematic of SFS. As shown, users can send and receive messages using standard pushbutton telephones. Messages are stored in a computer (an IBM Series/1 computer in ADS), which is connected to the telephone network and can be dialed from anywhere. Starting from the lower left of the figure and going clockwise, users can get or send messages from telephone booths, from their office, from conference rooms (where several people can use SFS at once), or from motels and home. A user's secretary can, via the secretary password, assist the principal.

SFS allows a person to *compose* (edit, review, insert, delete, annotate, and format) messages, *distribute*

these to any person in the world, and *receive* messages from any person in the world. (There is no distinction between messages, memos, documents, mail, notes, letters, etc. in SFS.) To do all this, SFS uses a pushbutton telephone (or its equivalent) as the terminal. The telephone was selected because it is the most ubiquitous terminal in the world, and users can thus send messages at almost any time and from almost any place. All principals have telephones in their offices and at home (but need not use their own telephones). Telephones are also in most other places where principals work, or else nearby, e.g., phone booths. With SFS, unlike with most other computer message systems, users need not seek out a special-purpose terminal. A principal gives commands to SFS via the pushbuttons on the telephone. Audio messages (as well as other media documents) are stored and modified in digital form but can be received or played out in analog or digital form.

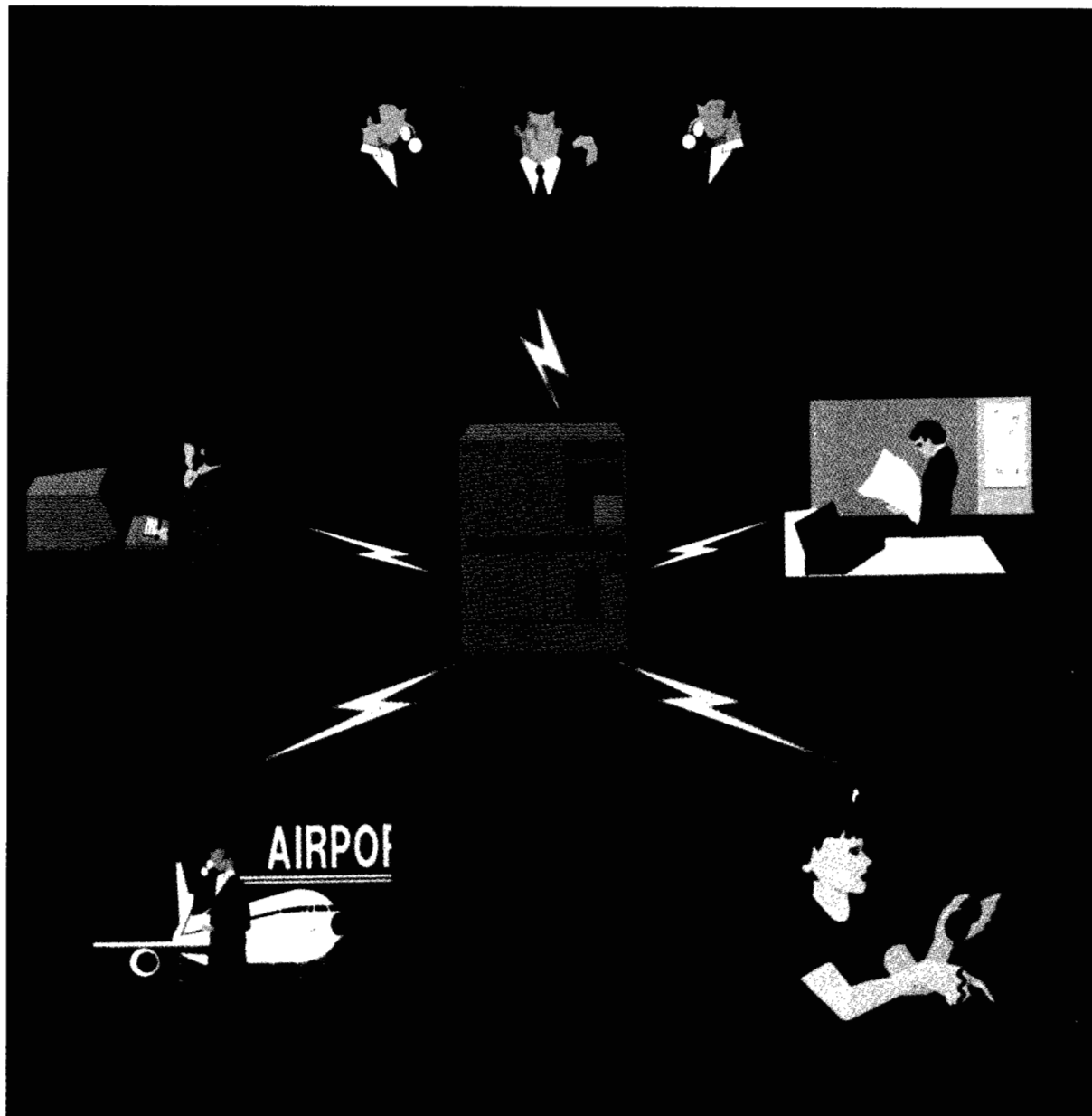
Key results. Our research in the last six years shows that principals use SFS directly (from our interviews and observational data), use it at all hours of the day (from interviews, observations, and usage analyses), find it is easy to learn (ADS customers tell us that new users often learn it with no training), find it very useful in their work (from self-reports, interviews, and surveys), and send content messages (from observations and self-reports). If it is unavailable for even a short period of time, they are disturbed (self-reports and survey data). These results are elaborated upon elsewhere.^{5,12} SFS is a cost-effective way of communicating and can pay for itself in less than one year (based upon displaceable costs and not counting any increase in principal productivity; see below).

What a person can do with SFS

Message creation or editing. A person, using only a pushbutton telephone as a terminal, can (1) create an audio message, (2) review and replace all or part of it, and (3) insert, delete, move, and edit segments of it. Although most of our work so far emphasizes spoken messages, we have used SFS to compose and distribute handwritten and typed messages as well. Figure 2 is an example of how a user calls SFS, records a message, and transmits it.

A person can annotate audio, written, printed, or pictorial messages received from other people. Figure 3 shows an example of a user listening to a message from his manager, adding the answer to her question, listening to both voices now in the message, and

Figure 1 Conceptual diagram of SFS (from original rendition by Jennifer Howard)



returning it to his manager as well as to the sales manager. Note that it is equally easy for a user to annotate a message anywhere in that message or to start over. Experience has shown that annotation is often used, and users often like to send both the initial inquiry and the response.

Unlike a tape recorder or conventional dictating system, SFS is voice-activated. A person does not turn it off and on while recording. In effect, SFS records only when a person is talking. For subsequent ease of listening, as well as storage compaction, pauses between words are encoded and automatically com-

pressed by sfs.¹³ Encoding and storing pauses, rather than simply deleting them, allows users to put the pauses back in if they want to.

Users create messages quickly, usually in a minute or two.¹² Because of this, and because of a terminal (the telephone) almost always being nearby, users indicate they can compose a message under almost any circumstances, including during a meeting with other people. Users volunteer that they like being able to send someone a message when "it occurs to me."

As SFS has evolved over the years, it has been used by principals much more as a message system (super-telephone) than as an enhanced dictation system (see

Gould and Boies⁵): Our laboratory experiments showed that it is easier and more efficient for people to speak the contents of documents (whereby the recipient will listen to them) than to dictate the documents (whereby the recipient will read them).^{14,15}

Message distribution. A user can send a message to any person in the world who has access to a push-button telephone (or device to convert a dial telephone into a pushbutton telephone) and who knows how to use a telephone. Most often, users send messages to other "registered" users, i.e., users whose names ("user IDs") are stored in SFS and are played out. Messages are sent by key-pressing the last name of a registered user, as shown in Figure 2. In the case

Figure 2 Example of a user calling SFS, recording a message, and transmitting it. Pressing 00 means "do it," or transmit it

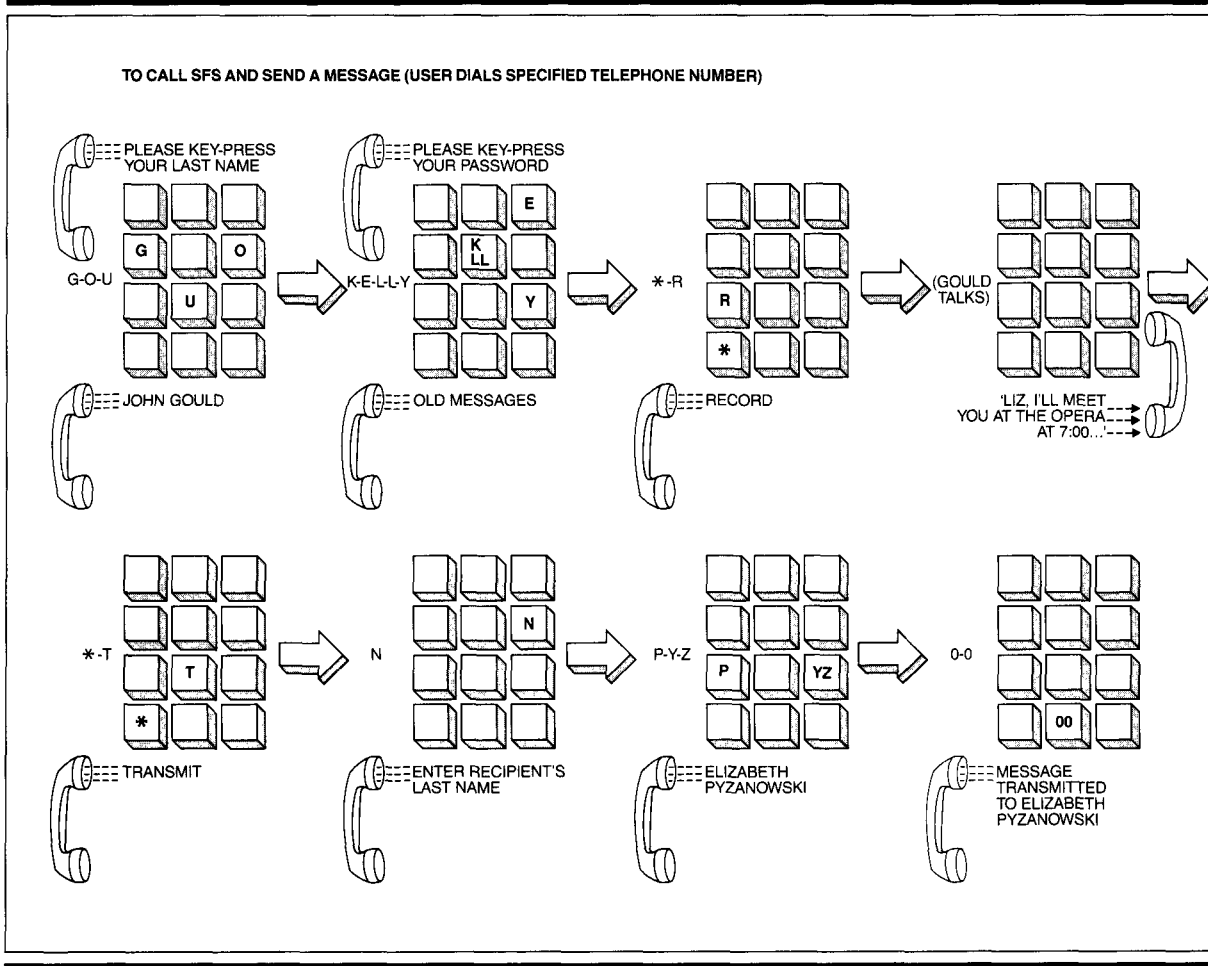
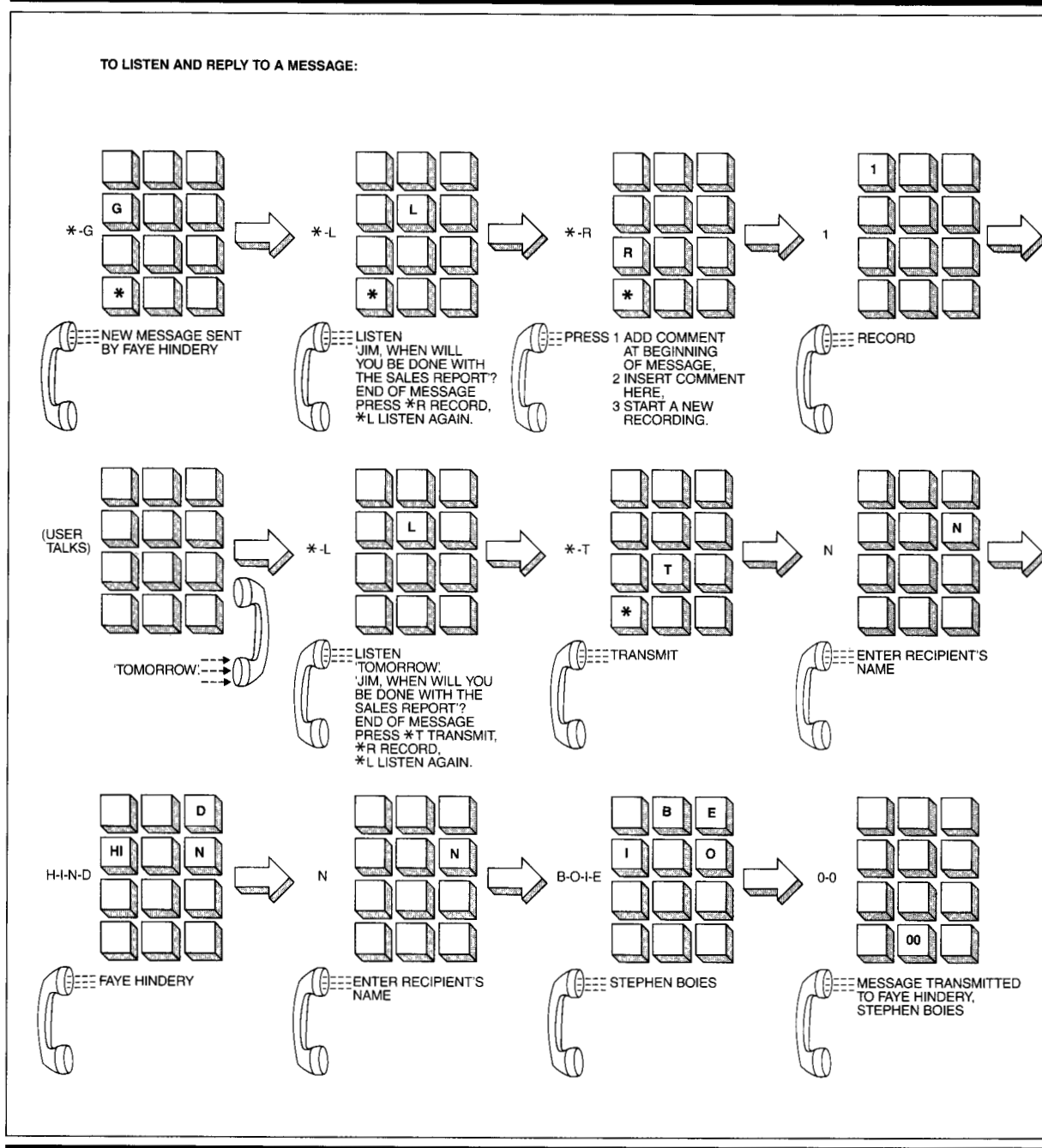


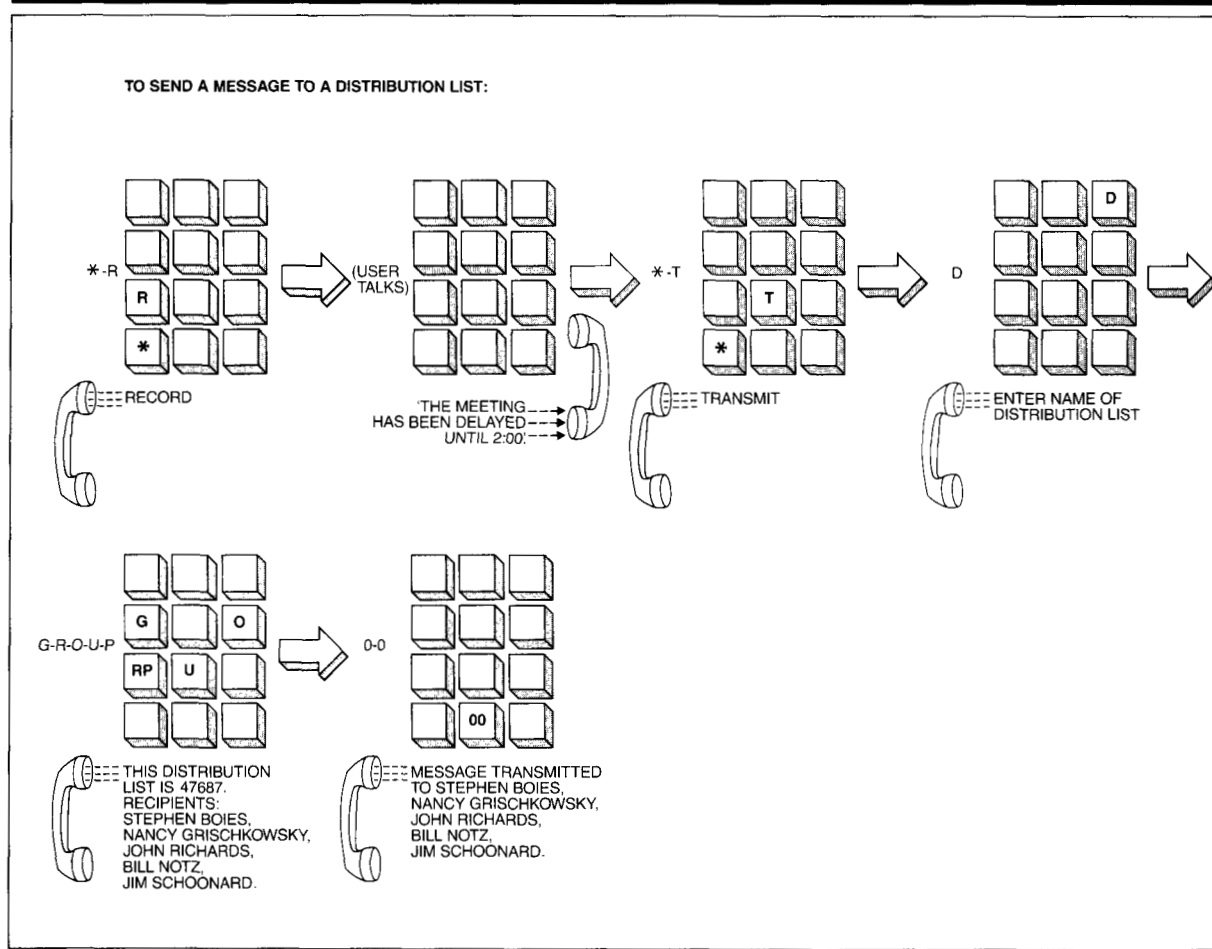
Figure 3 Example of a user (Jim) listening to a message, replying to it, and sending the reply back to the sender (Faye Hindery, his manager) and a copy of it to another person (Stephen Boies, the sales manager)



of ambiguities with names (e.g., several Smiths), SFS asks the user which Smith.

A user can send a message to more than one person at once, as shown in Figure 3, or to people on user-

Figure 4 Example of a user sending a message to a predefined distribution list. Users can give any name they choose to their distribution lists and can play out those names if they have forgotten them



defined distribution lists. Figure 4 is an example of a user sending a message to a distribution list. Users can play out the names of their distribution lists and the names of the people on them. The concept of distribution lists is further enhanced by SFS, which can store defined sets of people, e.g., group members and organizational charts, and automatically update these.

A user can actually send a message to *any* person in the world who has a telephone number, or receive a message from any such person. Figure 5 is an example of a user sending a message to a nonregistered person. In this case, the user key-presses the telephone number of the recipient. If a message is sent

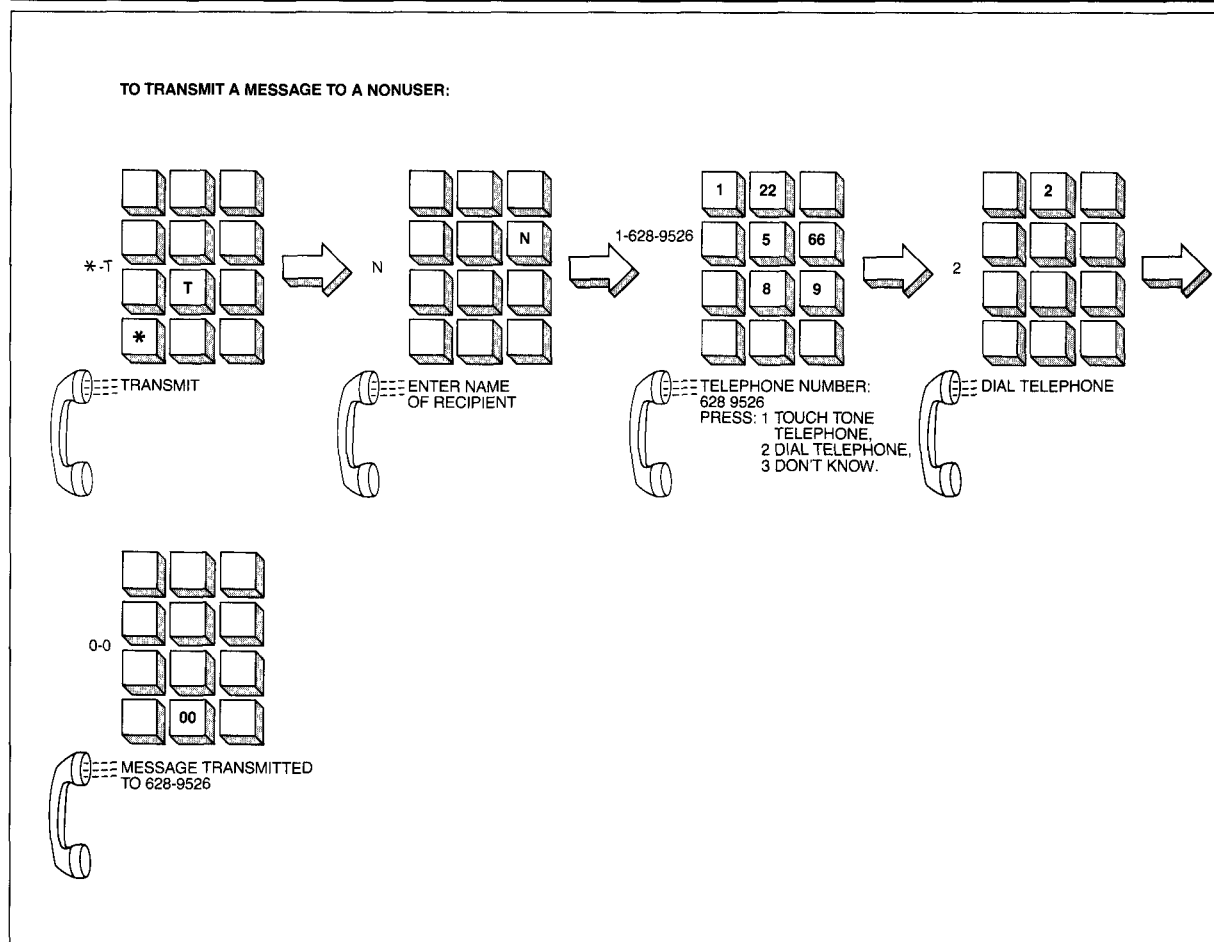
to a telephone number, the security may not be as great as when it is sent to a person's name because the message can be heard by anyone answering the telephone.

A user may send a request to several people, and as they respond, SFS notes this. SFS will make additional reminder calls to the remaining people, reminding them of overdue replies.

A user can page other people via radio in conjunction with a telephone.

Note that a recipient can receive a message at *any* telephone, which is very useful while traveling. The

Figure 5 Example of transmitting a message to a nonuser. Should a user make a mistake in keying a telephone number, he or she can key that number in again to delete it and then key in the correct number

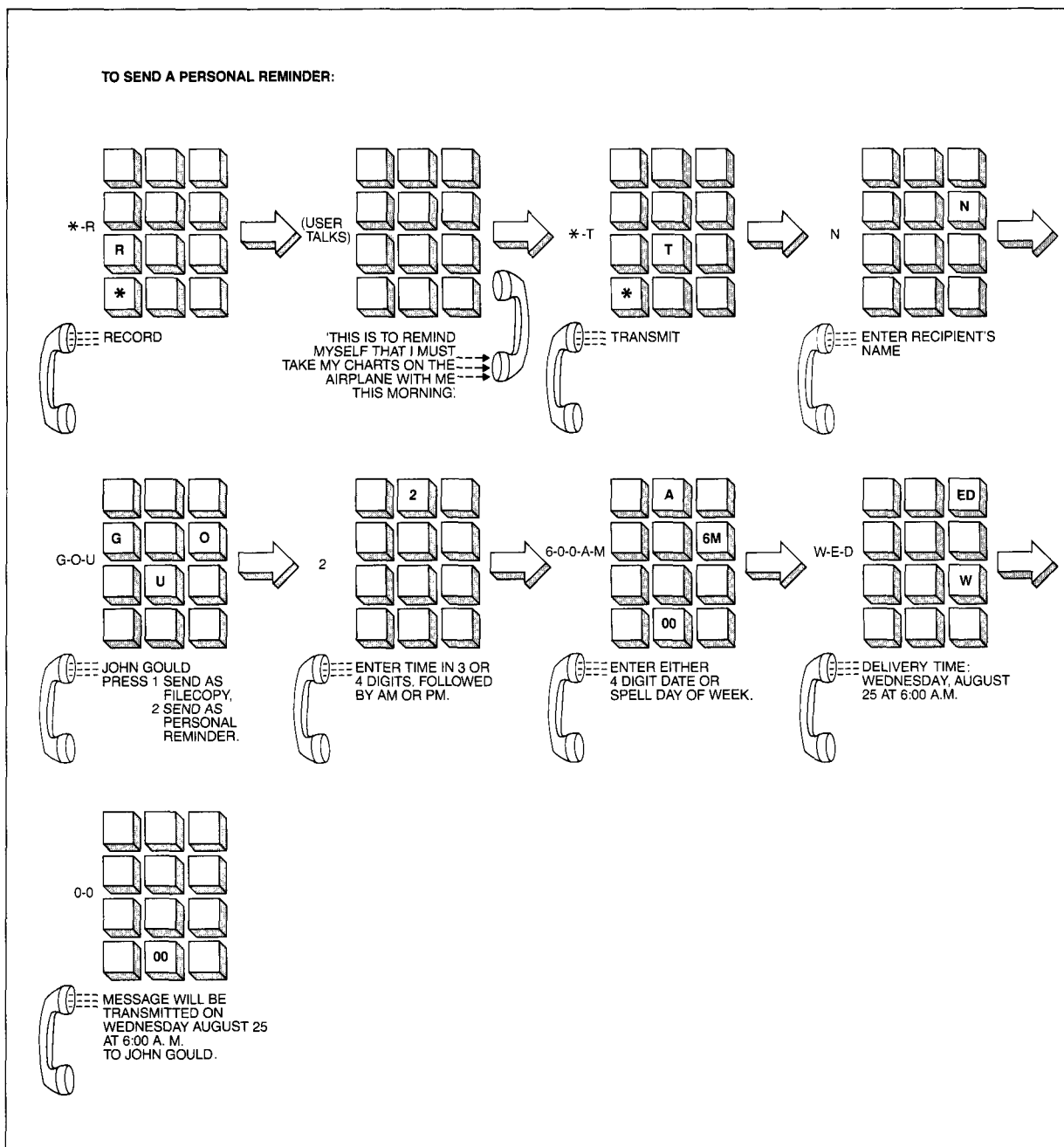


sender need not know where a registered recipient is or will be. Users can call SFS from any telephone, or be called by SFS at any telephone number included in the list of numbers where SFS can reach him or her. These numbers can correspond to a recipient's office phone, laboratory phone, and off-hours phone, for example. Users can easily change their numbers (from any telephone), which allows SFS to call them no matter where they are. The order and number of times SFS calls each number depends upon the hour of the day and whether the sender has assigned a special classification (e.g., urgent) to the message. SFS will call back in a few minutes if the recipient's phone is busy; it will try alternative numbers that the recipient has specified; and it will retry later if the phone is unanswered.

When a message is sent, SFS calls the recipient and notifies him about this. This call may be made at a time designated by SFS, or it may be made at times designated either by the sender or the recipient. SFS will automatically retry should a recipient be unavailable or should the number be busy.

A person can send a message to himself or herself to be delivered at any time (i.e., make his or her own phone ring at any time). Since one can send messages to oneself, SFS provides a reminder service. Figure 6 shows an example of a user sending himself a personal reminder, which in this case serves as a wake-up call as well. If the user makes a mistake keying in a time or date, he or she simply keys it in again.

Figure 6 Example of John Gould sending himself a personal reminder wake-up call. SFS will call him at home, his off-hours location, because of the time of day he entered



A user can send a message to a recipient and ask SFS to automatically notify him or her once the recipient has listened to it. Or, a user can ask SFS if a particular

individual has listened to a message that he or she sent to that individual. This is in contrast to other forms of noninteractive communication. Although

people can verify through the U. S. Post Office's Registered Mail system whether someone at the recipient's (presumed) location has *accepted* a document, this does not tell whether the intended recipient has actually *read* it. With SFS a user can learn whether a message he or she sent has actually been listened to. This information has proven valuable in reducing user uncertainty about whether a recipient, especially one who is traveling and whose whereabouts are unknown, has received a crucial message. SFS provides this information immediately, whereas U. S. Registered Mail, for example, requires several days, and even some electronic message systems have network delays.

A user can amend or cancel a message that has already been sent if the facts have changed or the information is no longer needed. We have adopted the policy that if the message is sent to several people, and if one or more people have listened to it, then the message cannot be amended or canceled.

A user can solicit a reply to his or her message, and both the sender and recipient can be automatically notified the day this reply is due, as well as if it is overdue. SFS automatically tells the user of the overdue status, and then plays out the text of the message.

Besides a New Message Box, an Old Message Box, and an Outgoing Message Box, each user has a Pending Message Box. The Pending Message Box contains a list of people who are overdue in replying to a user, as well as the people to whom that user owes replies. Users skip from one box to another with a two-key command. The most important box is activated when SFS calls a user or a user calls SFS. For example, if users have new messages, their New Message Box is automatically activated. If there are no new or pending messages, their Old Message Box is automatically activated when a user calls SFS (but in this case, SFS would not call a user).

There is a private bulletin board or shared message facility. If, for example, over a period of weeks a group of users wanted names of potential recruits for a new job, each user could send messages to this central collection point so that each member could hear the entire set (i.e., the messages from all group members) as it grows.

Message filing and retrieval. Each audio message has associated with it relevant statistics (e.g., time and date of creation, author's name, list of recipients, time at which each has heard it, and actions they

have taken on it). Upon retrieving a message, a user can play out these statistics (by pressing the W-key, i.e., the Who, What, When key).

A valuable application is filing a message to yourself containing important information that you use only occasionally but which is critical when you need it, e.g., the social security numbers of your children, project numbers, passbook numbers, credit card

A user can tailor SFS to his or her own needs.

numbers, zip code numbers, where to call if you lose your traveler's checks, automobile license numbers, and other hard-to-remember facts. These can be accessed at any time and from anywhere, e.g., if you should lose your credit cards while traveling.

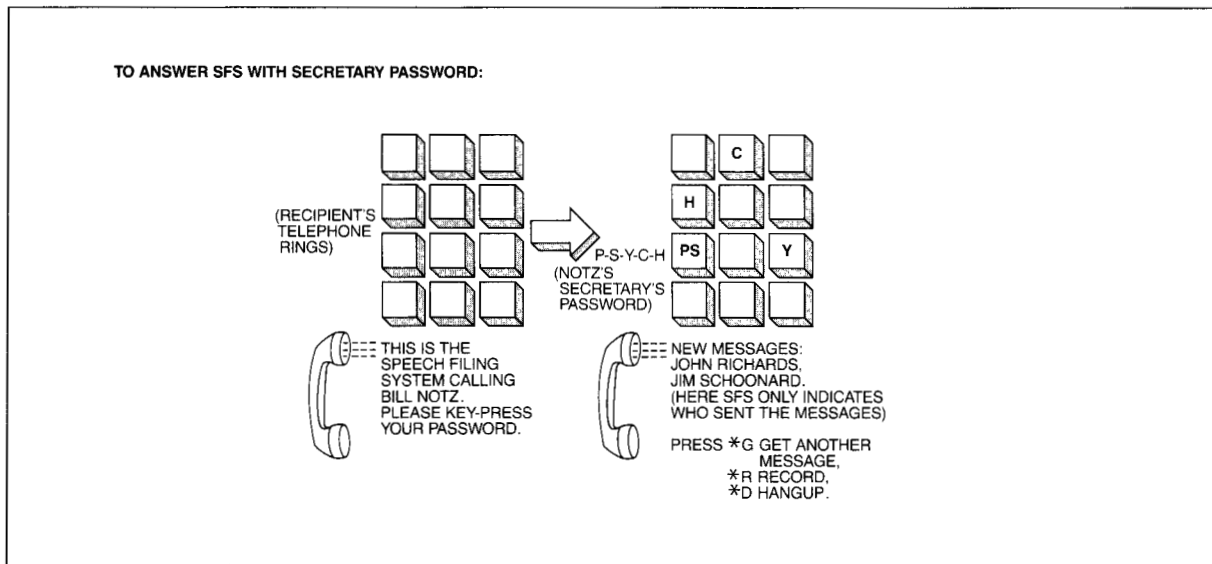
A sender or a recipient can classify or label a message with any heading he or she chooses. The recipient can give the message a coded name, via pushbutton telephone key presses, and can associate a spoken name with this coded name to provide a mnemonic name. Users can scan through the names of the files they have stored.

A person can retrieve a message by sender's name, by recipient's name, by file name, by date of origin, or by classification (e.g., personal, confidential, or secret).

Recipients can erase messages they receive. SFS also erases messages that have been saved past a certain time—unless a user requests otherwise. This time is set by a "customer" (i.e., the owner of the entire system). Retention dates can be limited by the sender. SFS provides an audit trail that may be retained indefinitely.

A person has great control over the message he or she sends. A person can *classify* a message (e.g., confidential, secret, or personal). A person can regulate what others may do with the message (e.g., listen only, listen and comment on it, listen, com-

Figure 7 Example of SFS calling a user and the secretary answering. With default options, SFS does not allow the secretary to listen to the messages, only learn who they are from



ment, and forward it, etc.), and how long they may retain it. A person can request responses to a message.

All these functions are accomplished by the user pressing the telephone pushbuttons.

Listening to messages. A user can listen to his or her messages from anywhere. Messages, of course, are heard in the voice of their creator. Pauses are automatically encoded when a message is recorded, so they can be contracted or expanded while listening. Listening to messages without pauses is much preferred over listening to messages containing the pauses of the author (as with traditional tape recorders or dictating equipment). Experiments have shown that comprehension of material heard via SFS's pause-deleted speech is equal to or better than that for the same material heard via careful reading or via extemporaneous speech.⁴ The limiting factor in listening to SFS messages is the quality of the telephone lines. A user can change listening speed, increasing it several times over the rate it normally plays out. Other listening aids are described below.

Personalization. A user can tailor SFS to his or her own needs via pushbutton telephone key-press commands. A user has four passwords: personal, secretary, family, and guest. A user can use all SFS func-

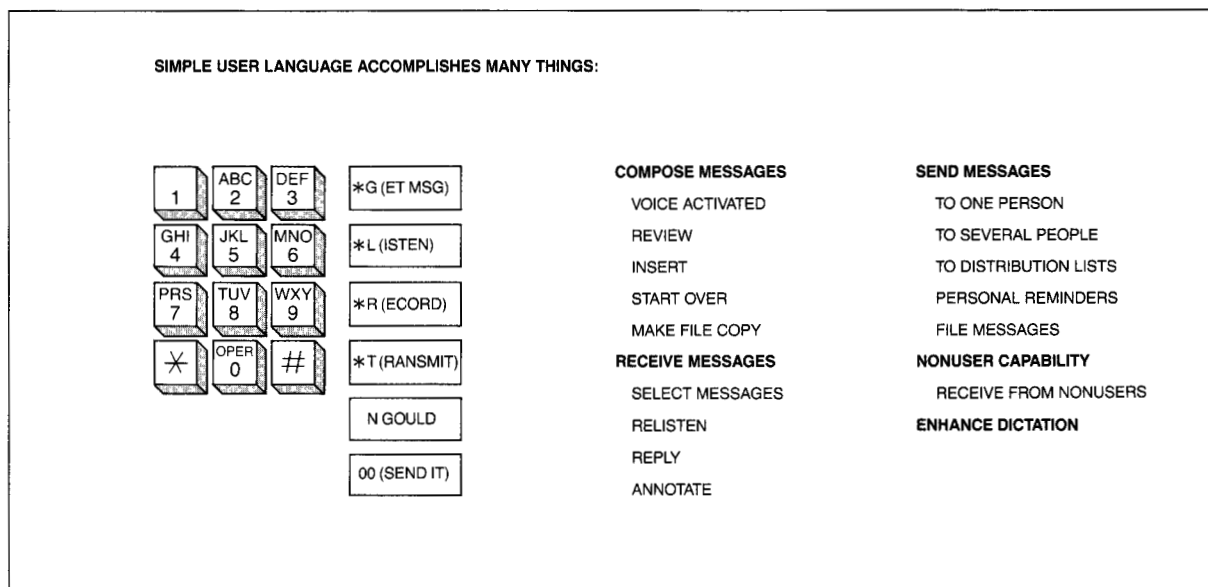
tions with his or her personal password. Initially, SFS assigns a subset of the functions to the other passwords.

Figure 7 is an example of what happens when SFS calls and a secretary answers. With the secretary password, a secretary can learn who the principal's messages are from, but not listen to their content. A principal can modify what his or her secretary's password (as well as the others) can do, assigning more function to it. For example, a principal may want the secretary to be able to listen to all nonconfidential messages.

Special attention, in the form of a special password and priorities, is given to a user's family. (Ever try calling home at 4:30 p.m. and get a busy signal?) With the family password, one accesses only the messages designated for the family, e.g., "I'll be home late tonight."

A user has three telephone numbers where SFS will call him or her. Users can call SFS and set these numbers from any telephone anywhere, which is useful when a user is at another location and wants SFS to call there. These numbers are usually the user's office number, secondary work number, and off-hours or home number.

Figure 8 The user command language that evolved from SFS and now runs on the ADS product



Several user interface languages are available, and the user can select the one he prefers. A new user, for example, may choose to use the system-driven, prompted language that experience has shown requires no user training.⁵ The experienced user, on the other hand, may not want the prompted, or audio menu, approach, but would rather choose to key-press the commands based upon his recollection of them.

Users can select the rate at which system messages automatically play out, making pronunciation faster or slower than originally recorded. It has been found that with experience, users like messages to play out faster.

SFS is designed to be able to monitor automatically user behavior and to provide feedback and help appropriate to the user's knowledge.

Results of using SFS

Useful functions. The two major user interface issues in any system are identifying the required useful functions and making these easy to learn and use. The first issue was initially accomplished with our behavioral analyses of principals' needs. The suc-

ceeding years of empirical research that studied the functions principals actually used and the additional functions they wanted led to major refinements (see Gould and Boies⁵) which proved ultimately to be successful, as mentioned earlier under "Key results."

Learning SFS. The second basic behavioral challenge was to make SFS easy to learn and use. The main problem centered on mapping the 100–200 different SFS functions onto the 12-key telephone set. The basic commands arrived at after several years of empirical and experimental research⁵ are shown in Figure 8. They are *G to get a message, *L to listen to a message, *R to record a message, and *T to transmit a message. Each of these commands is, in effect, a command mode. For example, within the Transmit mode the N key is the prefix for a user's Name. With these basic commands, a user can do all the things shown in Figure 8. In addition, there is a Customize mode (*C) that allows users to personalize SFS for themselves, e.g., set their passwords and telephone numbers. The Customize mode is entirely prompted—after pressing *C a user selects from a series of three-choice alternatives—because it contains much function that is only infrequently used.

The solution that evolved included identification of the basic functions people wanted, a small number

of commands to accomplish these functions (with only a few key-presses), feedback messages after every key-press, selective prompting messages, and almost no need for documentation. Experience has shown that users now require only a few minutes of training to learn these basic functions. Getting to

Users learn the basic functions in a few minutes.

this point, however, took years of human factors empirical research and several major iterations on the user interface.⁵

The self-help tools on SFS have in fact proven helpful primarily because they are integrated with the command languages and are easy to use. Pressing the # key causes suggestions about what the user might want to do next, conditioned by the user's context, to play out. Pressing the W key (i.e., 9 key) answers who, what, when, and where questions about a message a user is listening to or creating. The aim of this help system is not merely to aid the user in using SFS, but to aid the user with the general problem he or she is trying to solve.

From novice to experienced user. We have attempted to solve the general problem in SFS, found in all user-oriented interactive computer systems, of how to help a person use additional features of the system as he or she becomes more experienced. One simple method that new users in some customer locations have liked is to play out a "Helpful Hint of the Week" to a user upon first using SFS that week. In another approach, after a (programmable) number of uses by a given user, the SFS user interface can be automatically changed for that user. This approach has worked well in the case of automatically speeding up the rate of system message playback after a user uses SFS a certain number of times.

Handicapped users. SFS would seem to have advantages for the blind, who can listen to documents that might otherwise be available only in printed form. SFS ordinarily plays messages at about 200 words per minute, without being much affected by the rate at

which they are recorded. The sped-up speech facilities of SFS allow messages to play out at about twice this rate, which approaches the maximum rate of speech perception.^{16,17} The few manual controls and possibility of voice commands (see below) make SFS potentially attractive for the manually handicapped also.

Some observations on SFS usage. The following summarizes some qualitative observations on SFS usage. Users learn the basic functions in a few minutes. They require little or no documentation. SFS is truly principal-oriented. Principals use it directly. We know of no example in which a principal has asked his secretary to send a message for him or her. Much more is accomplished by noninteractive communication than we had initially imagined. Principals send content messages (as opposed to simply asking the recipient to return a telephone call). SFS appeals to a wide range of users, including those who dislike gadgets and had no intention of ever using a computer terminal. SFS is especially used by managers. SFS is used remotely (at night, while traveling, in other offices, etc.). Perhaps of most relevance, experience shows that users miss SFS if it is unavailable for a short time—as was occasionally the case with the research prototypes. All these observations are elaborated upon in Gould and Boies.⁵

What SFS does for principals. The following is a summary of informal positive comments we have heard from users of SFS over the last several years. Users like to be able to communicate with people at any time, from any place, and are especially appreciative of this when they do not know where other people are. It solves the time-zone problems, e.g., California to New York, Europe to U.S. (The fact that all times are based upon the SFS installation location does not seem to be a problem, probably because few timed deliveries are sent across several time zones.) Users say SFS relieves their anxieties about certain types of communication, and relieves what would otherwise be a larger (human) memory load. It helps them to control their interruptions and to communicate on relevant matters without getting into other, unwanted subjects. They like the personal touch of voice messages. SFS is said to eliminate the burden of routine matters. Users do not have to hunt around for people, and they can inform several people at once. It is easier and quicker to compose spoken messages than to compose written, dictated, or text-edited messages.^{14,15,18,21,22} Users say that SFS improves their intragroup and intergroup communication.

Table 1 Estimated ADS displaceable costs

Operating Expense	Cost Savings*
Reduction in long-distance telephone costs (Assumes 6 fewer long distance calls/mo.)	\$15.00
Reduction in memos—saved secretary time (Assumes 8 fewer memos/mo.)	51.60
Reduction in secretary message handling (Assumes 5 fewer messages/day)	26.65
TOTAL PER USER PER MONTH	\$93.25
COST OF ADS PER USER PER MONTH	—\$10.00
NET COST SAVINGS WITH ADS PER USER PER MONTH	\$83.25

*Estimates are based upon unpublished studies available to the authors from IBM sources of what principals do without ADS and informal studies of what principals did with SFS prototypes. The ADS cost per user assumes users already have a paid-for telephone and at least 200 users on an ADS, and includes maintenance charges.

Cost of SFS. An estimate of the costs of SFS can be obtained from the commercially available IBM Audio Distribution System. ADS costs about \$10.00 per month per user for an average system configuration,⁵ which is much less expensive than electronic message systems used from video terminals.¹⁹

ADS is a principal-support system. Generally such systems can be cost-justified only in terms of increased principal productivity, which is difficult to estimate or measure. However, given the way ADS impacts principals' behavior, it can, fortunately, be cost-justified in terms of displaceable costs (so-called "hard dollars") as well. Three normal operating costs, shown in Table 1, can be reduced by ADS. First, long-distance telephone costs can be reduced because customers can lease a single 800 number for ADS. Second, typing costs can be reduced because fewer typed memos are required. Third, telephone message handling can be reduced because of fewer missed telephone calls. As shown in Table 1, we estimate these displaceable costs to be currently \$93.25 per user per month in a typical office. Thus, the net cost saving is \$83.25 per user per month.

Additional SFS features

What has been described so far are the general characteristics and main functions of SFS, and how and why people are using them. In this last section we

describe some additional SFS features not yet mentioned.

Feedback. SFS provides the user with stored voice prompts, hints, suggestions, and menu choices for accomplishing useful work. SFS also responds to the

SFS has enhanced dictation facilities.

user with stored-voice confirming messages when key-presses are appropriate. When key-presses are inappropriate, SFS responds with stored-voice error messages, warnings, and helpful hints.

If the user is stuck or still does not know what to do after prompting, SFS aids in other ways. First, if a user, during a work session, has not done anything for a prescribed length of time (which is usually tens of seconds and depends upon what the user is trying to do), SFS detects this and tells the user about his or her options. What SFS tells the user is entirely dependent upon where the user is in the system. Second, a user presses the # key and SFS lists the options. Again, this is entirely context-dependent. A user can also press the W key. SFS then plays out answers to who, what, when, and where questions. Finally, a user can select a completely prompted language to guide him or her through the basic functions. This selection is made by changes in the user's profile (i.e., the user presses *C to get into the Customize mode).

Mnemonics. There are many levels of mnemonics, or memory aids, designed to make SFS easy to learn and use. The command names shown in Figure 8 illustrate verbal memory aids, e.g., G for Get, L for Listen. There are verbal mnemonics within a mode also. For example, in Transmit mode, the "D" key is the prefix for Distribution lists, the "N" key is the prefix for a user's Name, and the "W" key provides Who, What, When, and Where information.

An example of a spatial mnemonic is illustrated in Figure 9. SFS allows a person to skip around within a message. Imagine a person's forefinger resting on

the 5 key. To go to the beginning, or *top*, of a message, a user presses 5 and then moves his finger *up* to press the 2 key. To go to the end, or *bottom*, of the message, a user presses 5 and then moves his finger *down* to press the 8 key. To skip back a little bit (several seconds of speech), a user presses 5 and then moves his finger to the left to press the 4 key. To skip forward a little bit, a user presses 5 and then moves his finger to the right to press the 6 key.

Enhanced dictation. SFS has enhanced dictation facilities, both for creating a long document and for listening to it or another one. With traditional dictating equipment, authors cannot copy, move, or delete segments of speech. With SFS, however, they can do these operations anywhere in the document. They can also insert material anywhere in the document. (Of course, they can record over the last several words of what they said, as with traditional dictating equipment.) They can listen to a document (theirs or somebody else's) in sped-up speech. They can attach audio format markers, through key-presses, to various parts of the document. We conceive of these markers as sentence, paragraph, and topic markers. When pressed, each causes an audio tone of a different frequency to be heard by the user. They are hierarchically organized, provide listeners clues to document structure analogous to what punctuation and indentation provide readers, and facilitate listening. The user, again through key-presses, can, for example, skip to a new paragraph, or listen only through a particular topic.

Voice commands. Presently SFS does not include any speech recognition. Users give SFS commands by pressing buttons on their telephones. The SFS input stream contains pushbutton signals mixed together with voice signals. SFS automatically separates these two, and interprets the frequencies corresponding to the telephone button presses as commands. SFS could be controlled with voice commands instead of key-presses, as our early feasibility experiments demonstrated. SFS would recognize and separate out the speech patterns corresponding to the voice commands. A special dictionary might be needed for each user, however. Further, users could not use SFS command names or user names in their messages, unless other precautions were taken. The programming underlying SFS makes it possible to switch between command modalities.

Coded data collection. When SFS is used as a voice message system, it is storing and manipulating so-called *noncoded information*. SFS can also collect

Figure 9 Example of a spatial mnemonic, or memory aid, in the Listen mode

< >	<2>	< >
<4>	<5>	<6>
< >	<8>	< >
< >	< >	< >

coded information. That is, people can enter actual coded data remotely, through the telephone, e.g., social security numbers, customer orders, and research data. There are many important SFS applications of this function. In a study conducted by Eric Goldwasser in 1978 (unpublished), IBM field engineers called SFS, often from a customer's location, just after completing a service call. These field engineers were immediately prompted with a series of questions, e.g., the customer's name, the serial number of the machine repaired, etc. They were asked to key in part numbers that needed to be ordered, and the quantity required. (Appropriate branching was included here so that as many part numbers could be ordered as needed.) These data were automatically transferred, on a computer network, to the appropriate departments in IBM. This operation allowed quick processing of orders, without the errors introduced by delayed data entry. Further, it also provided fast and accurate tracking of the performance of the new computer processors that these field engineers were servicing. After entering the data, the field engineers could make voice comments about the customer or the new processors they were servicing, and these comments were relayed to product people.

Language for coded application design. Creating these prompting sequences for new applications was easily done using the SFS computer-assisted instruction facility designed for applications people by Jim Schoonard.²⁰

Remote control. SFS is designed in such a way that it could be used to remotely control analog and digital instruments and appliances, given appropriate telephone interfaces. To cite examples relating to the home, one could raise his home temperature just

before leaving work in the evening, or turn on a microwave oven from any telephone. Since most principals have irregular work hours, this feature has advantages over preset controls. SFS is also designed in such a way that it could be used in connection with biofeedback applications and other health-related instrumentation.

Mixed media. SFS, as mentioned earlier, is a document distribution system that can accept as input and distribute all media, e.g., written, spoken, typed, graphic, and facsimile. Users can edit and annotate other media documents just as they can edit and annotate audio messages.

Telephone answering. If SFS is attached to a telephone exchange (either public or private, as ADS is in Europe), incoming calls can be automatically transferred to SFS when a person does not answer the telephone. The caller can hear the recorded SFS message and then can leave a message which will be placed by SFS in the user's New Message Box.

SFS is basically different from conventional telephone-answering devices⁵ in that it has much more function. In addition, with telephone-answering devices, callers generally *expect* to talk directly with a person and are disappointed when they hear a taped message instead. In using SFS, users have no such violated expectations.

Generating automatic documentation. SFS provides automatic user documentation of itself, thus eliminating the errors that often occur when user documentation of a system is not kept in step with how the system actually operates. We keep documentation and system in step by using a virtual machine (VM) simulator to generate the documentation. The code (in the form of the command tables) that controls SFS (on an IBM Series/1 computer) is put into this VM simulator. The 12 function keys on the IBM 3270 display terminal are arranged in the same spatial array as the keys on a pushbutton telephone. A person uses the SFS simulator with these keys. Printed (rather than audio) system messages appear on the video screen. A user types, rather than speaks, his messages. As the document designer uses the simulator, the results can be printed, thus supplying the final documentation.

Summary

Communication is the main activity of business people. SFS was designed to improve communication.

Using a pushbutton telephone as the terminal, users can send messages to anyone in the world and receive messages from anyone in the world. Because of the years of human factors research that led to several iterative improvements in the user interface (described more fully in Gould and Boies⁵), SFS includes useful functions that are easy to learn and use. An IBM product, the Audio Distribution System, is based upon the SFS research and is now commercially available.

Acknowledgments

The work described in this paper is not just that of the authors. SFS was a team effort, with Stephen J. Boies serving as the leader. Other members of the group who contributed significantly over several years were John Gould, Nancy Grischkowsky, Bill Notz, John Richards, David Zeheb, and Jack Duffy. We thank Clayton Lewis, Bill Notz, and Mary Beth Rosson for their comments on earlier versions of the manuscript.

Cited references

1. S. J. Boies, "A computer based audio communication system," *AIIE Conference on Automating Business Communications*, (January 23-25, 1978), pp. 369-372. (Paper can be obtained from Management Education Corporation (MEC), Box 3727, Santa Monica, CA 90403.)
2. D. Zeheb and S. J. Boies, "Speech filing migration system," in H. Inose (Editor), *Proceedings of the International Conference of Computer Communication* (September 1978), pp. 571-574.
3. J. D. Gould, *An Experimental Study of Writing, Dictating, and Speaking*, Research Report RC-6186, IBM Corporation, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598 (1976).
4. D. Nix, *Two Experiments on the Comprehensibility of Pause-Depleted Speech*, Research Report RC-6305, IBM Corporation, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598 (1976).
5. J. D. Gould and S. J. Boies, "Human factors challenges in creating a principal support office system—The speech filing system approach," *ACM Transactions on Office Information Systems* (1983, in press).
6. *IBM Audio Distribution System Subscriber's Guide*, SC34-0400-1, IBM Corporation, 4111 Northside Parkway N.W., Box 2150, Atlanta, GA 30056; also available through IBM branch offices.
7. E. T. Klemmer and F. W. Snyder, "Measurement of time spent communicating," *Journal of Communication* 22, 142-158 (1972).
8. G. H. Engel, J. Groppuso, R. A. Lowenstein, and W. Traub, "An office communications system," *IBM Systems Journal* 18, No. 3, 402-431 (1979).
9. H. Mintzberg, *The Nature of Managerial Work*, Harper and Row, New York (1973).
10. R. R. Panko, "The EMS revolution (A survey of electronic message systems)," *Computerworld* (August 25, 1980).

11. R. Lowenstein, IBM Stamford, CT, personal communication based upon the results of an extensive unpublished field study of principals conducted at a large corporate headquarters building (1979).
12. J. T. Richards, *The IBM Research Speech Filing System: Analysis of the System 7/S370 Prototype*, Research Report RC-9114, IBM Corporation, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598 (1981).
13. S. J. Boies, W. A. Notz, and D. Zeheb, "Encoding and decoding digital speech," *IBM Technical Disclosure Bulletin* 19, No. 6, 2357-2359 (1976).
14. J. D. Gould, "An experimental study of writing, dictating, and speaking," in J. Requin (Editor), *Attention and Performance VII*, Lawrence Erlbaum Associates, Hillsdale, NJ (1978), pp. 299-319.
15. J. D. Gould, "How experts dictate," *Journal of Experimental Psychology: Human Perception and Performance* 4, No. 4, 648-661 (1978).
16. R. P. Carver, "Effects of increasing the rate of speech presentation upon comprehension," *Journal of Educational Psychology* 65, 118-126 (1973).
17. E. Foulke and T. G. Sticht, "Review of research on the intelligibility and comprehension of accelerated speech," *Psychological Bulletin* 72, 50-62 (1969).
18. J. D. Gould, "Writing and speaking letters and messages," *International Journal of Man-Machine Studies* 16, 147-171 (1982).
19. R. E. Rice, "The impacts of computer-mediated organizational and interpersonal communication," in M. Williams (Editor), *Annual Review of Information Science and Technology* 15, Knowledge Industry Publications, White Plains, NY (1981), pp. 221-249.
20. J. W. Schoonard, *How to Use Speakeasy*, unpublished manuscript (1979). (Obtainable from the author at the IBM Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, NY 10598.)
21. J. D. Gould and S. J. Boies, "Writing, dictating, and speaking letters," *Science* 201, 1145-1147 (1978).
22. J. D. Gould, "Experiments on composing letters: Some facts, some myths, and some observations," in L. Gregg and E. Steinberg (Editors), *Cognitive Processes in Writing*, Erlbaum and Associates, Hillsdale, NJ (1980), pp. 98-127.

John D. Gould IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598. Dr. Gould is a psychologist and research staff member at the Research Center. He has been interested in human factors and offices for almost a decade. He was responsible for the user interface on the Speech Filing System. Among his achievements was to help Stephen Boies.

Stephen J. Boies IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598. Dr. Boies is a psychologist and senior manager of office systems applications at the Research Center. He has been interested in human factors and offices for almost a decade. He managed and led the Speech Filing System project. Among his achievements is that John Gould works for him.

Reprint Order No. G321-5210.