WORKING NOTE AC96

Alan Creak
2 October 1996

REACHING BEYOND WORDS IN REHABILITATION COMPUTER SYSTEMS.

In computer-based communication systems such as are developed for rehabilitation purposes person-to-person communication is the important aim. Software design techniques do not of themselves take this into account, and may therefore lead to interfaces which are less satisfactory than they might be. An approach which encourages attention to the higher levels of human communication is suggested as an aid to software specification, and examples in software design and fault diagnosis are presented.

COMPUTERS MEDIATING IN HUMAN COMMUNICATION.

In an attempt to explore questions of design in developing computer software for rehabilitation purposes, Creak and Sheehan[1] have suggested a systematic approach to describing message-carrying signals in communication systems. They recommend that signals should be described in terms of certain attributes corresponding to different levels of abstraction in the system. Generally, the attributes may or may not be changed by system components and at interfaces between components. If loss of some aspect of the communication is to be avoided, the description of the signal should be unchanged across every interface (in effect a requirement that plugs and sockets must match exactly), and those of its attributes which describe the intended meaning in human terms must be preserved within the system components. Attributes describing details depending on the implementation – for example, the structure of the representation or the encoding technique used – may change as the signal moves through the system, but the requirement that the meaning be preserved ensures that the compete message can be conveyed. This view is not put forward as a design technique, but rather as an aid to system specification. By defining the attributes broadly it is hoped to include all significant aspects of the communication; by examining each interface or component with the attributes in mind, attention is drawn to possible areas of concern.

The intention of communication can be informally described as to copy an idea from the mind of the initiator to the mind of the recipient. Practically speaking, ideas are not accessible to computer – or any other – communication systems, and it is the initiator's responsibility to encode an idea in terms of the vocabulary provided by the chosen communication medium, and to do so in such a way that the recipient can decode the received message in terms which approximate to the original idea. The communication system is then required to handle the encoded message as provided by the initiator, and to deliver it to the recipient.

How can this vocabulary be defined ? In the archetypal communication method of face-to-face speech, the vocabulary used is not limited to words. Non-verbal communication through facial expression, gesture, and intonation are important, commonly functioning by qualifying the basic sense communicated in words. Few mechanical media provide for the use of such additional channels of communication, though their effects can be significant. If a mechanical system were required to make such provision, it would be necessary to define additional vocabulary – not necessarily verbal – to specify and represent the information carried in the new channels.

The notion of adding non-textual components to synthesised speech is not new. For example, there has been some work on generating emotional speech patterns[2]; the result was HAMLET (Helpful Automatic Machine for Language and Emotional Talk). (It is perhaps rather sad that of the six emotions implemented – anger, happiness, sadness, fear, disgust, grief – only one is positive.) The new vocabulary in this system appears as a set of actions used to select the desired emotion from a table presented on a computer screen.

Creak and Sheehan refer to the attributes at the highest level as the *human form* of the signal, which is contrasted with the *machine form*. The definitions given[1] are, respectively, "How the information is represented at the level of human perception" and "How the information is represented at the level of machine encoding". At the machine level, we describe the signal in terms of bits and bytes, how they are organised, and how they should be interpreted; at the human level, we describe the same signal in terms of the organisation and interpretation of units of speech (words, intonation, gesture, etc.). This is the most ambitious practicable goal for system design, as the most we can possibly undertake in designing the computer system is to transmit a message without losing any of the information given to us. If it is nevertheless misunderstood, then either the information given was inappropriate, or some part of it has been lost or altered in the transmission. The criterion for perfect design is that the human form of the signal must not change during the transmission.

An example will clarify the recommended approach. Consider a simple sentence :

Mary had a little lamb.

As it stands, it is a straightforward statement of ownership at some past time, and a satisfactory human-form description of the message is a report of the text. Quite commonly, though, that is not a complete description; the same words can be used to convey more by adding emphasis :

Sentence with emphasis :	*Possible implication :*
Mary had a little lamb.	(– but Freddie didn't.)
Mary *had* a little lamb.	(– but doesn't any more.)
Mary had a *little* lamb.	(– but Freddie had a big one.)
Mary had a little *lamb*.	(– not a little dog.)

(That list does not exhaust the possibilities – for example, each of those four statements can easily be spoken as a question.) In each case, the sentence with emphasis carries more meaning than the simple, unemphasised, sentence, and in each case the added meaning is different. The interpretation to be put on the emphasis depends quite strongly on the context – the examples listed above are no more than possibilities – but there is certainly more meaning in the sentence with emphasis than is conveyed by an unadorned copy of the words spoken, and it may well be important for a proper understanding of the text or discourse of which the statement forms a part. The emphasis must therefore be included in the human-form description of the sentence; as we are not committed to any formal constraints, we might suppose that the human-form statement of the first example could be written as

"The text 'Mary had a little lamb' with emphasis on 'Mary'."

That is what we wish to convey unchanged through the system. How is it conveyed in practice ?

In speech, we show emphasis by various means. Typically, we speak the emphasised word a little slower, a little louder, at a slightly higher pitch. We can suppose that, instead of relying on the single communications channel of the words themselves, we are now using two channels, the words and the prosody. In handwriting, we can use underlining or capital letters, while in printed text, as in the table above, we can use italic or bold characters; again, we are using two channels, the words and the appearance of the words, instead of only one. If we are restricted to a single channel, we can still cope, but we have to encode the non-verbal information in the same channel. Markup languages such as TeX and HTML operate on this principle by defining additional character sequences : "Mary had a little lamb". (A good example of the spontaneous evolution of such extensions is the set of abbreviations and "smileys" developed by addicts of the aggressively single-text-channel medium of electronic mail.) Such encoding methods are well suited to machine consumption; in text intended for reading they are comprehensible (if you know the code), but irritating if they appear frequently. As a more elegant alternative, it is always possible to make the implicit explicit by writing down in the text the phrases corresponding to those on the right-hand side of the list above.

However the emphasis is encoded, it carries additional information, so – at the level of information theory – it must be represented by bits in the message transmitted. These additional bits may be conveyed by auxiliary channels, or they may be included in a single channel with the primary communication, but if the information which they encode is lost the meaning is changed. The intention of introducing the human-form description is to focus attention on this possibility, so that steps may be taken to preserve the information, or, failing that, it is made clear where information might be lost.

The importance of the non-verbal component of communication, illustrated here by the example of emphasis, is demonstrated by its common use in speech, writing, and other media. Nevertheless, there is rarely any provision for such features in rehabilitation systems. A significant difficulty is that to represent the non-verbal component it is necessary to use some of the bandwidth of what might already be a very narrow channel; but in view of its importance, it is worth investigating how such provision might be made, and the approach advocated by Creak and Sheehan, by emphasising the need for communication of the complete message, offers a framework within which the question can be addressed.

In this note, I illustrate the utility of the approach by applying it to the notional design of an interface which would handle emphasis, and using it as the basis for discussion of a published case study. The first illustration gives an example of system design, while the second shows how the technique might be useful in discussing reasons for poor performance in practical systems.

DESIGNING AN INTERFACE FOR EXPRESSIVE COMMUNICATION.

In order to design an interface permitting non-verbal communication, it is first necessary to decide just what sorts of communication are to be supported. This follows from the discussion above; if the human-form statement must contain explicit note of the sender's intentions, some vocabulary in which this material is to be expressed must be provided, and that in turn requires a decision on what is to be conveyed. Here I have chosen to continue the discussion of emphasis of individual words. No implication that this is the most important non-verbal mode of communication is intended; it is no more than a convenient and realistic example.

We can reasonably assume a system such as is described by this scheme :

Person → Transducer →
Computer Input →
Software →
Computer output →
Transducer → Person

(The scheme, though fairly detailed, is oversimplified. The oversimplification is unimportant for this illustration, but a more complete analysis would obviously be necessary in a serious design exercise.) We wish to design the system so that messages with human-form descriptions including emphasis associated with individual words can be faithfully transmitted from person to person. To do so, we must define how the emphasis is to be encoded in each of the transitions between the system components shown in the scheme, and we must make sure that the components themselves can deal with the emphasis in an appropriate way.

Let us suppose that the person is accustomed to using a conventional scanner operated by a single switch. At any moment, some selection from an available set of characters is displayed on a screen, with the displayed set changed at intervals of about one second to cycle repeatedly through the complete available set. Pressing the switch when a set is displayed selects the set, and then the process is repeated to select a single character from the set. The scanner then reverts to its original mode of cycling through subsets.

We begin with the first transition, from Person to Transducer. The first task is to extend the vocabulary available at the interface in order to make it possible to represent emphasis. We must then select an encoding which the person can produce, and which the transducer can receive. These are the only constraints at this point, though we shall see that other considerations can limit the field of choice. Here are three possible encoding schemes which use the scanner, but allow emphasis to be added as required :

· Use markup symbols, as with – for example – HTML. Text could then be entered in the form "Mary had a little lamb". Simpler forms which would use fewer characters are obviously easy to devise, but the principle of the method is to add further conventional text to the message; no new vocabulary is required. At some later stage in the operation, a system component must be able to recognise the markup code and do something with it.

· Use some switch action which has no direct textual representation – for example, operating the switch twice instead of once when making any selection while entering a word could be taken to indicate emphasis for that word. (I refer to this double action below as a "double click", by analogy with the similar mouse operation.) This method introduces a new vocabulary item, which is carried in the same data stream as the original items. The next stage in the operation must be able to identify the new vocabulary, and handle it in a suitable manner.

· Use another switch, perhaps operated by the other hand or some other body part. Closing this switch at any time during entry of a word could be interpreted as associating emphasis with the word. Notice that the dexterity required for this operation is significantly less than that needed for the primary switch control. A new vocabulary item is again introduced, but in this method it is transported through a separate data stream; further on in the processing, some component must be able to bring the streams together.

What are the implications of these possible decisions ? Here is a brief discussion of the consequences for the system design at successive stages of the computation.

Person → Transducer.

Assuming that we know that the person concerned is capable of producing the symbols, the immediate implications involve only the transducer : the first two methods require a single switch, while the third requires two switches. The encoding techniques are different in all cases, with the first using only conventional scanner signals with the additional information encoded as recognisable sequences of conventional characters, the second using modified scanner signals with the additional information encoded as double clicks, and the third relying on a dual-channel input system with the emphasis signals separated from the conventional scanner signals. The choice will be (or, in an ideal world, would be) governed by the ability of the person who will be using the system; the fastest feasible method would normally be chosen.

Transducer → Computer Input.

In the next stage, we consider the link between the transducer and the computer. Here we are concerned with both the hardware connections and the software interface which manages the connections. For the first two methods, the scanner interface suffices, so the main point of interest is in the requirement in the third method for two input channels. In most cases, there would be little or no practical difficulty in satisfying the requirement.

For purposes of illustration, though, suppose that for some reason it was deemed essential that the input connection should be restricted to a single serial port. With this constraint, the third method would be more difficult to put into practice. In that case, we could choose to follow one of several courses in the design. For example :

· We could decide to abandon the third method.

· We could devise some sort of hardware multiplexer (not a difficult task, given the very simple signals involved) with which the two channels could be combined into a single serial stream – but that would incur the additional expense of complicating the coding at the computer input, and would require additional software in the interface to demultiplex the signals in order to reproduce the two original streams.

- We could accept the limitation, and give up the emphasis channel. We have failed to implement the system we required, but we have pinpointed the source of the failure, and could use this knowledge in future design activities.

Computer Input → Software.

The software has so far been (deliberately) left undefined, except for mention of a scanner. In practice, there will be more, if only to generate the required output, but the scanner is the first item in the sequence, and its requirements determine much of the detail of this step. A scanner works by changing the screen display at intervals, and it interprets each input signal as selecting the screen display current when it is received. When the selection is complete (usually after two input signals, for a conventional rectangular grid), the scanner generates an output signal which might be an ASCII character. How does this fit into our new system ?

With the first design, there is no difficulty; the scanner's function remains that of a simple character detector, and the emphasis signal remains in the same stream as the input text.

With the second design, there is some difficulty : how are we to deal with the double clicks ? Two questions must be addressed : how do we detect the double click ? – and what do we do with it once detected ? Conventional scanners are not equipped to perform either of these tasks. (They may allow for multiple clicks, but are likely to treat them either as accidents, in which case they are ignored, or as cancelling the first clicks, in which case they are useless for our purposes. It is very unlikely that they will be identified as a special case which we can trap and use in some way.) There are two ways to proceed. We can either write a new scanner which will identify the double clicks and present them as output, either as distinguishable symbols within the ordinary output stream or by sending output to an additional emphasis stream, or we can insert additional software between the input and the scanner to extract the double click. With this method, the original scanner can be maintained, but the additional emphasis stream reappears.

The third design is the easiest. We can continue to use a conventional scanner, and simply ignore the emphasis channel at this stage, leaving it to be handled by the next stage of software.

The rest.

What happens next ? More software must pick up the output from the scanner stage and do something with it. The details clearly depend on the form of output required, and further exhaustive elaboration would be tedious without adding anything of substance to the discussion. If the output is to be printed, the letters will be assembled into words, and the emphasis signals converted into whatever instructions are needed to display the emphasised text in (perhaps) italic; if the output is a speech synthesiser, an obvious analogous process can be carried out to add emphasis to the spoken phrase. In each case, the software used must be specialised to accept and interpret both text and emphasis streams, and to handle them appropriately. The significant conclusion is that the text and emphasis streams can be preserved throughout the process.

COMMENTS.

The example is contrived, but even so illustrates how the insistence on including emphasis affects the view of the system to be designed. The important point is that by taking explicit account of the requirement for faithful transmission of the human-form description, we have been led to the design of a system which can do it. Several different solutions to the design problem became evident, and their requirements in terms of hardware and software were made clear.

Why has this approach not commonly been adopted in rehabilitation computing ? I have no good answer to this question, but I speculate that the reason is simply that few people have asked for the full human-form communication. Instead of asking what people really want to do with computers, we have assumed that the requirements for rehabilitation systems are essentially the same as those for general computer use – but in the area of communication aids there is very little general computer use, so the assumption is not necessarily justified.

A DIAGNOSTIC APPROACH.

That discussion illustrates how informal analysis of a proposal keeping in mind specific requirements for the quality of communication can lead to constraints which must be observed when more formally specifying the system. The aim is design, and it is assumed that we know in general terms what we want.

Another point of view is possible. If the assumptions behind the approach are valid, then any failure in communication is the consequence of a mismatch between the capabilities of two components of the system, or of information loss of some sort within one of the components. To illustrate this view, I analyse a case in which a communication aid turned out to be unsatisfactory, and comment on the steps taken to remedy the defects observed.

Jinks, Young, and Henry[3] describe the case of Lee, who graduated from a simple low-technology communication board to a computer-based LOLEC (logical letter coding) system driving a speech synthesiser. Lee has physical disabilities which restrict his actions to pointing. He had originally used the communication board with great skill, choosing combinations of words to identify concepts by example – so he used "500 RUMMY" to mean "CARD GAMES". Using an encoding technique in which letter combinations were associated with prerecorded phrases, he rapidly became frustrated, and for any sort of demanding communication returned to the communication board, with which he was able to communicate more precisely and more rapidly. It is interesting to analyse this observation as a case of bad matching between items in the transmission sequence.

> *NOTE : In writing the previous paragraph, I have overestimated Lee's ability because of a misinterpretation of the original paper. In fact, Lee didn't use what I've later called the "grouping" channel7, in which classes are identified by giving some examples. I've nevertheless left the rest of the note unchanged because the idea of the "grouping" channel is interesting, and something like it is certainly used in some contexts; consider phrases like "Tom, Dick, and Harry", "nuts and bolts", "string and fencing wire".*

Consider the communication in which the combination of words was used :

JIM – GOOD – 500 RUMMY

The aim of the communication is to copy an idea in Lee's mind into someone else's mind :

$$\text{idea(Lee)} \rightarrow \text{idea(recipient)}$$

The medium of communication is, in all cases, a set of symbols, so we can depict the main features of the transaction in this way :

idea(Lee) → symbols → idea(recipient)

The system will work, at least potentially, provided that the symbols available are chosen so that Lee can encode his ideas effectively, and they can be decoded by the recipient. The vocabulary must therefore include ways to express the full range of ideas to be communicated.

The communication board.

With the communication board, the set of symbols contains 250 words (called *CBwords* below), and all detailed communication must be encoded in terms of this set. (Lee has also a limited repertoire of gestures, but there is no explicit mention of the use of this potential parallel communication channel in the article except for incidental activities.) In using his board, Lee must first express the idea which he wants to convey in terms of the words available, then point out the words by indicating their positions on the communication board. The recipient sees these actions, identifies the words, and then reconstructs the original idea. The essential features are :

idea (Lee) → CBwords (Lee)
→ CBpositions (Lee)
→ CBwords (recipient) → idea (recipient)

The main constraint is the restriction to the set of CBwords; a complete operational description of the communication would also include details of the word selection and so on, but these steps do not directly affect the transmission of the desired message. This is clearly an activity requiring encoding and decoding, so we can reasonably ask questions about these two processes, and about the code which is used.

Of these, the code is the most accessible. Consider the example : it expresses the intended meaning in "telegraphese", with little formal grammatical structure, and just about about enough semantic content in the selection of words from the CBwords to convey the desired message by implication. The effect of the constrained set of symbols is clear; "JIM – GOOD – CARDS" is apparently shorter, but presumably "cards" is not available on Lee's communication board, and to spell the word would require more selections and therefore take significantly longer.

How can we describe Lee's communication ? He is not simply communicating in words; the idea communicated is more than the sum of the words used. It is interesting to stretch the notion of communication channels, and to suppose that he is exploiting what could be described as a *grouping* channel, with additional meaning implied by the juxtaposition of selected words. This is comparable to the use of emphasis and prosody which was significant in the discussion above. Lee's grouping method is perhaps related to the "semantic compaction" principle behind Minspeak systems[4], with the difference that in Minspeak the association of ideas is used as a mnemonic device, while in Lee's case it is incorporated into the communication.

According to the preceding comments, a new communication channel must be associated with a corresponding vocabulary. In the case of the grouping channel, the "vocabulary" is the adjacent presentation of words which denote items which are both members of some class. It is appears occasionally in ordinary English, in phrases such as "Tom, Dick, and Harry" used colloquially to mean "everyone".

In exploiting the grouping channel, Lee used his own intelligence to select from the available symbols a suggestive combination of words which he expected would convey the message. In this expectation, he was relying on an equally powerful decoder – the human intelligence and experience of the receiver – to interpret the information carried by the grouping channel. (The importance of experience was made clear to me by my own first encounter with the description of Lee's method; I didn't understand his encoding immediately, because, while I could associate meanings with all the words, I didn't associate "500" with card games. On discovering that 500 was also the name of a card game, I was also able to interpret the grouping.)

Why should a "sentence" with so much missing be so clear ? The subtlety of the interpretation is illustrated by considering alternatives. There are several, because of the grammatical incompleteness of the sentence and the ambiguity of two of the words it contains. First, what else could Lee have meant ? It's hard to find anything convincing. "Jim plays 500 well, but there's something peculiar about it" ? That's a possible interpretation, but the other meaning is far more likely. "Jim is all right, but the other 500 are peculiar" ? That would only make sense in a context which resolved the reference to 500. "Jim thinks card games are good" ? That's possible, but less likely. Second, consider some variants. A change of one word in "Jim good 500 metre" transfers the emphasis to athletics; another – "Jim good Fred rummy" – becomes a character judgment. All in all, the meaning is remarkably precise.

We see that the encoding relies on much more than a predefined word – meaning equivalence. The nature of the word is important in defining what functions it can have in the utterance; the word order helps to associate words in specific ways; the basic meaning of each word further restricts possible combinations; and perceived relationships between word meanings (the two card games) can be significant in selecting between different interpretations.

Taken as a whole, this communication is an extreme example of the process which Creak and Sheehan[1] describe as adding information to a transmitted message. Only the words are transmitted, but we already know a lot about them. It works because the symbols chosen, though few, have a wealth of properties, both semantic and syntactic, and that these properties can be exploited both by the encoder and decoder, provided that both are able to identify the complex associations involved. While this decoding method, sometimes called *semantic parsing* , can be automated to some degree (for example, Demasco and McCoy[5] review possible techniques, and describe a system in which the aim is to reduce to the minimum the material which must be entered, relying on a semantic parser to expand the material into an appropriate sentence), attempts at automatic treatment inevitably come up against the "commonsense problem" – the requirement for a vast collection of knowledge about the world on which to perceive relationships between items. In any instance, there is also specific context : the conversational context in which the utterance is found, possible knowledge of Lee's interests and mannerisms of communication, and so on.

We can therefore account for the success of Lee's communication board technique, despite the restricted set of available symbols, by the excellent matching between encoder and decoder; while much detail is lost in encoding, the decoder has the right background to fill in the lost parts.

The initial LOLEC approach.

With the first LOLEC system, Lee could select one from a set of predefined messages (called *PDmessages* below) by entering letter combinations at a keyboard, and the messages were then spoken by a speech synthesiser. We can describe this process in terms comparable to those used in the previous example :

idea (Lee) → PDmessage (Lee)
→ letter combination (Lee)
→ KBposition (Lee)
→ PDmessage (device)
→ PDmessage (recipient) → idea (recipient).

The main differences from the point of view of conveying the intended meaning are the initial encoding by selecting a PDmessage rather than a set of CBwords, and the new task of encoding into a letter combination. Other changes presumably have negligible effects; Lee's selection of keyboard positions rather than communication board positions would be expected to make little difference provided that his physical dexterity was adequate for both tasks, and the recipient's task in comprehending a predefined message is likely to be much less demanding than that of interpreting Lee's selection of words. Difficulties arise, though, from both of the main differences.

The choice of letter combination was difficult because Lee couldn't remember the codes; for all that they were supposed to be "LOGICAL letter encodings", the logic depends on the spelling, and is compromised if your spelling isn't reliable. Indeed, there is a sense in which the method can act as an error amplifier; a spelling mistake in writing a sentence only rarely renders the sentence unintelligible, but a spelling mistake in a LOLEC system of this sort destroys the sentence entirely. Given a good memory for letters, the system will work well, but it is an essential prerequisite.

To eliminate this source of error, Jinks et al. provided a message directory, so that Lee could now select directly from a list of messages rather than havng to remember precise codes. Unfortunately, the cost of precise recall is slower recall; selecting the message now requires a search through a list, which is even slower. Lee, reasonably enough, became impatient.

On the choice of PDmessage rather than CBwords, two comments from the paper are revealing : Lee reverted to his communication board when he "had trouble finding the correct code or when he wanted to express something important"; and "his communication board message was more precise and was conveyed more rapidly". These remarks suggest that Lee had difficulty with the encoding step. Presumably this reflects the greatly reduced expressive power of predefined sentences over words; while the original article does not record the number of sentences provided, it is unlikely to approach the number which can be constructed from 250 words. (The reference to difficulties arising "after more than 30 codes were entered" suggest that the number of sentences might have been in the tens rather than the thousands.) The difficulty in expressing "something important" is also interesting; it is inevitable that a fixed repertoire of predefined sentences will not lend itself to the communication of any original ideas, and Lee would need the flexibility of his word-based communication board to say anything new.

What went wrong ? The computer-based system, while making it much easier for Lee to transmit well formed English messages, took away his freedom to construct his own word patterns, which, although grammatically far from English, were semantically much richer than grammatical sentences. The speech output available from the LOLEC system could be seen as a possible advantage, but it isn't much use if you can't say what you want. The meanings of sentences are too precise and complete; the meanings of words, in contrast, are moderately precise and not at all complete, and they can be bent and manipulated by their contexts. Lee had been exploiting this malleability in his grouping channel, but that was not available in the LOLEC system. Two sentences, however well expressed they may be, cannot interact in the same way as two words. How many sentences would you need before your set included "Jim is good at card games" ? And how could you compose that message from others in any plausible way ? So far as the analysis is concerned, a significant indication of this encoding difficulty is seen as the decrease in the difficulty of the recipient's decoding task. The recipient now merely understands an ordinary sentence, and is no longer an active participant in the transmission of ideas; the vast power of the original decoder – the human intelligence and experience – is not engaged.

It is perhaps necessary to emphasise that circumstances are important; this is not a condemnation of LOLEC systems. Both sorts of communication system have their places, but this type of LOLEC system turned out to be inappropriate for Lee's skills. Someone with less well developed motor skills than Lee, and therefore unable to select from a communication board with the same facility, could well find the same level of LOLEC system very satisfactory. The critical argument presented above is only valid if an alternative, and better, communications technique is accessible.

The second LOLEC approach.

The next step for Lee was to address one of his difficulties with the LOLEC system; the restriction to predefined messages was removed by replacing them with the words from his communication board, while retaining the LOLEC selection method. In view of Lee's proficiency at exploiting the communication board vocabulary to good effect, this could reasonably be expected to be a constructive development. The steps required are now :

idea (Lee) → CBwords (Lee)
→ letter combinations (Lee)
→ KBpositions (Lee)
→ CBwords (device)
→ CBwords (recipient) → idea (recipient).

This scheme restores the active engagement of the recipient, so that Lee can once again use his skill in using words, including the grouping channel – but in order for this strategy to be effective, the recipient must receive the words, and Lee's difficulty with letter combinations, already seen to some degree with the sentence selection, interferes with the chain of events before the recipient has a chance to cooperate. In fact, Lee had difficulty with the encoding stage (words → letter combinations), just as he did with the letter codes used to select the sentences. Presumably he could select the words well enough, as he was doing that with the communication board, but the task of encoding them was significantly more difficult than the direct selection he had used with the communication board.

This obstacle was something of a surprise, as Jinks et al. were aware that "Lee was able to recall some of the first letters of longer words", and they had expected that ability to be a good basis for this choice of strategy. It seemed, though, that the longish letter strings necessary to select from the comparatively large number of words were too confusing.

The third computer approach.

As the stumbling block was the process of encoding the words, the next step was clearly to revert to an encoding process which Lee could manage – the communication board. Instead of using the board itself, though, Jinks et al. transferred the words to a membrane keyboard, so that Lee's selections could be detected directly by the computer software and the speech synthesiser could still be used. The system can now be described in this way :

idea (Lee) → CBwords (Lee)
→ CBpositions (Lee)
→ CBwords (device)
→ CBwords (recipient) → idea (recipient).

The combination of Lee's skills and the recipient's skills can now be exercised as with the original communication board, with the added advantage of spoken output.

That description of the system is presented in terms of the membrane keyboard alone. In fact, the direct selection interface was provided in addition to the standard keyboard used for the LOLEC experiments, thereby building on Lee's existing skills while providing for future expansion – practically impossible with the original communication board – through some encoding technique.

COMMENTS ON THE EXAMPLE.

The example shows that it is possible to present a plausible interpretation of the results reported by Jinks et al. from the point of view advocated above. Hindsight is a wonderful faculty, though, and I have used it liberally.

The example does *not* show that the advocated approach could have been used effectively by Jinks et al. before they began their work with Lee in order to identify the problems and – perhaps – to achieve a satisfactory solution directly. The method is not sufficiently formal to be used in this way; it gives no guidance on how to analyse the communication process into steps (in developing the example above, I experimented with several different analyses), nor what sort of criteria should be used in describing the steps (what is "the vast power of the original decoder – the human intelligence and experience", and what does it mean ?).

DISCUSSION.

In this note, I have shown that a certain way of describing the flow of information through communication systems can assist in drawing up system specifications, and throw some light on the performance of such systems in practice. I can add, from my own experience in following this approach in other systems, that it has given me helpful insights into the relationships and interactions between the many activities which must operate at different stages of the communication process and at different levels of complexity.

The approach therefore appears to have some value as an informal descriptive technique. It is less obvious that it can be developed as a useful method for system design (which was the original intention) as well as for description. For such a development to be possible, it would be necessary to conduct the analysis as a much more formal process, with standards established for the identification of the steps in the communication, and for their description.

Further work in this direction will begin with careful analysis of a set of cases covering a wide range of communication techniques. The aim will be to establish standards as described above, and thereby to develop the method into a useful tool for designing communication systems.

REFERENCES.

1 : G.A. Creak, R. Sheehan : *The representation of information in rehabilitation computing,* Auckland Computer Science Report #54, Auckland University Computer Science Department, July 1991.

2 : E. Abadjieva, I.R. Murray, J.L. Arnott : "An enhanced development system for emotional speech synthesis for use in vocal prostheses", in Reference (6), paper number 1.2.

3 : A. Jinks, C. Young, K. Henry : "From communication boards to keyboards", *Communication Outlook* **13#2**, 4 (December, 1991).

4 : B. Baker : "Minspeak", *Byte* **7**, 186 (September, 1982).

5 : P.W. Demasco, K.F. McCoy : "Generating text from compressed input : an intelligent interface for people with severe motor impairments", *Comm.ACM* **35#5**, 68 (May, 1992).

6 : *ECART 2 : Proceedings of the European Conference on Rehabilitation Technology*, 26-28 May, 1993, Stockholm, Sweden.

7 : A. Jinks, in a letter to me, 8 June 1996.