TESSA, a system to aid communication with deaf people

Conference Paper · January 2002

DOI: 10.1145/638249.638287 · Source: DBLP

7 authors, including:

Stephen James Cox
University of East Anglia

Mike Lincoln
The University of Edinburgh

Judy Tryggvason
University of East Anglia

Mel Nakisa
Buckinghamshire New University

TESSA, a system to aid communication with deaf people

Stephen Cox, Michael Lincoln and Judy Tryggvason
School of Information Systems,
University of East Anglia,
Norwich NR4 7TJ, U.K.

Melanie Nakisa
Royal National Institute for Deaf People,
19–23 Featherstone Street,
London EC1Y 8SL U.K.

Mark Wells, Marcus Tutt and Sanja Abbott
Televirtual Ltd.
Anglia House,
Agricultural Hall Plain,
Norwich, U.K.

ABSTRACT
TESSA is an experimental system that aims to aid transactions between a deaf person and a clerk in a Post Office by translating the clerk's speech to sign language. A speech recogniser recognises speech from the clerk and the system then synthesizes the appropriate sequence of signs in British Sign language (BSL) using a specially-developed avatar. By using a phrase lookup approach to language translation, which is appropriate for the highly constrained discourse in a Post Office, we were able to build a working system that we could evaluate. We summarise the results of this evaluation (undertaken by deaf users and Post office clerks), and discuss how the findings from the evaluation are being used in the development of an improved system.

Keywords
Interactive systems, translation systems, Aids for the Deaf, speech recognition, avatars.

INTRODUCTION
There has recently been considerable research activity in developing automatic systems which can understand and output speech to provide information services or to perform transactions with customers [10]. Most of these systems have been developed for use over the telephone network with the goal of replacing completely or assisting a human operator [5]. A key aspect of them is that they operate in a rather restricted domain of discourse (e.g. train timetable enquiries [9], e-mail access [17] directory enquiries [18]) and this gives them some robustness to the difficult problems of variability and "noise" in the language used by the speakers, the speech signal and the telephony channel. There has also been work on interactive speech-to-speech translation systems e.g. [12]. These systems are designed to provide translation of conversational speech between languages with a potentially very large vocabulary [13]. We have been developing a system which combines aspects of both kinds of systems mentioned above. It is an interactive translation system but it operates in a very restricted domain and is designed to assist in the completion of a transaction between a Post Office (PO) clerk and a deaf customer. The system translates the clerk's speech into British Sign Language (BSL) and displays the signs using a specially-developed avatar. A comprehensive approach to the task of enabling humans who cannot sign to communicate using sign-language would clearly require the development of a general purpose speech to sign-language converter. This in turn requires the solution of the following problems:

1. automatic speech to text conversion (speech recognition);
2. automatic translation of arbitrary English text into a suitable representation of sign language;
3. display of this representation as a sequence of signs using computer graphics techniques.

However, recent research into "formulaic" language has shown that cross-language communication is possible using a limited set of pre-defined phrases, provided that the discourse between the participants is highly constrained in its topics and scope [16]. We chose to develop a system for use in a PO because many of the transactions are highly predictable and hence much of the associated language can be pre-defined. This enables us to sidestep or simplify many of the difficult problems mentioned above by defining a limited set of phrases that can be recognised and displayed in sign-language.

Although this imposes restrictions on what can be "translated", it is still likely to form a useful system because the task underlying the translation is a narrow one. Our aim in developing the current system was to provide a speech-to-sign translation system which had limited generality, but which could be used to accomplish transactions and hence whose benefit to the deaf community could be objectively evaluated. The information obtained by deploying and evaluating a simple working system should be invaluable in the development of more complex and sophisticated translation systems that may be of higher generality.

The system has been developed with the collaboration of Consignia (the UK Post Office), and research continues as part of the European Union fifth framework project, ViSiCAST [2], which aims to benefit deaf citizens by facilitating access to information and services by using sign language.

OVERVIEW OF THE SYSTEM

Design philosophy

Our goal was to develop a system to enable a Post Office counter-clerk to communicate with a deaf customer using automatically-generated sign-language, and hence to aid completion of a transaction. A priori, it might seem that recognising the clerk's speech and displaying it as text to the deaf customer would be adequate. However, for many people who have been profoundly deaf from a young age, signing is their first language so they learn to read and write English as a second language [4]. As a result, many deaf people have below-average reading abilities for English text and prefer to communicate using sign language [15].

Two variants of sign language are possible for communication with deaf people in the UK: British Sign Language (BSL) and Sign-supported English (SSE). BSL is a fully developed language with it own syntactical and semantic structures [3], whereas SSE uses the same (or very similar) signs for words as BSL, but uses English language word order [8]. Using pre-stored SSE "words" would enable sentences to be translated into sign language, but SSE is not popular with the deaf community and it is very important that the system is acceptable to deaf users. Another approach is to use whole BSL phrase units rather than words. This is possible only if a small number of phrases is required, and these phrases can be recorded in BSL rather than SSE. If recording of the signs is done correctly, phrases can be concatenated to a certain extent e.g. amounts of money can be slotted into a carrier phrase such as "The cost is...". Although this approach imposes considerable restrictions on the meanings that can be conveyed in BSL, and hence on the dialogue, the limited nature of the transactions in the Post Office should mean that most transactions could be completed in this way.

System components

Figure 1 is a diagram showing the structure of the system. The Post Office clerk wears a headset microphone. The speech recogniser is constantly active and responds when the clerk utters a phrase that matches a "legal" phrase from the grammar. The screen in front of the clerk displays a menu of topics available e.g. "Postage", "DVLA", "BillPayments", "Passports".

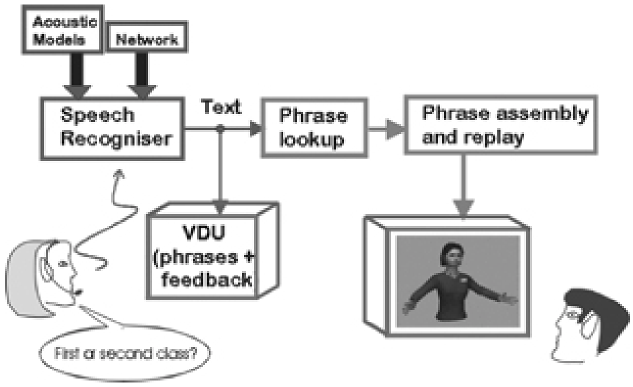

Figure 1

Speaking any of these words invokes another screen showing a list of phrases relevant to this category which can be recognised. However, this is only an *aide-memoire* to the clerk; all phrases are active (i.e. can be recognised) at any time, so that switching between categories is seamless. In trials, we found that the clerk could remember many of the most commonly-used phrases without consulting the screen.

Prior to designing the system, we obtained transcripts of recordings of PO transactions at three locations in the UK, in all about sixteen hours of business. Inevitably, much of the dialogue transcribed was in the nature of social interaction and had little to do directly with the transaction in hand. However, analysis of these transcriptions was essential for estimating the vocabulary which would be needed by our system to achieve a reasonable coverage of the most popular transactions. At the end of this analysis, a set of 115 phrases was prepared which we estimated should be adequate to cover about 90% of transactions performed. This set of phrases was changed and extended after trials with users (see the section on evaluation) and the total number of phrases currently available in the current system is about 370.

Speech recognition

Our original system used the Entropic speech recogniser. The recogniser requires a set of acoustic models for matching the input speech signal and a network that guides the search of the recogniser during recognition. The acoustic models can be adapted to the voice of each user ("speaker adaptation"), a process which takes about an hour, and the individual's models are then stored for later use. Speaker adaptation of the models greatly increases the recognition accuracy and hence the usability of the system.

The network constrains the speech recogniser to a finite number of predefined paths through the available vocabulary. These paths define the set of allowed phrases and consist of a start node (usually denoting silence, or background noise) followed by a number of word nodes or sub-networks, finishing with an end node (again denoting silence). Sub-networks are useful ways of defining phrase segments that can vary. For instance, a sub-network called “one2hundred” represents the legal ways of saying the integers between one and 100, and this can be inserted at any appropriate point into the network. There are other sub-networks called “amounts-of-money”, “days-of-the-week”, “countries” etc. The network can easily be changed, so that that phrases can be altered or added to the system without the need for any re-compilation. The recogniser can be operated on a “best-match” basis, so that a phrase which is phonetically “close” but not identical with a phrase in the network will be recognised as the latter. This allows some flexibility for the speech of the clerk. (For instance, the phrase “Put that on the scales, please”, which is not present in the network, would be recognised as “Please put it on the scales”)

Note that, because there is no separation of speech and language decoding in this system, it does not suffer from inaccuracies in the speech decoding process being forwarded to a language translation process that is also imperfect, an effect that can make more complex systems fail to translate correctly even quite simple phrases. By using pre-stored phrases, in effect we trade flexibility and range for accuracy.

The system described here is the first stage towards a more sophisticated system which will incorporate the techniques used in “speech-understanding” systems to enable a much wider range of transactions to be completed. In our current research system, we are experimenting with using a probabilistic language model recogniser followed by a language processor that attempts to map the output from the recogniser to the correct phrase. This has the benefit of allowing the clerk complete flexibility in what he or she says to the recogniser (as long as the words used are within the vocabulary of the recogniser) at the expense of requiring some language “understanding” to determine the correct sequence of signs to be output. At time of writing, we do not know whether this system will be less accurate than the system that uses a network. In addition, the system can obviously be adapted to translate to another spoken language (using either displayed text or pre-recorded/synthesised speech output) as well as to sign language, and this possibility is also being explored.

System software

The system software has the task of enabling communication between the speech recognition module and the avatar module and of controlling the overall progress of a transaction. The sign assembly system is written in TCL and the recognition module incorporated as a TCL extension. The avatar module is written in C++ and communication between this and the other system components is performed using a remote procedure call system via TCP/IP socket connections.

The signing avatar, TESSA

The simplest way of signing the set of phrases defined for the application would be to store video-recordings of a person signing each phrase and concatenate the appropriate phrases in response to the output from the speech recogniser. However, we have been developing an experimental system that uses a virtual human (avatar) to sign television subtitles [14]. A long-term goal is to produce a “text-to-sign synthesizer” that will be capable of synthesising signs from a much less restricted vocabulary and building such a system using concatenated video clips would not be viable. Another advantage of using an avatar is that different figures can be rendered onto the avatar's frame, so that a single set of recordings of signs can be used to drive different virtual humans. Conversely, multiple human signers can be used to generate the signed content of the system while using the same avatar for the output signing, making it easy to expand and update the signed content. In addition, concatenation of signing is more fluent and controlled for avatar than for video signing, as the exact positioning of the avatar can be manipulated. For these reasons, we decided to display the signs using an avatar, TESSA. Research into methods for capturing signing movements directly from video has been reported e.g. [1]. This approach is highly desirable as it obviates the need to record signs by attaching motion sensors to a human, with the attendant problems of invasiveness, motion restriction, calibration, sensor fusion etc. Unfortunately, capture from video is not yet robust enough to record high quality motion. The alternative is to capture signs using separate sensors for the hands, body and face. This technique appears to capture sufficient movement to generate true and realistic signing from a virtual human. The motion is captured as follows:

1. Cybergloves with 18 resistive elements for each hand are used to record finger and thumb positions relative to the hand itself.
2. Polhemus magnetic sensors record the wrist, upper arm, head and upper torso positions in three-dimensional space relative to a magnetic field source.
3. Facial movements are captured using a helmet mounted camera with infra-red filters and surrounded by infra-red light emitting diodes to illuminate Scotchlight reflectors stuck onto the face. Typically 18 reflectors are placed in regions of interest such as the mouth and eyebrows.

Figure 2

Figure 2 shows this configuration in use. The sensors are sampled at between 30 and 60 Hz and the separate streams integrated, using interpolation where necessary, into a single raw motion-data stream that can drive the virtual human directly. The system is calibrated at the beginning of each session but, in practice, the main variation lies between signers. For example, the considerable cross-talk between glove sensors depends heavily on how tightly the gloves fit. It is particularly important to ensure good calibration at positions where fingers are supposed to just touch the thumb and where hands touch both each other and the face. These positions are important to clear signing and, to reduce computation times, there is currently no collision detection to prevent body parts sinking into each other. Where individual signs or segments are to be added to the lexicon then signs are altered manually using a custom editor program, and marks at the beginning and end of each sign aid blending for concatenation.

The motion-data stream is displayed using an avatar (virtual human). In common with many avatars, a three-dimensional “skeleton” is driven directly from the motion-data. The skeleton is wrapped in, and elastically attached to, a texture mapped three-dimensional polygon mesh that is controlled by a separate thread (event loop) that tracks the skeleton. We use one of the latest PC-accelerated 3D graphics cards to render the resulting 5000 polygons at 50 frames/s using Direct-X on a Pentium class PC. Because TESSA is a full three-dimensional model, her position and pose can be changed by the user during use, an extremely valuable feature that enables users to select the optimal viewing angle and size. In addition, the identity of the virtual human can be changed.

EVALUATION

It is essential that the system conveys useful information in a way that is helpful and acceptable to deaf users. The extent to which TESSA met this aim was assessed using the following evaluation methods

1. Evaluation of the quality of the signs;
2. Evaluation of the difficulty of performing a transaction with TESSA;
3. Questionnaires to the deaf users and Post Office clerks.

In this section, we present a summary of the results of these evaluations.

Six pre-lingually profoundly deaf people whose first language is BSL and three PO clerks took part in the evaluations of the system. The evaluations took place over three sets of two days. Two deaf people and one clerk attended for each pair of days. The first day started with completion of the first part of a questionnaire. Each deaf participant then alternated between identifying a block of signed phrases and attempting a block of staged transactions. At the end of the second day, all participants completed the remainder of the questionnaire and gave any general feedback. BSL/English interpreters were present throughout.

Quality of signing

The quality of TESSA's signing was measured in two ways: intelligibility of signs, and acceptability of signs to deaf users. The first of these measurements is an objective one and is clearly important in establishing a baseline for the current system against which future avatars may be evaluated. However, it is well known that intelligibility on its own is inadequate for assessment of these systems: for instance, synthetic speech can sound fully intelligible but be disliked by users [6]. Hence we also included a subjective measurement of “acceptability” of signing.

Intelligibility

The deaf participants were presented with each signed phrase and asked to write down what they understood. From the 115 distinct phrases, 133 phrases were generated by incorporating days of the week and numbers to ensure that each day and each number (units and tens) was presented at least once. Signed phrases were presented on the screen without text. The deaf participant could control presentation of each phrase and was allowed to repeat each phrase up to a maximum of five presentations. Phrases were presented in blocks of between 20 and 24, in groups according to broad categories, for example, postage, bill payment, amounts of money. Accuracy of identification of phrases was assessed in two ways:

1. By the accuracy of identification of complete phrases;
2. By the accuracy of approximate “semantic sign units” within the phrase.

For example, the phrase “It should arrive by Tuesday but it's not guaranteed” requires five sign units, so “should arrive Tuesday not guaranteed” would score 100% and "should arrive Tuesday" 66%. The 133 phrases gave a total of 444 sign units. While these units were not all distinct (for example, the sign for "pound" was presented several times), identification of each presentation of a unit was scored separately. An experimenter judged the accuracy of responses for both measures on the basis of written responses from each deaf participant. Once each phrase had been scored for accuracy of identification, each deaf person was re-presented with each phrase not identified correctly along with the text of the intended phrase. With an interpreter and experimenter, they were asked to indicate whether the signs were considered inappropriate or whether they were just not clear. Any signs considered inappropriate were not necessarily wrong; rather they may have represented different regional variations in sign to those used by the deaf participant. Variation in signs is a more difficult problem to contend with than variations in accent or dialect in spoken languages, as hearing people can use a standard written language as a reference, which is not available to those who communicate using only signs [8]. The average number of times each phrase was presented before an attempt at identification was made was 1.8. Attempts at identification were made after one presentation for the majority of phrases (51%) and required more than two presentations for 20% of phrases. Figure 3 shows that the average accuracy of identification of complete phrases was 61% and ranged from 42% to 70% across deaf participants. For the identification of sign units in phrases, average accuracy was 81% and ranged from 67% to 89%.

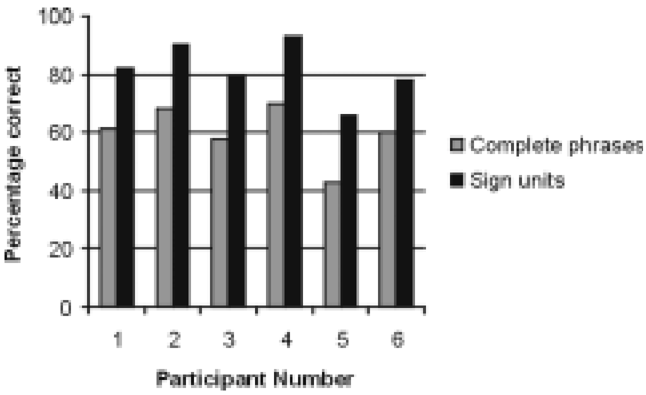

Figure 3

Subsequent analysis of the sign units which were wrongly identified indicated that on average 30% of errors (6% of all sign units) were due to signs considered inappropriate and the remaining 70% (13% of all sign units) were due to unclear signing.

Acceptability

Participants were asked to rate how acceptable the phrase was as an example of BSL on a 3-point scale (1="Low" 2="Medium", 3="High"). Table 1 shows the percentage of phrases that were rated in each category of acceptability. The average acceptability rating was 2.2 and ranged from 1.7 to 2.8.

Acceptability Rating		% of phrases
High	3	20.2
Medium	2	43.2
Low	1	36.6

Table 1

Discussion

Accuracy of identification of the signed phrases was 61% for complete phrases and 81% for sign units, with quite a wide range in accuracy across deaf participants (ranges of 28% and 20%, respectively). This range inaccuracy suggests it is important to use many sign-language users for a true assessment of signed content of these systems. In future, it may be more appropriate to use more than six deaf people from a range of UK regions to assess sign quality. The majority of identification errors (70%) were due to signs being unclear rather than due to inappropriate signs. The percentage of errors for inappropriate signs did not differ greatly between subjects, with personal averages ranging from 4.7% to 6.6%. This pattern might suggest that the same signs were considered inappropriate by all deaf participants. However, inspection of the pattern of errors across deaf participants for each phrase indicated that this was not necessarily the case. Of the 46 phrases where one or more sign was considered inappropriate by any deaf participant, in 34 (74%), a sign was considered inappropriate by no more than two of the deaf participants. This result suggests that regional variations or differences in personal signing style may have played a role in phrase intelligibility. Ratings of acceptability were also given across the scale with 20% of phrases rated as highly acceptable and 63% in one of the top two categories, indicating that there is scope for improving the quality of the avatar's signing.

Transactions

Staged Post Office transactions were used to compare completion times and ease and acceptability of communication with and without TESSA. Each deaf participant attempted 30 transactions with a single Post Office clerk. Transactions were selected by the Post Office as those achievable with the phrases available. There were 18 distinct transactions: 6 were denoted "simple", 6 "average difficulty" and 6 "complex". The average difficulty and complex transactions were attempted twice by each deaf participant/clerk pair, once with an open counter and once behind a fortified counter where a transparent screen separates clerk and customer. Use of different counter styles did not appear to affect performance hence results are not reported separately here. Half of all transactions were attempted with TESSA and half without. The phrases presented with or without TESSA were counter-balanced between deaf participants. Practice transactions were performed with TESSA at the start of each session so that the clerk, deaf participant and interpreter could get used to using TESSA and the format of the evaluation. Transactions were performed in blocks of six, three with TESSA and three without. The approximate time taken to successfully complete each transaction was recorded. On completion of each transaction, both deaf participants and clerks were asked to rate each transaction for acceptability on a 3-pointscale from 1—"Low" to 3—"High".

Timings

Figure 4 shows the results from these experiments. Errorbars show the 95% confidence intervals of the mean transaction times. On average, transactions took longer to complete with TESSA than without [$F(1,178)=61.2$, $p<0.001$]. Average times for transactions were 57s without TESSA and 112s with TESSA.

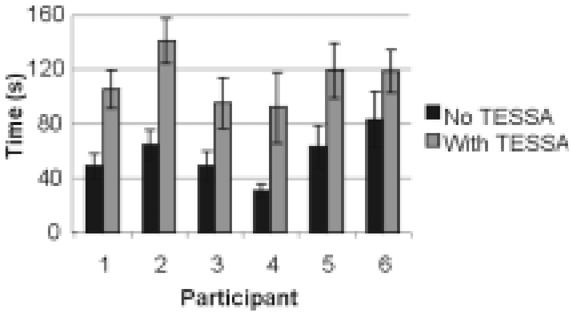

Figure 4

Acceptability

The deaf participants rated acceptability of transactions completed with TESSA as slightly lower than transactions completed without TESSA. On a 3-point scale of acceptability (1="Low" 2="Medium", 3="High"), average ratings were 1.9 with TESSA and 2.6 without. The corresponding figures for the clerks were 2.5 and 2.6.

Discussion

Compared to transactions without TESSA, transactions performed with TESSA took on average nearly twice as long to complete, and the deaf participants, and to a lesser extent the clerks, rated communication as less acceptable. The main reason most likely to have contributed to these effects was the somewhat disjointed communication with TESSA. As expected, it took the clerks some time to learn which phrases were available and to locate the phrase they wanted so they could read it out word for word. The clerks had only about an hour of practice using the system before the trials. These difficulties should decrease substantially with training and experience on the system. Moreover, the next version of the system, which will incorporate some speech "understanding", will not require phrases to be repeated verbatim. Additional factors may have contributed to the longer transaction times and poorer ratings with TESSA:

1. The transactions covered by the system were drawn from the most commonly occurring transactions in the PO, for example, buying stamps, cashing a cheque. The transactions used in the trials therefore tended to represent situations in which communication was fairly easy without TESSA.
2. The deaf participants were not truly representative of a cross-section of the deaf community in that they were all fairly good communicators and all had reasonable written skills. Hence they were able to complete the simple transactions, by lip-reading/speaking and writing notes or asking the clerk to write things down where necessary.
3. The clerks either were "deaf aware" or soon became deaf aware as a result of spending two days with the profoundly-deaf participants. Communication without TESSA was fairly easy as they used good eye contact, spoke clearly and were prepared to write things down if they were not understood.
4. There was a delay of a few seconds between recognition of the spoken phrase and the signing of the phrase. Not only did this add to the overall transaction time but the delay often resulted in loss of attention and the need for the sign to be repeated or the clerk to repeat the phrase.

Questionnaires

Questionnaires to both deaf participants and clerks were used to obtain subjective views of previous experiences of communication in the PO, and how these experiences differed in the trials and were anticipated to differ in real life using TESSA.

Deaf participants were asked about ease of communication in the PO, previously, in the trials with TESSA, and anticipated in everyday life with TESSA. Responses indicated that four of the six thought that communication with TESSA was "Manageable", including one participant who usually found communication in the PO "very difficult". However, two participants said they thought communication with TESSA was "Very difficult". All clerks said communication was "Slightly easier" or "Much easier" with TESSA than without, and that in everyday life they anticipated that communication would be "Much easier" with TESSA. In addition, all clerks said that they would prefer to have TESSA available as an option to use when communication became difficult, even though they all thought transactions would take "Slightly longer" with TESSA.

Discussion of evaluation results

The deaf participants provided much constructive feedback about how TESSA could be improved. Their main points were:

1. Facial expressions need to be improved. Clearer handshapes, finger configurations and lip patterns are required, especially for numbers and finger-spelling.
2. The delay between the end of the spoken phrase and the beginning of signing needs to be reduced.
3. The appearance of the avatar needs to improve. In particular, a clearer distinction should be made between the face and hands and the clothing, which should be plain.
4. All deaf participants said they would prefer to see both BSL and text rather than just BSL or just text. They also thought that SSE should be available as an option.

When asked to comment on the use of avatars for signing in general, all deaf participants thought that avatars would be most useful for more complex communication needs, e.g. explaining forms to claim social benefits.

The fact that the responses from the deaf participants were not more generally positive does not seem unreasonable at this stage in the life-cycle of the project. To the best of our knowledge, this is the first time that avatar technology has been used to produce sign language, and we are encouraged by the levels of intelligibility and acceptability of the signs reported in the evaluation.

All clerks said they would prefer to have the system available as they thought it would make communication with deaf customers easier and more effective. Use of the system for multiple spoken languages and with text sub-titles would ensure more frequent use and hence greater likelihood that the system would be used with deaf people. The clerks also commented that they would like more phrases and an "unconstrained" speech system, where phrases need not be spoken verbatim.

General comments and future work

Our goal in developing this trial system was to establish whether the introduction of a limited speech-to-sign translation system for the PO counter clerk would be beneficial to deaf users whose primary means of communication was sign language. Although some of the feedback from the evaluation was critical, we are encouraged by the following points:

1. Some deaf participants who said that communication in the Post Office "usually upset or worried them" said they thought using TESSA in the Post Office would not bother them at all. Most said they would prefer to have TESSA available in the Post Office for use when communication became difficult.
2. Feedback from the Post Office clerks was generally very positive, despite the very limited time they had to train with the system.

These evaluations, although limited in extent, have indicated that there is much scope for improvement of TESSA, have given some insight into how these improvements could be achieved and provided baseline outcome measures against which improvements can be assessed. The majority of aspects identified for improvement are planned for further development within the ViSiCAST project. Primarily, the development of an "unconstrained speech" system, where phrases need not be repeated word for word, will enable much more natural communication and should greatly reduce the time taken for transactions, so is also likely to be more acceptable to both deaf customers and clerks. Other aspects to be explored include research into facial modelling, which should improve avatar facial expressions and lip patterns. New data gloves are also being used to improve recording of finger movements and hand shapes. New models of the avatar and clothing will also take account of the comments made by the deaf participants. Less formal evaluations are planned within the deaf community to assess the views of more deaf people and further formal evaluations will continue through the lifetime of the ViSiCAST project. In tandem with these developments, the ViSiCAST project is also doing basic research into the general problem of converting arbitrary English text into a representation of sign language [11], and developing a synthetic avatar that can sign these representations without the need for motion capture [7]. These will feed into the application described here to increase its flexibility and sophistication. The problem of two-way communication is also being addressed by research into sign-language recognition.

ACKNOWLEDGMENTS

The development of TESSA was sponsored by Consignia (the UK Post Office) and the evaluation by the European Fifth Framework Programme under the auspices of the ViSiCAST project. We would like to thank the deaf participants, PO clerks and interpreters for their help in the evaluations.

REFERENCES

1. T. Ahmad, C. J. Taylor, A. Lanitis, and T. F. Cootes. Tracking and recognising hand gestures, using statistical shape models. *Image and Vision Computing*, 15(5):345 – 352, May 1997.
2. J.A. Bangham, S.J.Cox, M. Lincoln, I. Marshall, M. Tutt and M. Wells. Signing for the deaf using virtual humans. In *IEE Colloquium on Speech and Language Processing*, April 2000.

3. D. Brien. *Dictionary of British Sign Language / English*. Faber and Faber, 1992.
4. R. Conrad. *The deaf school child*. Harper and Row, 1979.
5. R. D. Johnston et al. Current and experimental applications of speech technology for Telecom services in Europe. *Speech Communication*, 23(1–2):5–16, 1997.
6. R. D. Johnston. Beyond intelligibility—the performance of text-to-speech synthesisers. *BT Technology Journal*, 14(1):100–110, January 1996.
7. J. R. Kennaway. Synthetic animation of deaf signing gesture. In I. Wachsmuth, editor, *4th International Workshop on Gesture and Sign Language Based Human--Computer Interaction*, Springer series Lecture Notes in Artificial Intelligence. Springer Verlag, 2001. To be published.
8. J. Kyle and B. Woll. *Sign Language: The Study of Deaf People and their Language*. Cambridge University Press, 1985.
9. L. F. Lamel et al. The LIMSI RailTel system: field trial of a telephone service for rail travel information. *Speech Communication*, 23(1–2):67–82, 1997.
10. B. Mazor and B. L. Zeigler. The design of speech-interactive dialogs for transaction-automation systems. *Speech Communication*, 17:313–320, November 1995.
11. E Safar and I Marshall. The architecture of an English-text-to-Sign-Languages translation system. In G Angleova et al., editors, *Recent Advances in Natural Language Processing (RANLP)*, pages 223–-228, September 2001.
12. W. Wahlster, editor. *Verbmobil: Foundations of Speech to Speech Translation*. Springer-Verlag, 2000.
13. A. Waibel. Interactive translation of conversational speech. *Computer*, 29(7), 1996.
14. M. Wells, F. Pezeshkpour, M. Tutt, J. A. Bangham, and I. Marshall. Simon - an innovative approach to deaf signing on television. In *Proc. International Broadcasting Convention*, pages 477–482, September 1999.
15. D. Wood, H. Wood, A. Griffiths, and I. Howarth. *Teaching and talking with deaf children*. John Wiley and Sons, 1986.
16. A. Wray. The functions of formulaic language: an integrated model. *Language and Communication* 20 (2000) 1–28.
17. P. J Wyard et al. Spoken language systems—beyond prompt and response. In F. A. Westall, R. D. Johnston, and A. V. Lewis, editors, *Speech Technology for Communications*, pages 487–-520. Chapman and Hall, 1998.
18. O. Yoshioka, Y. Minami, and K. Shikano. A speech dialogue system with multimodal interface for telephone directory assistance. *IEICE ransactions on Information and Systems*, E78D(6):616–621, 1995.