Language Technology and Patients with Limited English

Harold Somers
Centre for Computational Linguistics
UMIST, Manchester
Harold.Somers@umist.ac.uk

1 Introduction

This paper describes a proposed framework for the use of language technology to provide computer-based help for patients with limited or no English. Aimed at users of the Health Services who are disadvantaged by their (lack of) linguistic skills, the system will assist the patient in different ways at different stages of their interactions with health-care providers. In its full conception it will embrace a wide range of NLP technologies. Although the research is based on the UK model of health-care provision, there are clear messages for anyone interested in language technology and under-resourced languages, whatever the application.

Focusing on the GP's clinic, it will provide a kind of FAQ help-desk and act as a kind of Receptionist to help determine whether the patient needs to see the GP or some other health-care specialist. If a GP consultation is indicated, the computer can be used for history note-taking. During the consultation itself, it can act as a mediator between the doctor and patient. Afterwards, in help-desk mode again, it can help the patient understand the diagnosis, any tests needed, and the proposed treatment regime.

We propose in the first instance to develop systems aimed at Urdu- and Somali-speaking patients, focusing on respiratory problems (e.g. asthma).

2 Patients with Limited English

In many parts of the UK and other "western" countries, there are recent or long-term immigrants, refugees, and asylum seekers and other people whose command of English, while often adequate for day-to-day activities such as shopping and other domestic chores, is not sufficient for more formal situations such as interactions with health services, especially visits to their GP. There is no shortage of literature reporting disparities in health, health-care, and social care provision in these communities and communication difficulties are identified as a major factor (e.g. McAvoy and Sayeed 1990; Chalabian and Dunnington 1997; Woodhead 2000). The problem is also well recognised in other countries (e.g. USA, Uba 1992; Jackson 1998).

People in this situation will only rarely be lucky enough to find a homolingual GP, and even then will still have to communicate with other persons (receptionist, community nurse, pharmacist, specialist). Some may take with them an "interpreter", typically a family member or someone from their religious community, or else will just "muddle through". The outcome is undesirable in either case, for numerous reasons. In a recent systematic literature search of a range of medical and social science journal databases since 1990, on barriers to accessing health-care experienced by refugees in the UK (Jary 2001; Hays 2002), language difficulties were identified as the largest single barrier to care and as such repeatedly identified as a major concern for refugees (e.g. Tang and Cunninghame 1994; Lam and Green 1994). In a study in London, 53% of GPs felt that language difficulties were a problem (Ramsay and Turner 1993). Effective communication is important in all areas of health care (Voelker 1995), from finding out about services available through to complying with treatment.

There have been only a few suggestions for initiatives to tackle this problem (reviewed in Jary 2001), including a cheap national specialist medical telephone interpreting service, with hands free conferencing to enable concurrent discussions and examination if needed (Jones and Gill 1998; Wolmuth 1996), and multilingual phrase cards for use by health-care practitioners and receptionists (simple words like days of the week could make a significant difference to people trying to access health care). Further initiatives urgently need to be developed.

There can be no doubting the importance of doctor–patient communication, which has for many years been the focus of medical attention. Everything in medical practice arguably derives from the consultation, during which the doctor must acquire and impart information, and set up a relationship with the patient; the consultation itself can also have a therapeutic role. Valuable consultation time can be saved by having the patient complete a pre-consultation questionnaire which allows information to be expressed which may be given reluctantly in a hurried interview. There is a considerable literature on the structure of the consultation, from various angles including the linguistic, pragmatic, ergonomic, social and of course medical aspects.

Use of computers in the doctor–patient consultation paradoxically has been recognised as both potentially detrimental and potentially hugely helpful. The early use of computers on the consultation desk was seen as a threat, detracting from interaction with the patient, reducing eye contact and rapport build up. More recently the help of computers to increase communication and rapport has begun to be recognised. Computers can help in accessing records of other interactions, reducing the need for repetition. A recent systematic review of UK literature in the 1990s (Mitchell and Sullivan 2001) concluded that Primary Care computing systems can improve practitioner performance, particularly for health promotion interventions.

3 A Technology-based Solution

As mentioned above, it is proposed to introduce various "modes" of technology to alleviate the situation. The most intricate of these is during the consultation itself, where we envisage a kind of interactive phrase-book, designed to run on the typical PC that might be found on a GP's desk.

At the core of the system is a hybrid multi-engine embedded MT system: essentially an EBMT system with a "translation memory" (TM) extracted from corpora of doctor–patient interviews, supplemented with a simple rule-based MT (RBMT) system and a word-by-word lexical look-up facility. It will have a highly flexible interface: a simple set-up like in a chat-room, where each user types at a keyboard with the results shown on a split-screen is not practical when one of the users may not be a regular computer user.

The system in this mode has two users: the doctor and the patient, with significantly different profiles of computing experience. Accordingly, the user-interfaces will be quite different for the two users, while necessarily being integrated. Whereas the doctor can be expected to use the keyboard and mouse, and be comfortable with a sophisticated GUI, the patient's interface presents a number of problems. Both speech- and text-based interfaces are proposed.

It should also be remembered that some patients will not need to use the system for every part of the interview, their English being sufficient for some interactions. In addition to the "Consultation mode", we will simultaneously develop a "Reception mode" with an interactive FAQ/help system and a "History mode" involving a computer-aided patient interview system.

In the following sections, we give some more details about the design features of the different modes of the proposed system.

3.1 Multi-engine MT system

MT has now proved itself viable under conditions of restricted input and interactive use. Particularly effective is an architecture which tries various strategies in parallel and then tries to reconcile the results. This is the "multi-engine" approach seen in the PANGLOSS and DIPLOMAT systems (Frederking et al. 1994, 1997). The engines that our system will use will be an EBMT/TM system (Somers 1999), a rule-based transfer system (Trujillo 1999:121ff), and a simple lexical look-up system; it is to be expected that the input from the doctor will usually go through the EBMT system, while the patient's input, being more varied, may more often be translated by RBMT or on a word-by-word basis. In the proposed scenario, it is an example of an "embedded" MT system (cf. Van Ess-Dykema et al. 2000)

EBMT is akin to case-based reasoning (CBR) (Kolodner 1993; Somers and Collins 2003) in that new translations are composed on the basis of past translations, as provided by the "example base" of utterances taken from a corpus of doctor–patient interviews, manually translated into the target language. This method gives a very high quality of translation when the input can be matched against an appropriate example. The match does not have to be exact: as in CBR, a partial match can lead to a successful outcome.

RBMT and word-by-word translation methods tend to result in more stilted translations, closely following the syntax of the source language. In our scenario, this is more likely to be used for translating the patient's replies into English: thus the burden of understanding a less polished translation will normally fall on the doctor, who will gain experience of the system with use, and – on the evidence of early users of less sophisticated MT systems (cf. Church and Hovy 1993) – will quickly get used to its quirky style.

The notion of "restricted input" relates to the widely accepted notion of "sublanguage"-based approaches to MT (Kittredge and Lehrberger 1982), especially inasmuch as a corpus can help to define the sublanguage (cf. Deville and Herbigniaux 1995; McEnery and Wilson 1996:147ff; Sekine 1997).

The experience of the DIPLOMAT project (Frederking et al. 1997) is especially relevant to this project, since their system was developed specifically with rapid development of new language pairs for use in a dialogue situation between an experienced user and a naïve interviewee who may have little experience of computers, and may not even be literate. Versions of DIPLOMAT have been developed for English–Croatian and English–Haitian Creole, for

use in the field to allow English-speaking soldiers on peace-keeping missions to interview local residents.[1] An additional feature of DIPLOMAT is the use of speech-recognition and synthesis front and back ends, and the extensive use of on-screen interactive correction by both participants. As the language pairs indicate, it has been tested in the former Yugoslavia, and in Haiti. The success of the DIPLOMAT project gives a strong indication of the viability of the current project.

3.2 Corpus of doctor–patient interviews

Transcribed corpus data from doctor–patient interviews is readily available in the British National Corpus, which contains about 100 examples of short (300–900 words) medical consultations in GP surgeries or hospitals, already annotated for POS tags and some other aspects. Several other similar corpora have been collected, e.g. by Thomas and Wilson (1996), Wynn (1999). Other researchers have collections of tape-recordings,[2] and there are even conferences dedicated to the analysis of doctor–patient discourse.[3] Data from consultations where an interpreter was present may also be relevant: a small amount of such data has already been made available to us (Cambridge 1997).

This corpus will serve multiple purposes, and accordingly we should distinguish various of its characteristics. For example, transcriptions of interpreter-mediated interviews, and interviews where the patient has a poor command of English, will be useful as an indication of how such interviews tend to proceed. They will not however serve as a direct model for the system, which aims to bypass some of the difficulties that arise in such situations. For most of our purposes, what is important is not so much the verbatim transcripts, but the model of the discourse and the examples of the kinds of things that are said (cf. Passonneau and Litman 1997; Berthelin et al. 1999). This being the case, the utterances in the corpus can legitimately be "cleaned up". The corpus will be marked up, especially for dialogue function in a TEI-conformant manner.

Another purpose of the corpus is to provide a source of examples for the EBMT system, and so a parallel target version will have to be provided. It will also serve as a "training corpus" (Leech 1992) for the development of the translation lexicon and the RBMT system. To some extent, some of this linguistic information can be extracted semi-automatically (see for example Brent 1993; Smadja 1993; Melamed 2000; Véronis 2000).

3.3 Dialogue model

An important element in Dialogue MT (Boitet 1993, 1999) is a model of the dialogue. The dialogue model can be used for the translation component to reduce the search space for the EBMT system (cf. Somers et al. 1994), and to inform target-word selection in RBMT and word-by-word translation (the correct translation may depend on the context, i.e. the discourse function of the utterance, cf. Somers et al. 1990; Somers and Jones 1992). The model also plays a role in the interface, simplifying and determining the options offered in the menu-driven mode for both doctor and patient (cf. Alm et al. 1989).

3.4 The doctor's interface

Doctors greet and observe patients in all doctor–patient encounters, and in the UK the consultation proceeds normally these days in the presence of a computer which is used for recording all personal details, history taking of a problem, diagnosis, and treatment. Thus it is a small step to consider the possibility of using a computer to aid communication as part of the existing situation.

For the doctor's interface, three main possibilities are envisaged: typing at the keyboard, augmented by auto-completion; or a menu-based approach, enriched by dynamic domain knowledge; finally, and more ambitiously, a speech interface is also planned.

The menu-based interface, which is also appropriate for the patient's interface, involves "intelligent" menu-driven selection. Several script- or frame-based interfaces have been reported, for example the UNICORN system (Dye et al. 1997; Iwabuchi et al. 2000), which is specifically aimed at multilingual communication, DRAFTER (Hartley and Paris 1997), for multilingual document preparation, Floorgrabber (Alm and Arnott 1998) and Frametalker (Higginbotham et al. 2000) for users with communication difficulties. The "intelligence" derives from domain knowledge and a discourse model which permit the interface to be simplified by determining the options

1. See also http://www.avt-actii.lmowego.com/
2. See for example ww2.mcgill.ca/ Psychiatry/ transcultural/ prmary.html
3. For example, the Conference on Medical Interaction, 18-20 October 2000, at the University of Southern Denmark, Odense. See http://www.conversationanalysis.net/Conferences/Medcal/program_doc-pat.htm.

offered. This type of interface is most appropriate when the consultation is following a predictable course, and "standard" questions or comments are being made, for example "How long have you had this problem?".

In the keyboard-based typing interface, the doctor simply types the input, or parts of it that the patient does not understand. Typing is aided by auto-completion proposals based on the corpus, an idea already demonstrated by Langlais et al. (2000) with their TRANSTYPE project. Typing is necessitated when what the doctor wants to say is not sufficiently similar to anything that the menu-driven interface is offering, for example a much more specific question or comment which relates to things the patient may have said earlier, e.g. "When did your step-mother pass away?".

3.5 The patient's interface

Some patients will be highly experienced in using computers while for others, a keyboard- or mouse-driven interface may not be appropriate. Therefore, a range of interfaces must be made available to the patient. We can include simple interfaces like a drop-down menu, as in the doctor's interface. If the patient's language involves a different character set (as is the case with Urdu), it is not viable to assume the patient might want to use the keyboard: character-handling of non-Roman writing systems is not a problem as such (and is necessary for output), but we cannot assume that the patient can quickly learn to use an Urdu keyboard, or, worse still, to learn a set of mappings from a QWERTY keyboard. The problem may be less acute for patients whose language uses the Latin alphabet. All these issues represent an important and innovative aspect of the research proposed here: we need to discover the best way to integrate all the possibilities so as to provide an interface that both doctor and patient are comfortable with, that promotes an equitable exchange (rather than giving one or other user excessive control), and makes best use of their respective skills and experience. There are important socio-cultural issues here which we cannot address fully in this paper

Of relevance here is the field of Augmentative and Alternative Communication (AAC) and in particular the work on picture-based communication (PBC) interfaces (Blenkhorn 1992; Loncke et al. 1999). AAC is usually focused on disabled users, and AAC techniques have apparently not been applied to users whose only "handicap" is lack of a shared language.[4] Langer and Hickey (1999) report on growing contacts between the AAC and NLP research communities, as evidenced by dedicated workshops (e.g. Copestake et al. 1997) and journal editions (Langer 1998). The work of Grisedale et al. (1997) is also of interest: they developed a GUI for healthcare workers in rural India, like us facing the problems of inexperienced computer users and a non-Roman writing system. HCI issues are of paramount importance here: robustness and flexibility are essential; alternative modes of input, such as touch screens, may be preferred, since the patient may lack experience of mouse manipulation. Johnson (2003) is currently investigating how suitable PBC is for this application, whether paper-based on computerized. An important issue is the cultural as well as linguistic suitability for non-English-speaking non-disabled users of symbols developed for English native language disabled users.

3.6 "Reception mode": FAQ/Help facility

Consultations often include obtaining answers to the same series of questions (such as how long has the problem been continuing). This may lend itself to identification of a series of frequently asked questions. It is easy to spend too much time during the doctor–patient consultation on routine symptomatic inquiry and too little on observing the patient and listening to their spontaneous talk. One possibility therefore is to transfer some of the functionality of the system to a pre-consultation computer-mediated help-desk and interview (cf. Osman et al. 1994). By "help-desk", we mean a simple on-line interface containing potted texts in answer to frequently asked questions (FAQs).

These interfaces can be run with a simulated natural-language interface as is the case for example with Microsoft Word, which is actually based on key-word matching. This could be installed on a computer terminal in the Health Centre reception area, so that potential patients could get relevant information without even making an appointment with the GP. There has been a considerable amount of relevant work in this area, notably on Tailored Patient Information (TPI) systems (Buchanan et al. 1995; Cawsey et al. 1995; Reiter and Osman 1997). Navigation of the help facility can be system-led or patient-led. In the latter case it would work in much the same way as the help facility in, say, a word-processor offers "Type in your query here". In the former case, the user is lead through the interaction with a structured database depending on the choices made at each point. Different start points might relate to basic symptoms (answering the question "Do I need to see the doctor?"), general procedure ("What can I expect when I go to the hospital?") or, after diagnosis, what the course of treatment involves, e.g. general informa-

4. Personal communication: Pat Mirenda, editor of the journal *AAC Augmentative and Alternative Communication*. See also Johnston (2003).

tion about the drugs or therapy that have been prescribed, and the likely outcomes and progress of the patient's condition. The system can have an even more rudimentary use, since many patients come to the health centre with non-medical problems, such as questions about housing benefit, and they can be redirected to the appropriate agency.

3.7 "History mode": Computer-mediated interviewing

Many services in general are finding it helpful nowadays to gather basic information from the patient prior to meeting with the professional. This is the important element of "history" note taking which can be partly accomplished using computer-mediated interviewing techniques, which can make better use of the time the patient spends in the waiting room. These widely-used techniques have been found to be particularly useful in sensitive applications like taking patient's medical details (Lilford et al. 1985), where decreased time pressure leads to fuller responses, especially when questions are of a sensitive or embarrassing nature. Most systems are based on flexible multiple-choice questionnaires, while the use of free text (e.g. Peiris et al. 1995) is more complex, and brings us into the area of conversation systems (cf. Wilks 1999). An on-line consultation might be appropriate in the case of patients returning with chronic problems.

4 Conclusion

We have presented here a proposal for a highly innovative multi-modal system. While plan-based communication or authoring tools have been proposed previously, the multilingual profile coupled with the dialogue situation for the doctor's and patient's interfaces is quite novel. The application of AAC techniques to use by non-handicapped but linguistically disadvantaged users is likewise a new idea. This presentation has focused on the language technology aspects, but the work has a simultaneous impact for researchers in primary care, implying research on doctor–patient communication, access to health services by, and improving the quality of access and quality of care to hard-to-reach groups (Lovel et al. 1998), reducing perceived time wasting with perceived difficult patients, developing training agendas for health care professionals, and agendas for community development initiatives (Moran et al. 2000) so that newly arrived communities make better use of the local health services and get a better quality of care not only in the UK but in other countries across Europe, Australasia and North America. It is at the moment a proposal, but we hope in due course to be able to report on its implementation, and on results of trials and evaluations.

Acknowledgement

This work is being undertaken in conjunction with ongoing research at the School of Primary Care in the Faculty of Medicine, University of Manchester under the direction of Dr Hermione Lovel. Principal researchers on this project are Marianne Johnson and Zeinab Mohammed. A version of this paper appears as Somers & Lovel (2003).

References

Alm, N. and J.L. Arnott. 1998. 'Computer-assisted conversation for nonvocal people using prestored texts'. *IEEE Transactions on Systems, Man, and Cybernetics* **28**, 318–328.

Alm, N., J.L. Arnott and A.F. Newell. 1989. 'Discourse analysis and pragmatics in the design of a conversation prosthesis'. *Journal of Medical Engineering and Technology* **13**, 10–12.

Berthelin, J.-B., B. Grau, I. Robba and A. Vilnat. 1999. 'A cross-corpus vision of conversation and dialogue'. *Language Technologies – Multilingual Aspects: Workshop in the framework of the 32nd Annual Meeting of the Societas Linguistica Europaea*, Ljubljana, www.limsi.fr/Individu/jbb/cross-corpus.html.

Blenkhorn, P. 1992. 'A picture communicator for symbol users and/or speech impaired people'. *Journal of Medical Engineering and Technology* **16**, 243–249.

Boitet C. 1993. 'La TAO comme technologie scientifique : le cas de la TA fondée sur le dialogue'. In A. Clas and P. Bouillon (eds) *La Traductique : Études et recherches de traduction par ordinateur*, Montréal: Presses de l'Université de Montréal, pp. 109–148.

Boitet C. 1999. 'Dialogue-based MT and self-explaining documents as an alternative to MAHT and MT of controlled languages'. *Machine Translation Review* **10**, 6–15.

Brent, M.R. 1993. 'From grammar to lexicon: unsupervised learning of lexical syntax'. *Computational Linguistics* **19**, 243–262.

Buchanan, B., J. Moore, D. Forsythe, G. Garenini, G. Banks and S. Ohlsson. 1995. 'An intelligent interactive system for delivering individualized information to patients'. *Artificial Intelligence in Medicine* **7**, 117–154.

Cambridge, J. 1997. *Information Exchange in Bilingual Medical Interviews*. MA dissertation, Department of Linguistics, University of Manchester.

Cawsey, A., K. Binsted and R. Jones. 1995. 'Personalised explanations for patient education'. *Proceedings of the 5th European Workshop on Natural Language Generation*, Leiden, Netherlands, pp. 59–74.

Chalabian, J. and G. Dunnington. 1997. 'Impact of language barrier on quality of patient care, resident stress, and teaching'. *Teaching and Learning in Medicine* **9**, 84–90.

Church, K.W. and E.H. Hovy. 1993. 'Good applications for crummy machine translation'. *Machine Translation* **8**, 239–258.

Copestake, A., S. Langer and S. Palazuelos-Cagigas (eds) 1997. *Natural Language Processing for Communication Aids: Proceedings of a workshop sponsored by the Association for Computational Linguistics*, Madrid.

Deville, G. and E. Herbigniaux. 1995. 'Natural language modelling in a Machine Translation prototype for health-care applications: a sublanguage approach'. *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation, TMI* 95, Leuven, Belgium, pp. 142–157.

Dye, R., N. Alm, J.L. Arnott, G. Harper and A.I. Morrison. 1997. 'A script-based AAC system for transactional interaction'. *Natural Language Engineering* **1**, 1–13.

Frederking, R., S. Nirenburg, D. Farwell, S. Helmreich, E. Hovy, K. Knight, S. Beale, C. Domashnev, D. Attardo, D. Grannes and R. Brown. 1994. 'Integrating translations from multiple sources within the Pangloss Mark III Machine Translation system'. *Technology Partnerships for Crossing the Language Barrier: Proceedings of the First Conference of the Association for Machine Translation in the Americas*, Columbia, Maryland, pp. 73–80.

Frederking, R., A. Rudnicky and C. Hogan. 1997. 'Interactive speech translation in the DIPLOMAT project'. *Spoken Language Translation: Proceedings of a Workshop Sponsored by the Association for Computational Linguistics and by the European Network in Language and Speech (ELSNET)*, Madrid, pp. 61–66.

Grisedale, S., M. Graves and A. Grünsteidl. 1997. 'Designing a graphical user interface for healthcare workers in rural India'. *Human Factors in Computing Systems: CHI 97 Conference Proceedings*, Atlanta, Georgia, pp. 471–478.

Hartley, A. and C. Paris. 1997. 'Multilingual document production from support for translating to support for authoring'. *Machine Translation* **12**, 109–128.

Hays, L. 2002. *Barriers to health care for refugees and asylum seekers.* Special Study Module, Medical Undergraduate Course, Manchester, Supervisor Dr Hermione Lovel.

Higginbotham, J., B.J. Moulton, G.W. Lesher, D.P. Wilkins and J. Cornish. 2000. 'Frametalker: development of a frame-based communication system'. *Proceedings of the CSUN Conference on Technology and Disability*, Los Angeles. Proceedings online http://www.csun.edu/cod/conf2000/proceedings/0156Higginbotham.html.

Iwabuchi, M., N. Alm, P. Andreasen and K. Nakamura. 2000. 'The development of UNICORN – a multilingual communicator for people with cross-language communication difficulties'. *Proceedings of the CSUN Conference on Technology and Disability*, Los Angeles. Proceedings online http://www.csun.edu/cod/conf2000/proceedings/0025Alm.html

Jackson, C. 1998. 'Medical interpretation, an essential service for non-English speaking immigrants'. In S. Loue (ed.) *Handbook of Immigrant Health*, London: Plenum Press, pp. 61–79.

Jary, D. 2001. *What are the barriers to accessing health care experienced by refugees in the United Kingdom?* Special Study Module, Medical Undergraduate Course, Manchester, Supervisor Dr Hermione Lovel.

Johnson, M. 2003. *Do communication impairments offer insights into alternative communication strategies for (non-disabled) patients with limited or no English?* M.Phil thesis, Centre for Computational Linguistics, UMIST, Manchester.

Jones, D. and P. Gill. 1998. 'Breaking down language barriers. *British medical Journal* **316**, 1476–1480.

Kittredge, R. and J. Lehrberger (eds) 1982. *Sublanguage: Studies of Language in Restricted Semantic Domains*, Berlin: Walter de Gruyter.

Kolodner, J. 1993. *Case-Based Reasoning.* San Mateo, CA: Morgan Kaufmann.

Lam, T. and J. Green. 1994. 'Primary health care and the Vietnamese community: a survey in Greenwich. *Health and Social Care* **2**, 293–299.

Langer, S. (ed.) 1998. 'Special issue on Augmentative and Alternative Communication'. *Natural Language Processing* **4**.1.

Langer, S. and M. Hickey. 1999. 'Augmentative and Alternative Communication and Natural Language Processing: current research activities and prospects'. *AAC Augmentative and Alternative Communication* **15**, 260–268.

Langlais, P., G. Foster and G. Lapalme. 2000. 'Unit completion for a computer-aided translation typing system'. *Machine Translation* **15**, 267–294.

Leech, G. 1992. 'Corpora and theories of linguistic competence'. In J. Svartvik (ed.) *Directions in Corpus Linguistics*, Berlin: Mouton de Gruyter, pp. 105–122.

Lilford, R.J., P. Bingham, G.L. Bourne and T. Chard. 1985. 'Computerized histories facilitate patient care in a termination of pregnancy clinic: the use of a small computer to obtain and reproduce patient information'. *British Journal of Obstetrics and Gynaecology* **92**, 333–340.

Loncke, F.T., J. Clibbens, H.H. Arvidson and L.L. Lloyd. 1999. *Augmentative and Alternative Communication: New Directions in Research and Practice.* London: Whurr.

Lovel, H.J., Z. Mohamed and R. Moran. 1998. 'Needs assessment with a hard-to-reach group (the example of Somali refugees in Manchester)'. *Needs and Outcomes Assessment in Primary Health Care*, World Health Organisation, Geneva.

McAvoy, B. and A. Sayeed. 1990. 'Communication'. In B.R. McAvoy and L.J. Donaldson (eds), *Health care for Asians*, Oxford: Oxford University Press, pp. 57–71.

McEnery, T. and A. Wilson. 1996. *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Melamed, I.D. 2000. *Empirical Methods for Exploiting Parallel Texts.* Cambridge, Massachusetts: MIT Press.

Mitchell, E, and F. Sullivan. 2001. 'A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97'. *British Medical Journal* **322**(7281), 279–282.

Moran, R., Z. Mohamed and H.J. Lovel. 2000. 'Taking action on refugee health needs'. *2nd International Conference on Women's Health*, Edinburgh.

Osman, L., M. Abdalla, J. Beattie, S. Ross, I. Russell, J. friend, J. Legge and J. Douglas. 1994. 'Reducing hospital admissions through computer supported education for asthma patients'. *British Medical Journal* **308**, 568–571.

Passonneau, R.J. and D.J. Litman. 1997. 'Discourse segmentation by human and automated means'. *Computational Linguistics* **23**, 103–139.

Peiris, D.R., N. Alm and P. Gregor. 1995. 'Computer interviews: an initial investigation using free text responses'. In M.A.R. Kirby, A.J. Dix and J.E. Finlay (eds) *People and Computers X: Proceedings of HCI '95, Huddersfield, ...*, Cambridge: Cambridge University Press, pp. 281–288.

Ramsay, R. and S.Turner. 1993. 'Refugees' health needs'. *British Journal of General Practice* **43**, 480–481.

Reiter, E. and L. Osman. 1997. 'Tailored patient information: some issues and questions'. *ACL Workshop From Research to Commercial Applications: Making NLP Technology Work in Practice*, Madrid, pp. 29–34.

Sekine, S. 1997. 'A new direction for sublanguage NLP'. In D.B. Jones and H.L. Somers (eds) *New Methods in Language Processing*, London: UCL Press, pp. 165–177.

Smadja, F. 1993. 'Retrieving collocations from text: Xtract'. *Computational Linguistics* **19**, 143–177

Somers, H. 1999. 'Review article: Example-based Machine Translation'. *Machine Translation* **14**, 113–157.9

Somers, H. and B. Collins. 2003. 'EBMT seen as case-based reasoning'. To appear in M. Carl and A. Way (eds) *Recent Advances in Example-Based Machine Translation*, Dordrecht: Kluwer Academic.

Somers, H.L. and D. Jones. 1992. 'Machine Translation seen as interactive multilingual text generation'. *Translating and the Computer 13: The Theory and Practice of Machine Translation – A Marriage of Convenience?*, London, pp. 153–165.

Somers, H. and H. Lovel. 2003. 'Computer-based support for patients with limited English'. *Proceedings of the 7th International EAMT Workshop on MT and other language technology tools Resources and tools for building MT*, Budapest, pp. 41–49.

Somers, H., I. McLean and D. Jones. 1994. 'Experiments in multilingual example-based generation.' *CSNLP 1994: 3rd Conference on the Cognitive Science of Natural Language Processing*, Dublin, Ireland.

Somers, H., J. Tsujii and D. Jones. 1990. 'Machine Translation without a source text'. *COLING-90, Papers Presented to the 13th International Conference on Computational Linguistics*, Helsinki, Vol. 3, pp. 271–276.

Tang, M. and C. Cunninghame. 1994. 'Ways of saying'. *Health Service Journal* **104**, 28–30.

Thomas, J. and A. Wilson. 1996. 'Methodologies for studying a corpus of doctor-patient interaction'. In J. Thomas and M. Short (eds) *Using Corpora for Language Research*, London: Longman, pp. 92–109.

Trujillo, A. 1999. *Translation Engines: Techniques for Machine Translation.* London: Springer.

Uba L. 1992. 'Cultural barriers to healthcare for southeast Asian refugees'. *Public Health Reports* **107**, 545–548.

Van Ess-Dykema, C., C.R. Voss and F. Reeder (eds) *ANLP/NAACL 2000 Workshop: Embedded Machine Translation Systems*, Seattle, Washington.

Véronis, J. (ed.) 2000. *Parallel Text Processing: Alignment and Use of Translation Corpora.* Dordrecht: Kluwer.

Voelker, R. 1995. 'Speaking the languages of medicine and culture'. *Journal of the American Medical Association* **273**, 1639–1641.

Wilks, Y. (ed.) 1999. *Machine Conversations*. Boston: Kluwer.

Wolmuth, P. 1996. 'Removing the barriers'. *Health Visitor* **69**, 93–94.

Woodhead, D. 2000 *The Health and Well-being of Asylum Seekers and Refugees*. London: King's Fund. http://www.kingsfund.org.uk/asar.PDF

Wynn, R. 1999. *Provider–Patient Interaction: A Corpus-Based Study of Doctor–Patient and Student–Patient Interaction.* Kristiansand, Norway: Høyskoleforlaget.