



Translation Project Workshop Report

Pritti Mehta
Genetic Interest Group

A workshop convened by the Genetic Interest Group on 1st December
2003.

ATTENDING

Mushtaq Ahmed – Yorkshire RGC
Lucy Arnold – Contact a Family
Tom Barclay – GIG
Dr. Angela Barnicoat – Great Ormond Street
Bernice Burton – Sickle Cell/Thalassaemia Services, Newham PCT
Tracy Bussoli – Guy's Hospital
Meredith Carter – GIG
Tara Clancy – Manchester RGC
Dr. Katrina Hargreaves – UK Newborn UK screening programme
Professor Anna Kessler – NW Thames RGC
Rehana Khan – Manchester RGC
Dr. Pritti Mehta – GIG
Eva Nazeem – Manchester RGC
Emma Newall – NW Thames RGC
Dr. Ainsley Newson – UCL
Stuart Nicholls – Nowgen
Simon Outram – South Bank University
Shanta Patel – Leicestershire RGC
Kashmir Randhawa – NW Thames RGC
Dr. Anand Saggarr-Malik – SW Thames RGC
Professor Reinier Salverda – UCL
Dr. Alison Shaw – Brunel University
Lucy Ullmann – GIG
Lizzie Winchester – Guy's Hospital
Melissa Winter – GIG

APOLOGIES

Asfa Ahmed – West Midlands RGC
Jo Ayres – Nowgen
Professor Shirley Hodgson – St. Georges
Anna Lane – GIG
Buddug Williams – GIG/Wales Gene Park
Dr. Charlotte Augst – CancerBACUP
Dr. Siddharth Banka – Hull Royal Infirmary
Christine Williams – NHS Sickle Cell/Thalassaemia Screening Programme

THANKS

Many thanks to all that participated in this event, and helped to make it such a useful and successful day. May I also acknowledge Lucy Ullmann who kept note of the meeting and also the contributions of three students. Kavita Sangha and Parul Vansadia of Brunel University, who helped to review the quality and effectiveness of current translations. Our third student was Meredith Carter of George Mason University USA, who just completed a 3-month internship with us (Sept – Dec 2003). She helped to develop the lexicon, gathered data on language needs, and also was involved in organising the workshop.

Disclaimer

This document aims to report on the discussions that took place during the Translation Project Workshop, and highlights future outcomes. Whilst every effort has been made to represent these discussions accurately, the report may inevitably represent the author's interpretations, and therefore responsibility for content lies solely with the author.

Pritti Mehta
GIG

INTRODUCTION

The London *IDEAS* Translation Project aims to develop and deliver meaningful and accessible patient information to London's linguistically diverse communities, and ultimately to improve patient access to genetic services. Resources are available to translate up to 27 patient information leaflets into 12 languages, and there is limited provision for providing this information in audio formats.

Although this is a local initiative, material generated will have national application and be made freely available to all Regional Genetics Centres (RGC-s). The city is the most linguistically diverse capital of the world, inhabited by over 45% of Britain's minority ethnic populations. London's RGC-s also serve over 14 million people – almost ¼ of the UK population. Therefore, whilst this project will not fulfil the needs of even the local population, London is extremely well placed to pilot this work.

The North-West Thames Regional Genetic Service in collaboration with the Genetic Interest Group carries out this project. The workshop organised by Dr Pritti Mehta of the Genetic Interest Group, aimed to seek professional advice and suggestions over some of the key issues arising from the project. These were:

- 1) To help identify which languages to include in this project;**
- 2) To discuss the North-West Thames Regional Genetics Service's proposed topics for translation;**
- 3) To introduce and discuss the proposed model for translation of genetic information.**

The discussion is divided into 3 sections, each focusing on one of the key aims listed above.

Tom Barclay of the Genetic Interest Group chaired the workshop. The programme began with presentations from Prof. Anna Kessling and Emma Newall of North – West Thames RGC, Pritti Mehta (GIG) and finally, Reinier Salverda, Professor of Dutch Literature (UCL), who has a specific research interest in London's Multilingualism. Copies of AK's, EN's and PM's presentation slides and RS's papers can be made available upon request. The following highlights the key point of each presentation.

Prof. Anna Kessling; Co-Director Knowledge Dissemination Unit (KDU) London IDEAS GKP

- Presented background information about the London IDEAS GKP;
- Provided information about Knowledge Dissemination Unit (KDU) of the Knowledge Park, which specialises in training, information & translating; and
- Explained how the translation project was conceived and falls into the remit of London IDEAS GKP.

Emma Newall; Genetic Associate North-West Thames Regional Genetics Centre

- Presented the proposed topics for translation;
- Explained how she was developing standardised English leaflet versions, paying particular attention to style and content and reducing ambiguity; and
- Presented results of back-translation, and explained how this helped to identify potential problems associated with translation of complex genetic/health information.

Dr Pritti Mehta – Genetic Interest Group

- Presented the results of the review of current patient information translations, which involved back-translation and focus groups;
- Introduced proposed paradigm for translation, based on community consultation and appropriate validation and evaluation methods;
- Proposed lexicon of genetic information; and,
- Presented current audits used for assessment of language needs.

Prof. Reinier Salverda - UCL

- Presented his current papers examining Multi-lingualism in London;
- Discussed the importance of getting the information right;
- Highlighted the need for visual information and provided key examples; and,
- Explained the importance of correct delivery and dissemination of the translated product.

DISCUSSION

The Key Languages for Translation.

The following, recent assessments of language needs, were highlighted in order to focus on, and possibly identify particular languages for inclusion in this project.

- Languages of London project - a study looking at languages spoken by London's school children, listing over 300 languages;
- RGC-s perceptions of language needs and the actual data obtained during a one month multi-centre pilot study conducted by GIG;
- Language line data on requests for interpreters from health sector in London and UK; and,
- Translation agencies data on top 20 Language requests from public and health sectors

These audits clearly demonstrated the diversity of language needs and hence the challenge that lay ahead in prioritising particular languages for the project. Variation between the different data sets was also noted. This was in part attributable to regional and local bias, but was also due to more subtle factors that would need to be teased out. To give example the group looked at the way Urdu was ranked amongst the different lists. Whilst appearing top on the RGC-s list of perceived language needs, it, in contrast, ranked number 15 in the Language Line (UK) data on requests for interpreters. This difference, however, was likely to be because Urdu speakers prefer to use family members to act as interpreters.

It was noted that the Languages of London project had listed Urdu and Hindi together as one category. Whilst these languages are very similar when spoken their written forms are considerably different. Urdu is the national language of Pakistan, is written in Arabic script and has borrowed many words from Persian and Arabic. Hindi on the other hand, has borrowed both the script and many words from Sanskrit. Urdu is written from right to left, unlike Hindi, which is written from left to right, but under the line rather than on it.

Whilst such audits may be very useful in helping to determine language choices, the following issues should also be considered:

- Genetic need; this may differ greatly between groups and therefore must be reflected in the choice of languages. Is there any data that may help to inform this? What types of diseases are most common? Need to look at disease frequency investigations.
- Dynamics of languages; Language needs will change from year to year. Important to predict changes in language needs in the future especially with countries joining the EU. E.g. Polish is likely to become important next year with Poland joining the EU. Need for Mandarin (or Cantonese) might also become more widespread, as the Chinese population is forecasted to grow over the next few years.
- London has high turnover of immigrants/asylum seekers probably not penetrating the system; they are likely to be under-represented, not likely to speak English fluently; are highly mobile; may live in temporary or shared accommodation, and maybe working long hours. These are all factors that may reduce uptake.
- Literacy levels of different populations and the efficacy of written versus audio material for certain groups

The following points about specific languages were highlighted:

- Sylheti: the language spoken by 95% of Bangladeshis living in the U.K. This language is distinct from Bengali. The writing system for Sylheti is called Sylheti Nagri, last taught in schools in Sylhet more than 50 years ago. Sylheti is therefore now predominantly a spoken language and therefore audio translations in Sylheti will be more useful for the British Bangladeshi population.
- Although Punjabi has a rich literary tradition it is more of a spoken language. There are also several dozen dialects. Amongst the main ones are Western/ Lahnda Punjabi (dominant in Pakistan) and Eastern/ Gurumukhi Punjabi (dominant in India). In deciding which Punjabi dialect to choose for translation, it is important to consider that the Pakistani populations will also make use of the western Punjabi audio translations and that their genetic needs are considered to be greater than that of the Sikh Punjabi population. The Pakistani population is also thought to be larger than the Sikh Punjabi population.
- Hindi is the national language of India and the Lingua - Franca of Indian cinema. It might therefore be useful to provide audio translations in this language.

The group noted that African languages are often under-represented and briefly discussed which languages might be needed by the Black African populations living in Britain. The following points were raised.

- Largest black African populations groups living in Britain are Nigerian, Ghanaian, Ugandan, Somali, and Sierra Leonian. Whilst there are many community languages and dialects spoken by Ghanaians, Ugandans and Nigerians, these populations also read and speak English. There are some exceptions amongst older generations, e.g. older Nigerians might prefer Yoruba or Ibo, and older Ugandans might prefer Bagandan.

- Among the Somali Community there are fewer languages. Most share English and Somali, and Arabic is still the liturgical language among Somali speakers. However large proportion for whom Somali is the only language and there are low literacy levels following conflict (female literacy rate about 24%). There is therefore a very strong case for providing both written and audio forms of Somali.
- English is also the official language of other African nations, such as Liberia, Nigeria, and Cameroon. People from former Zaire and other African populations such as Guinea, Gabon, Congo, Chad, Niger, Morocco know French. Similarly, populations from Mozambique, Guinea Bissau, Cape Verde, Angola know Portuguese.

The project should consider generally targeting some of the most widely spoken languages of the world, as these would in turn have wider relevance and future value.

For languages such as French, Spanish, and Portuguese it might be feasible to obtain genetic literature direct from the country of origin. In the case of languages such as Somali and Turkish, relevant patient information leaflets might be developed by countries such as Holland, where there is already a substantial presence of these minority ethnic groups. It was however suggested that such information might need to be translated back into English, in order to validate it and obtain English or dual language versions, and therefore would incur the same translation costs.

Other points to note;

- Languages such as Albanian (scores high in Brent), Romanian, Kurdish, Farsi are also potentially important.
- Current GCSE language choices might give indication of what is needed.

To conclude the group identified the following languages for inclusion in the project. These are listed below in alphabetical order. Consensus was, that the final decision be based on evaluation of the following factors: number of speakers; and language, genetic and social needs.

Albanian,
 Arabic,
 Farsi,
 French,
 Hindi,
 Gujarati
 Kurdish,
 Chinese (Mandarin or Cantonese?)
 Portuguese,
 *Punjabi,
 Somali,
 Spanish,
 *Syhleti
 Turkish,
 *Urdu,

* Audio translations of these languages would have greater efficacy than written translations.

The Key Topics for Translation

The following points/suggestions were made regarding the proposed list of topics for translation.

- Whether Cystic Fibrosis (CF) should be included as it predominantly affects people of Northern European origin. It was noted however, that a number of cases have emerged within consanguineous Pakistani families. It was also noted that all newborn babies would be screened for CF, thus highlighting a potential need for information about CF screening in languages other than English.
- Whether to combine certain leaflets such as 3 and 7 (bowel and colon cancer).
- Members of the group commented on the absence of sickle cell and thalassaemia from the list, however, it was noted that discussions had taken place with the haemoglobinopathy screening committee about collaborating over the production of such leaflets.
- The group suggested that it might also be useful to develop leaflets on the following topics;
 - Losing a baby including post mortem examinations and counselling
 - Undiagnosed malformation syndromes
 - Deafness
 - Insurance
- Certain topics are often given higher priority than others and therefore might already have been translated into some of the key languages. It might be more meaningful to translate topics that have not been previously translated.

The Proposed Protocol for Translation

The groups' comments regarding the proposed protocol were as follows;

- Whilst community consultation and piloting with patients is a very important part of the protocol, recruiting and working with lay people might prove to be somewhat problematic.
- Clients should work very closely with translators and guidelines should avoid unnecessary restrictions and allow for words that are not in dictionary to be translated. In this regard, the group agreed that it would be very important to develop a lexicon of key genetic terms - as proposed.
- It might be helpful to write things in the new language from scratch, rather than trying to translate from English.
- In order to generate wider interest the project should be discussed at an international level such as at the Genetics and Population Health Conference to be held in Perth, Australia in August 2004. The American Society of Human Genetics should also be contacted and at least be made aware of the project's aims
- Interpreters need to be provided with information about genetics, therefore might also benefit from receiving genetic information leaflet translations.

- Legal and other specialist translators exist, but there is a need for medical translators. Certain Universities (Bradford and DeMontfort) offer medical translation courses. It might also be useful to develop genetic translation courses.

Future Outcomes

Having considered a list of 20 languages, including those proposed by the workshop participants, the project team has decided upon the first 11 that will be prioritised in this project. In the main, choices have been based on number of speakers living in London (extrapolated from languages of London project), genetic service's perceived level of need, and information known about genetic need. Data on the linguistic needs of minority ethnic groups living in Britain is scarce. As a proxy, female literacy rate data for the country, or state of origin (based on UNESCO estimates for year 2000, and National literacy mission for year 1991) has been considered. Current UK asylum statistics for 3rd quarter 2003 (Home office) have also informed this decision. The information is summarised in the table overleaf, which also includes data obtained from translation agencies and Language Line. Less weight is however, given to the Language Line data, as it can be an inadequate indicator for communities where family members are preferred as interpreters.

The first six languages that will be prioritised are Arabic, Gujarati, Punjabi, Sylheti, Turkish, and Urdu. These are the most widely spoken minority languages in London, and are amongst the top choices for both London and UK-based RGC-s. Female literacy levels for country/state of origin ranges from 28 – 76% across these different population groups. Pakistani, Bangladeshi, Arabic and some Turkish populations are also likely to have an increased genetic need due to higher incidence of consanguinity.

The following 5 language choices are likely to be Chinese (Mandarin), Portuguese, Farsi, French, and Somali. This decision is again based upon number of speakers in London, and RGC perceived level of need. Somali, Chinese (Mandarin), Farsi, and French are also amongst the respective languages of some of the largest population groups (Somalis, Chinese, Iranian, Iraqi, Dem. Rep. Of Congo), currently seeking asylum in Britain. Such populations are likely to have increased social and linguistic needs. Somali female literacy rate is 24%.

The 12th language will be selected from the remaining 9 options; once more data on the social, linguistic and genetic needs has been gathered.

The project team has decided that whilst the chosen topics for translation will remain the same, attempts will be made to combine certain leaflet topics where possible. New topics suggested during the workshop, will be given priority after completion of the original set, if time and funding allows.

A proposal to orally present this work at Genetics and Population Health Conference in Perth, Australia, (August 2004), has been accepted.

Table 1: Summary of language data

Language	Estimated number of speakers in London	UK-RGC perceived need	London-RGC perceived need	Applications for Asylum; 3 rd Quarter 2003, indicating country of origin	Female Literacy Rate	UK-based Language Translation line data	agencies
Bengali+							
Sylheti	119,900	15	4		30	5937	10
Urdu	114455	21	4	(Pakistan) 495	28	5689	10
Punjabi	143000	12	2		58	5934	10
Gujarati	138000	13	2		61	0	9
Turkish	67000	10	4		76	16845	7
Arabic	49500	10	4	(Iraq) 690	48	8292	8
Cantonese +						3990/	
Mandarin	46000	5	1	(China) 965	82	2074	10
Portuguese	26900	5	3		90	12088	3
French	25300	4	2	(Dem.Rep.Congo)380	50	12954	4
Somali	19037	3	2	(Somalia) 1440	24	9957	5
Farsi	14900	4	3	(Iran) 860	69	10440	4
Spanish	24500	2	2		87	6148	3
Tamil	17700	3	1		62.66	2635	3
Vietnamese	15800	1	1		91	2034	3
Hindi	11445	9	2		45	0	8
Italian	11300	2	1		98	0	0
Kurdish	6200	1	1	(Turkey) 530	45	36046	4
Polish	6600	1	1		70	5156	0
Albanian	3900	3	2		97	7268	3
Dari/Pushto	1600	0	0	(Afghanistan) 470	69	2005/ 2157	1

NB.

Asylum seeker languages have been determined by considering the official language of the country of origin however there are some caveats.

- Arabic is the official language of Iraq. Arab refugees from southern Iraq speak Arabic, However there are also Iraqi Kurds and Assyrians who speak Assyrian and a Kurdish dialect Sorani respectively.
- Farsi is the official language of Iran. Iranian refugees include Armenian, Assyrian, Kurdish minorities who may speak their respective languages at home, as well as Farsi and English.

Majority of refugees from Turkey are Kurds, whilst they speak Kurdish, they write Turkish, however they have low literacy rates.

Whilst Languages of London project list Urdu and Hindi speakers together they report that there were approx 10x more Urdu speakers than Hindi speakers. The individual proportion of Urdu and Hindi speakers has thus been calculated from the total number of Urdu/Hindi speakers.

The French literacy rate is based on the figures derived for the Dem. Rep. Of Congo. Portuguese and Spanish literacy rates are based on figures for Portugal and Spain Female literacy rates for the following Portuguese speaking populations are given; Mozambique (29%), Guinea Bissau (28%), Cape Verde (66%).

