Lynda Proffitt

University of Westminster Conference on the Future of Interpreting

Working Session ll Research and the Profession

Saturday July 1st 2006

When I was first approached to give a presentation at this conference, I was asked if I would speak on "English as an International Language". As an interpreter, constantly exposed to my colleagues' complaints about speakers, my natural reaction to the expression "English as an International Language" or "Global English" or "Globish" is to say "dream on"! As the oft-quoted David Crystal has said: "everything depends on just how great a command of English is considered acceptable to count as a 'speaker of English'. He also refers to 'new Englishes' the new dialects of English which have emerged in, say, India or West Africa and the likelihood that they will become increasingly divergent. Although he optimistically holds out hope for what he calls a ' World Standard Spoken English', with the speaker deliberately avoiding non-standard vocabulary and possibly also non-standard pronunciation and grammar for the purposes of international communication, he nevertheless admits that 'it has hardly yet been born'. Where does that leave the poor interpreter?

A survey of AIIC interpreters in 2004 contained a section on stress factors in the profession. The report of the survey stated : "… dissatisfaction is primarily related to pervasive changes in the working environment restricting the interpreters' ability to do their job properly. The top dissatisfaction factors are the rise of Global English that leaves many interpreters frustrated – and/or bored, if they happen to work in the English booth and the lack or tardiness of materials to prepare properly".

As an indication of the way things are going, 65% of all the speakers taking the floor over 2 days of the international meeting covered by my research were NNSs (Non-Native Speakers) of English.

What adds to the interpreter's stress of course is that many participants at conferences, whatever their mother tongue, happily follow the proceedings in English until a difficult-to-comprehend NNS of English takes the floor. That is the moment when they reach for their headphones, tune into their preferred language channel and expect the interpreter to compensate for any difficulties they are having in understanding the speaker.

The research I carried out 9 years ago as part of a Masters Degree therefore had its origins, not in English *as* an international language but in English *used* as an international language and it is that research which I am going to describe here today. It was entitled "Simultaneous Interpretation of the Non-Native Speaker of English: Perceptions and Performance"

When I began looking at the research literature in the field, I realised that, although there are plenty of general theories of interpreting out there, virtually no empirical research had focused on the particular aspect of interpreting which interested me, the difficult NNS of English, and I decided to try to design a small-scale exploratory experiment to see how a sample of interpreters actually performed when interpreting a difficult NNS, how some of the theories might fit the scenario of interpreting a NNS, and whether there were any pointers for further research. I shall briefly describe the technicalities of the experiment itself and then go on to give you an overview of the results and possible explanations for them.

Once I had obtained agreement for the whole project from a conference organiser, the first problem was to select the input material for the experiment. I wanted authentic spontaneous speaker input, (since read statements are a separate field of research in themselves), but I was at the same time attempting to control for as many between-subject (interpreter) and within-subject variables as possible. One thing that has always struck me is that the speakers complained about by interpreters are not necessarily the speakers that I, sitting bored in the English booth, would have identified as difficult through listening alone. For the purposes of research, I could not rely on my own judgement to select the input material.

So I began by distributing a questionnaire to interpreters at a 6 language meeting, asking them to fill in a rating sheet at the end of each half-hour shift over a period of 2 days for any speakers, whether NS (Native Speakers) or NNS, who had been particularly easy or particularly difficult to interpret. I asked the interpreters to assess those speakers for 7 different features of the statements they had interpreted: Sentence structure, Accent, Intonation, Choice of Expression, Complexity of ideas, Delivery rate, Length of pauses. At the same time I kept a record of precisely who had spoken in the meeting over those 2 days and for what length of time.

There were some interesting results from my statistical analysis of the 73 rating sheets produced, which I do not have time to go into here (Tables 2&3). I eliminated responses relating to speakers speaking a language other than English and those relating to read statements. This gave me 50 responses. What is of particular relevance for the outcome of the experiment was the analysis of questionnaire results for delivery rate. To my surprise, unlike sentence structure, which proved to be the most significant of all the distinguishing features between speakers identified as difficult or easy, the rate of delivery was not a significant distinguishing feature.

For my input material, I compiled a 15 minute tape of brief, spontaneous statements in English, two by a NS described as particularly easy to interpret by 3 different booths on basis of questionnaire, (under 'further comments' they had written "what all speakers should be like", "one of the best speakers to interpret") and two by a NNS categorised by 3 different booths as particularly difficult to interpret ("impossible to understand", "not intelligible"). The interventions contained roughly the same number of words for each speaker, but the NS was faster (175 words per minute at his fastest, as against a fastest speed of 115 words per minute for the NNS). Of course with spontaneous speech, speed was not a variable I could easily control, particularly given the way I had chosen the input material. I gave the interpreters the acronyms and technical terminology mentioned in the statements in advance.

The idea was to establish a control baseline for my subjects when interpreting the easy NS, against which I could compare their performance when they interpreted the difficult NNS.

Next find your victims. I managed to convince 9 experienced interpreters to act as subjects. 3 each from English, French and Russian booths (i.e. the languages I can transcribe). The F and R booths all had English B's or had lived in the UK for more than 10 years.

Why did I need an English booth ? I asked the 3 English (monolingual) booth interpreters to work from English into English but not to shadow the speakers word for word. They were requested to process what they were hearing as if they were producing a "relay" interpretation for other booths. This meant, in theory, that the processing load was the same for the English booth as for the 2 other booths – all that was missing in their case was the language switch itself. Comparison of the monolingual and bilingual outputs for the NNS might, I thought, provide an insight into the demands of the difficult NNS on processing capacity with and without the language switch.

Immediately after interpreting the tape, the interpreters filled in a questionnaire with a section for comments about the speakers and asking if they had been aware of using any particular strategies.

The transcripts of the interpreters' output were rated by outside bilingual interpreters using the Carroll scale, originally devised for translations and first used for interpreters by Gerver, and subsequently Andersen), with 2 raters for each transcript. The material was transcribed to reflect verbal content only, not paralinguistic features. The Carroll scale rates for Intelligibility and Informativeness and, whilst as imperfect as any other rating system, it does have the advantage of providing fairly explicit guidelines for inexperienced raters.

When it came to analysing the results, all my prior expectations were confounded.

In the post experiment questionnaires the interpreters' comments were in line with those made in the questionnaires used to select the input: the NNS was "very unclear" "English of speaker was very bad" etc . They were aware of using waiting strategies for the NNS. There was no mention of speed as a problem with the NS.

However, based on total scores (i.e. taking into account both Intelligibility and Informativeness) and regardless of target language, in the 'bilingual' mode the 6 interpreters of the French and Russian booths obtained better overall scores for the difficult NNS than for the easy NS, whereas in the 'monolingual' English to English mode, scores were relatively constant for both the NNS and the NS. The monolinguals therefore outperformed the bilinguals with NS input but with NNS input their results were of the same order of magnitude (Figs. 4&5).

To briefly address some of the questions raised by my findings:

Disregarding all the possible artefacts of the experiment itself and bearing in mind that any results on such a small sample of interpreters are necessarily highly tentative, what could account for these performance outcomes?

Could it all be down to delivery rate differences between the NS and NNS after all? Did the slower speed of the NNS and/ or longer pauses give the interpreters more time for disambiguation, more time to mentally re-rehearse the phonological input and use more top-down processing involving other linguistic or extralinguistic clues?

Perhaps, as Strolz postulated in 1992, a higher level of alertness may be triggered by difficulty of understanding – so the interpreters may have simply gone into a higher gear with the NNS and drawn on additional reserves of capacity.

What then can account for the discrepancy between the interpreters' perception of the speaker as a difficult subject for interpretation and their relatively satisfactory performance when interpreting that speaker ? It seemed to me that the extra processing capacity devoted to comprehension of the NNS and the need to go that extra mile in extracting meaning had skewed the interpreters' perception of the interpretability of the speaker. That explanation would tally quite neatly with some research into the comprehension of NNS carried out outside the field of interpreting. Munro & Derwing (1995) showed that awareness of the extra processing time required causes accented input to be evaluated by listeners as difficult to understand, despite the fact that the input can be shown to be well understood by the listener. With the NNS there may be none of the automacity in the processing of the input which equates with saved capacity and there may be a greater requirement for top-down processing. drawing on extralinguistic and contextual knowledge, as would be predicted by Chernov's Probability Prognosis Model.

Why the better performance of the monolingual English booth for the NS, but a performance comparable with the bilingual booths for NNS input? It may be explained by a degree of shadowing, i.e. for the English booth there was less of a processing demand for lexical and syntactic restructuring when interpreting the NS, whereas producing a 'relayable' version of the NNS demanded capacity equivalent to that of interpreting a foreign source language, thus explaining the convergence of scores for all booths when interpreting the NNS.

Does this kind of research have any relevance for the profession or should the difficult NNS simply be considered as one of the extremes of the speaker continuum? Well yes I think it does have relevance.

For instance, although not a directly stated aim of my research, one of the ideas I have long entertained, speaker sensitivities permitting, is that greater use could be made of an otherwise idle English booth as an English to English relay booth for other booths for difficult NNSs (and indeed NSs of English with a strong dialect).

The results of my research pretty well knocked that idea on the head, at least in the case of the NNS selected for input. With the NNS input, the language switch seemed to have no discernable bearing on the performance outcome. What would be interesting, however, is to test the English to English relay idea over a more varied range of NNSs, alongside a control group of interpreters working directly from the speaker. Would the normal message loss of the relay system negate potential benefits, if any, for the other booths?

Another example of an area where I believe research of this kind applies is in the old B to A versus A to B argument. There are those who maintain that difficult NNS input is best

served by the A into B language combination – they would be proved wrong on the basis of my findings. When interpreting the NNS even English As to As had no edge over Bs to As.

I leave it to the trainers in the room to decide whether NNS oriented research has any relevance for their work.

Thank you for listening.

Table 2: Means, Standard Deviations and t-test results for all spontaneous speakers

Feature	Difficult speakers (n = 23)		Easy speakers (n = 27)		t-test
	Mean	Std Dev.	Mean	Std. Dev.	t
Sentence Structure	8.52	1.56	5.30	2.26	5.84**
Accent	7.74	2.26	5.22	2.21	3.97**
Intonation	7.48	1.93	5.36	2.04	3.69**
Choice of Expression	7.43	2.35	4.81	1.94	4.20**
Complexity of Ideas	7.22	2.45	4.96	1.93	3.64**
Delivery Rate	6.61	2.08	5.93	1.75	1.26 n.s.
Length of Pauses	5.29	2.55	6.12	1.75	1.27 n.s.

** p = <0.01

Table 3: Means, Standard Deviations and t-test results for NNSs only

Feature	Difficult speakers (n = 22)		Easy speakers (n = 19)		t-test
	Mean	Std Dev.	Mean	Std. Dev.	t
Sentence Structure	8.77	1.02	5.89	2.06	5.43**
Accent	8.04	1.76	5.84	1.89	3.86**
Intonation	7.68	1.93	6.06	1.52	3.09**
Choice of Expression	7.32	2.34	5.00	1.94	3.43**
Complexity of Ideas	7.32	2.46	4.84	2.09	3.45**
Delivery Rate	6.45	1.99	5.26	1.41	2.18*
Length of Pauses	5.50	2.42	5.99	1.78	0.56 n.s.

** p = <0.01
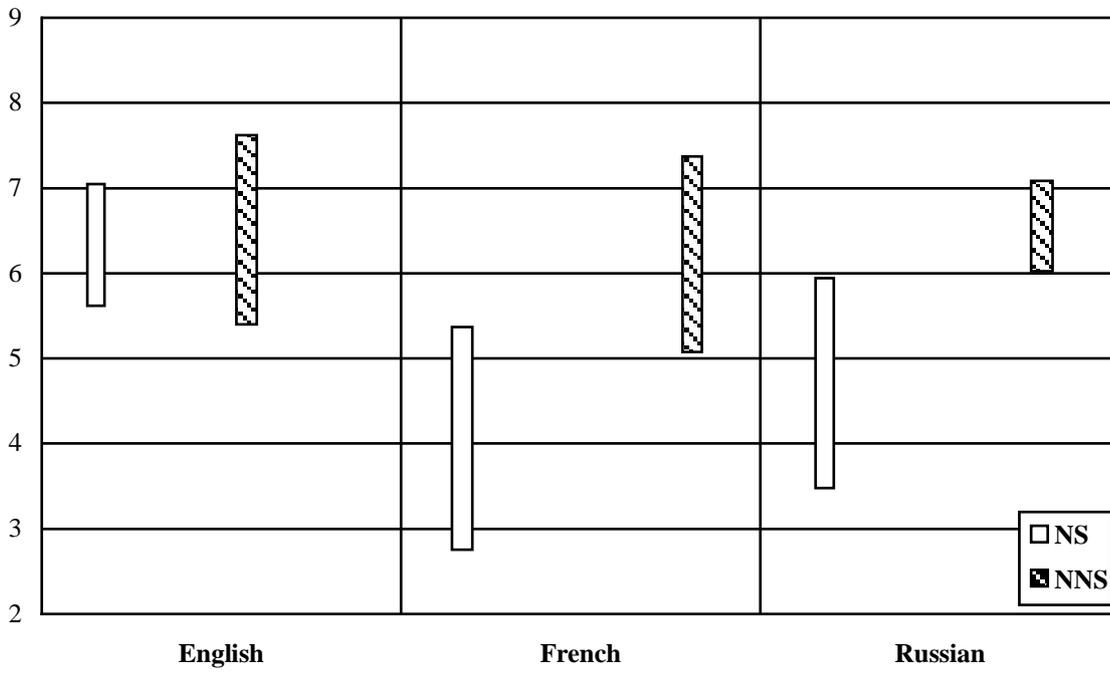*   p = <0.05

Figure 4: Ranges of Total scores by language and type of input

Figure 5: Means of Total scores: