

PRESENTATION BY PIA NOTHLEV

Respeak seminar in Barcelona 19th June 2009

Thank you, for inviting me to this seminar. There are so few possibilities of getting knowledge in this area, so I am very grateful to have connected with you and through you this conference.

It was not a hard job to convince me to go to Barcelona, although it was only two weeks ago.

I have very few people to exchange views and experience with in this job, and apparently my superior officer understands this well, which I am very grateful for.

TV 2 introduction

TV 2 | Danmark is the second national broadcaster in Denmark. We started in 1988 as an alternative to the monopoly, Danmarks Radio.

Although we are a commercial channel, we were also born with public service obligations – and since 2004 we get no license from the government. By now you have probably already guessed that it is not an easy game to get this little niche going. It is argued that live subtitling does not “earn money”, which is probably true. But in my mind, maybe more hard-of-hearing people will watch TV 2 if we make good live subtitling.

There is about half a million hearing impaired people in Denmark.

Our challenge at the moment is to meet the public service obligations that are facing us in 2012 – they require that we live subtitle ALL live programs from news over sports to live entertainment. The present public service obligations will be re-negotiated next year, so with the financial challenges TV 2 has these days we cannot foresee what will happen.

Introduction

I have been involved in the live subtitling area for about 5 years now. In 2004 we started live subtitling using dual keyboards and I was hired as a live subtitler.

In 2005 I became involved in the development and introduction of a speech recognition system funded partly by the government, which – according to our public service regulations – had to be used instead of dual keyboard.

For two years now we have used Philips SpeechMagic speech recognition system (integrated with Wincaps) for live subtitling of one daily newscast from Monday to Friday.

We have also live subtitled various election programs, election nights and royal events with respeak.

On one occasion we have done interlingual respeak, namely the inauguration speech made by president Obama and. When

Obama spoke in Prague the same respeaker (who is a trained interpreter) did relay interpretation of the interpreter via respeak. This, I understand, is not done very often.

It is a certainly a tour de force and you need a lot of preparation (we did not have the speech beforehand). The problem with interpreting and respeaking is that a lot of hearing people will also want to see it, and they are not used to the way it is running – scroll mode and with the inevitable errors, so we had some complaints.

System challenges

The system was built on a medical context used for Danish hospitals and private doctors. After that it was tuned with lots and lots of text from the Danish parliament, Danish papers and archive material from Danmarks Radio and TV 2.

The Danish language area being as small as it is (only 5,5 million people speaking it) it was hard to find someone who was willing to make a Danish speech recognition system, but Philips (now Nuance) in Vienna made the attempt, which we are now using.

TV 2 and DR, who worked together on the project together with the Ministry of Science and the Ministry of Culture, had a large number of requirements for the system, one of which was “easily and quickly to load new words into the context”. This is still one of the issues that we struggle to have a way of solving, as the present one is not good enough. In their latest email to me the producers stated that they don’t know how to handle our use-case.

We have no perfect tool to prepare a newscast in a short time. There is a "train word" tool, but unfortunately it prevents the system from improving acoustically, which of course is essential.

Another thing that may be special to us is that our respeakers only work 2 hours per day and have only 1 hour for preparation.

We have a fair recognition rate of about 97%, but the tool we have to measure this is not based on the number of right or wrong WORDS but on the number of right or wrong characters, which makes it somewhat obscure as the same characters can be in a word and if moved around can mean something completely different.

There is also a challenge in keeping proper nouns in the context down to a number that the system can handle. We often see erroneous recognition where proper nouns are recognized instead of any other word as proper nouns take precedence to other words.

The version we use is a 5.1, but already in 2006 Philips released a version 6.1 with new features. However, this version is without the feature that we use for unknown words: Train word. So there is no easy way to load new words and proper nouns before a news cast.

Language challenges

The Danish language contains about 20 vowel sounds (in Spanish there is only 5!). This makes speech recognition quite challenging – sometimes there is just one consonant in a word.

Words are often spoken the same way but the meaning but the spelling is different. The meaning of words is produced by tiny little nuances of the vowels. This makes it hard for a recognizing engine to distinguish the right words.

Respeakers

At the moment there are 8 respeakers who are all freelance, including me. However I am working on a daily basis and also run the system, hire new respeakers, train them and make reviews and have contact with our suppliers. I also do a lot of administration and maintenance of the system.

When we started respeaking, the same people who had up until then done the typing were asked if they wanted to continue, so none of us were chosen for this special job. It turned out, however, that everyone had respeaker qualities we could build on.

In the meantime 3 have left and I hired 3 new respeakers over the past two years. It takes a long time – and is therefore quite expensive - to train new respeakers. We tend to let them go on air before they are really ready, but that is good training too.

Training of respeakers

The training consists of

- 1) Initial training for one hour
- 2) reading a lot of text material to the profile, at the same time practicing punctuation.

- 3) How to go about in the technical system
- 4) Verbatim respeak of speeches – the (former) prime minister's new year's speech has been a very good practice, as he is very clear and not too fast speaking. This also gives good practice in split attention.
- 5) Short news casts – using 2 colours
- 6) Longer news casts – using four colours after a while and finally punctuation (usually this skill tends to be overlooked a bit but I advise to make short sentences with many full stops.
- 7) Editing skills – usually a course of 2-4 hours and some exercises.
- 8) How to prepare a news cast
- 9) Watching a lot of television with subtitles is also recommendable.

Editing

We believe in editing! We do it

- partly because we believe that no one is able to read verbatim, and certainly not the hard-of-hearing as they seem to have another ability to read than hearing people.
 - and partly because there is a lot of extra information on news casts that we don't need to repeat.
- a) we try to avoid saying the same thing twice,
 - b) we avoid saying things that can be seen on screen (in graphics fx) and
 - c) we try to avoid things that are implicit.

Another thing that I find necessary is to change the spoken words into written language, because spoken language is harder to read and also harder to recognize due to the fact that the system was fed with millions of written texts in order to tune the context.

I am not sure as to what extend we edit percentagewise. But editing is – as mentioned earlier – mainly done because it is impossible to hear and respeak each and every sentence because the ST rate is so high on news bulletins, so more often than not sentences are skipped because it is not possible – at least with our training – to do it all.

Viewer reception/perception

We do not get very many reactions from viewers, so we assume that they accept what they see. But we strive to do it better all the time and after having read Pablo's papers on scrolling versus block mode I will certainly try a few things out. In my point of view people would get used to reading scroll mode after a while.

I find scrolling better for faster spoken programs because block mode has such a long delay, whereas it may be ok to use block mode in fx royal transmissions, etc.

I understand from Pablo's research that hard-of-hearing or deaf people think scrolling is not edited as opposed to block mode.

I wonder whether this is also the case in Denmark. We are probably more used to subtitles than many other European

countries as ALL our foreign programs are subtitled (as opposed to dubbed).

Therefore in Denmark they probably know that all subtitles are edited.

Addition

We are not able to correct on the fly – we can correct a wrongly spoken word respeaking right away “erase word” and then say the right word, but there is no way of correcting wrong recognitions.

Every respeaker corrects his/her respeaks after transmissions and I adapt the respeaks afterwards.

We use different colors for different speakers.