



R&D White Paper

WHP 070

September 2003

A distributed live subtitling system

M. Marks

Research & Development
BRITISH BROADCASTING CORPORATION

A distributed live subtitling system

M. Marks

Abstract

There is an increasing requirement for broadcasters to provide subtitles for their television programmes. New technology allows this challenge to be met through alternative means of subtitle generation and more flexible operating practices.

The BBC has several national television services and many regional variations, which result in considerable difficulties for efficient live subtitling. A solution was required which has the flexibility to allow any subtitler to select easily and subtitle any service from any location, using speech recognition as well as the traditional method of stenography. The challenge of providing these facilities, with the resilience of a broadcast-critical service, has been met by a networked system of subtitling workstations and subtitle inserter gateways. This paper describes the development of this system and its roll-out in a complex broadcasting environment. It is helping the BBC to achieve the goal of subtitling all programmes by 2008 whilst minimising the additional costs involved.

This document was originally published in the Conference Publication of the International Broadcasting Convention (IBC 2003) Amsterdam 11th-15th September 2003

Key words: television, subtitling, speech recognition.

White Papers are distributed freely on request.
Authorisation of the Chief Scientist is required for
publication.

© BBC 2003. All rights reserved. Except as provided below, no part of this document may be reproduced in any material form (including photocopying or storing it in any medium by electronic means) without the prior written permission of BBC Research & Development except in accordance with the provisions of the (UK) Copyright, Designs and Patents Act 1988.

The BBC grants permission to individuals and organisations to make copies of the entire document (including this copyright notice) for their own internal use. No copies of this document may be published, distributed or made available to third parties whether by paper, electronic or other means without the BBC's prior written permission. Where necessary, third parties should be directed to the relevant page on BBC's website at <http://www.bbc.co.uk/rd/pubs/whp> for a copy of this document.

A DISTRIBUTED LIVE SUBTITLING SYSTEM

Matthew Marks

BBC Research and Development Department, UK

ABSTRACT

There is an increasing requirement for broadcasters to provide subtitles for their television programmes. New technology allows this challenge to be met through alternative means of subtitle generation and more flexible operating practices.

The BBC has several national television services and many regional variations, which result in considerable difficulties for efficient live subtitling. A solution was required which has the flexibility to allow any subtitler to select easily and subtitle any service from any location, using speech recognition as well as the traditional method of stenography.

The challenge of providing these facilities, with the resilience of a broadcast-critical service, has been met by a networked system of subtitling workstations and subtitle inserter gateways. This paper describes the development of this system and its roll-out in a complex broadcasting environment. It is helping the BBC to achieve the goal of subtitling all programmes by 2008 whilst minimising the additional costs involved.

INTRODUCTION

Subtitles for the hard of hearing were in the original specification for teletext, defined in the early 1970s. Since then, not only has the amount of television increased rapidly, but the obligations on the broadcasters to increase the proportion of subtitled material have also increased, with a UK target of 100% before the end of the decade.

Subtitles can be divided into two categories: those that are generated for pre-recorded programmes and are played out automatically as the programme is broadcast, and those that are generated live because the programme is not based on a script. BBC R&D Department has already applied innovative techniques to the production of pre-recorded subtitles (1).

This paper describes a project to overhaul the production of live subtitles, allowing more efficient use of subtitling staff and the widespread deployment of speech subtitlers to supplement stenographers. It also provides the facility to re-use the subtitles as a valuable source of programme metadata, and the versatility to adapt to the changing requirements of the BBC as a large national broadcaster.

REQUIREMENTS

A survey of the BBC's subtitling arrangements revealed that simultaneous live subtitling of some 25 television services is routinely carried out, with more during special occasions, such as sporting events. Regional opt-outs from national broadcasts produce peaks in demand, requiring a subtitler located in every region to subtitle a service simultaneously. Most of the subtitles are provided by stenographers, using phonetic principles to produce text rapidly from a "steno" keyboard and translation software. Broadcast stenography requires a high degree of skill, and stenographers are in short supply. Speech recognition (re-speaking) is also being used, but the systems available are expensive and have some limitations.

A completely new live subtitling system was proposed, with the following requirements:

- Production of subtitles by speech recognition and stenography, and by feeds from existing live subtitling applications;
- Subtitle production from any location convenient to the subtitler, including teleworking;
- Scalable, with a low cost per seat and the capacity to subtitle every service simultaneously;
- Subtitle insertion into the broadcast signal at any convenient location;
- Intuitive user interface, hiding network topology and presenting broadcast services by name;
- The ability for subtitlers to hand over between each other (especially useful where two share the subtitling of a long programme);
- Support for subtitling of main and reserve broadcast chains;
- Support for subtitles to be distributed to more than one service during simulcasts;
- Production of plain text for use elsewhere;
- Monitoring and control facilities;
- Reliability of a broadcast-critical system.

SYSTEM ARCHITECTURE

It can be seen that the project is an exercise in distributed computing, and to take advantage of a large amount of accumulated knowledge in this field, it has been built using CORBA (2), over the BBC's TCP/IP networks. CORBA is a system of "middleware" which enables any number of software components written in different languages, compiled for different architectures, and running on different machines, to locate each other and communicate in a standardised fashion. It is efficient in its data requirements, which is important because some home working subtitlers have to connect via analogue dialup as ISDN is not available in their area.

CORBA is a widely used and trusted open standard, with free implementations available: we used TAO (3) as it fulfilled all our requirements, and we already have experience of this implementation at BBC R&D.

SYSTEM COMPONENTS

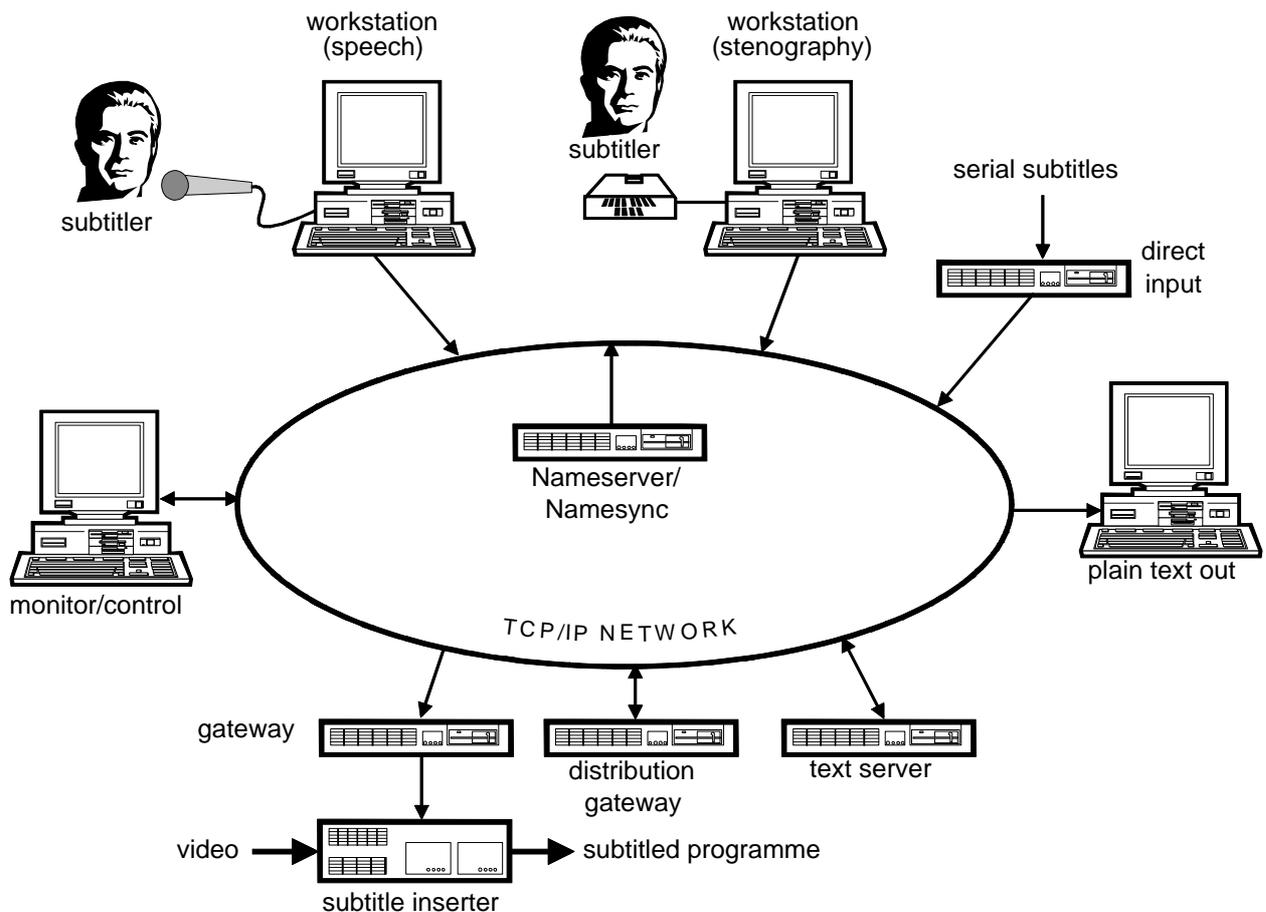


Figure 1 – The Live Subtitling System

The system consists of the following software components (Figure 1):

- Workstations: to generate the subtitles;
- Direct Inputs: to take subtitles from other systems;
- Gateways: to feed the subtitles to the insertion equipment;
- Distribution Gateways: to allow simulcasting of one subtitle stream to several Gateways;
- Text Servers: to provide several streams of plain text for other uses;
- Nameservers: to allow all the distributed components to locate each other;
- Namesyncs: to keep the information in multiple nameservers in synchronism;
- Monitor/Control software: to configure Direct Inputs, monitor the status of Gateways, and view any stream of subtitles passing through the system.

SYSTEM DESCRIPTION

There can be any number of subtitling Workstations, each running on a standard PC. These generate and format the subtitles, and send them to a main and reserve Gateway pair, located by consulting a CORBA nameserver. A Gateway runs as a service on a dedicated PC in a technical area, and drives the subtitle insertion equipment via a serial port. Nameservers are maintained in synchronisation using Namesync processes.

Two Workstations may be connected to a Gateway at once, and the Gateway controls handover between them, allowing one or other to go on-air as appropriate.

In addition to Workstations, there are Direct Inputs, which allow existing subtitling systems to be incorporated, by providing a serial port input which mimics a subtitle inserter. A Direct Input runs as a background process with no user interface, and so the Monitoring and Control software is used to set its destination Gateway and hence the broadcast service to be subtitled. It has a separate connection to the Gateway, which will put its subtitles on-air regardless of the Workstation situation. Normal handover procedure will therefore not be followed when subtitles from a Direct Input are to be used.

Workstations also produce plain text, which can be logged locally and also provided for redistribution elsewhere. Because the text is associated with a particular broadcast service rather than the Workstation which produced it, it is passed to the Gateway where it is then accessed by a Text Server with reference to the service served by the Gateway.

Each Gateway also provides a monitoring port, allowing its subtitles and status to be viewed over the network by the Monitoring and Control software.

Simulcasting is handled by Distribution Gateways. These are further background processes running on dedicated PCs, which appear to Workstations as normal Gateways (and thus as another broadcast service to the subtitler), and to normal Gateways as Workstations. They can therefore be slotted in to allow one source of subtitles to be fanned out to several broadcast services. The Monitoring and Control software is used to configure the simulcast provided by each Distribution Gateway.

Reliability Considerations

Nameservers

It is essential that there is more than one nameserver, running on separate (and preferably physically remote) machines, to avoid a single point of failure. On startup, Gateways register themselves with all known nameservers, and on shutdown, they de-register. To locate Gateways (and hence broadcast services to subtitle), Workstations can work through a list of nameservers in the event that one is not responding.

This creates the problem of maintaining synchronisation between the nameservers in the case of equipment or network failure. Often nameservers are arranged as a hierarchy of masters being duplicated by slaves, with slaves taking over if masters fail. This approach creates difficulties restoring the hierarchy when nameservers are restarted. Instead, the scheme adopted here is to run a Namesync service alongside each nameserver (which can be safely done on the same machine). Namesync regularly updates its nameserver with any additional Gateways found on the next available nameserver in a list of addresses provided.

By this method, nameservers can be configured to synchronise each other in a ring, which automatically shrinks to exclude any failed nameservers, and grows again when they are reinstated. Before adding a new Gateway to the nameserver, Namesync checks that the Gateway is alive, to avoid stale details being propagated. However, it does not check the Gateways already registered, because a momentary loss of communication while

Namesync is active could then lead to a functional Gateway becoming isolated until it is restarted. Genuinely stale details will instead be overwritten when a failed Gateway restarts.

Gateways

Each broadcast service has a main and a reserve Gateway, which subtitle the corresponding programme chains. Several Gateways can run on one PC, but main and reserve Gateways for the same service should be on separate machines to avoid a single point of failure.

The subtitles are sent from the Workstation to each Gateway separately, so that failure of one will not result in loss of subtitles to the other. Moreover, both Gateways perform the control functions of handover, etc, so that the Workstation can still operate properly if one stops responding. Gateway processes are multi-threaded to prevent them locking up if one of the connected devices stops responding.

The plain text output is only made available as a single network stream, to reduce the risk of overloading the Gateway with multiple requests for text. For the same reason, there is only one monitoring output.

Distribution Gateways

These contribute to the reliability of the system by relieving the Workstation of the task of having to send subtitles to many destinations during simulcasts. This would not only increase the load on the Workstation, but would also increase the amount of traffic on its network connection, which may be low bandwidth if the subtitler is working from home.

Distribution Gateways are duplicated and multi-threaded, as with the normal Gateways, and those for the same service should be run on separate machines. Care has to be taken to ensure that their configuration (the list of services to which they pass on subtitles) is synchronised.

Workstations and Direct Inputs

It is obviously impractical to duplicate the Workstation a subtitler is using, beyond having a standby available. A Workstation, however, must cope with the loss of one Gateway, and is therefore multi-threaded, and will silently try to reconnect to a failed Gateway at regular intervals. If both Gateways fail, the Workstation “disconnects” from the broadcast service and informs the subtitler, because there is no point in continuing to subtitle.

It is possible for a subtitler to take control of a service that is already being subtitled. This will be necessary if a Workstation fails, or its user does not respond. Moreover, a third subtitler has the ability to disconnect one of the two Workstations already connected to a service, allowing emergency access to a service which may otherwise be blocked.

THE WORKSTATION

The Workstation runs as an application on the subtitler's PC, and is used either by a stenographer or a speech subtitler. In stenography mode it takes its input from a stenography translator application. In speech mode, it uses IBM ViaVoice, a mass-market PC dictation package which is entirely controlled by the Workstation via its API.

To subtitle a programme, the subtitler first connects to a broadcast service chosen from the list of names retrieved automatically from a nameserver. A second step is required to go on air, which will involve a handover procedure if someone is already subtitling on that service. Initiating a handover gives a signal to the other subtitler's Workstation that someone is waiting to take control, and the other subtitler hands over when they are ready.

Subtitlers listen to the programme sound, and watch the picture (if available) to check if the subtitles need to be moved because they are obscuring an important part of it. Subtitle colour can be changed to indicate different speakers, new lines can be started, and the subtitle can be cleared from the screen if required. Home workers may monitor the picture off air, but if possible will have a separate audio feed via ISDN (reverting to analogue telephone in case of failure) to avoid the delays in the transmission chain. This audio feed is automatically switched to the correct studio output at the broadcast centre. With audio supplied in this manner, subtitling can be carried out anywhere in the world.

The Workstation is capable of generating a log file of the subtitle text, to be stored locally or on a networked drive or remote computer if the network supports it.

Speech Mode

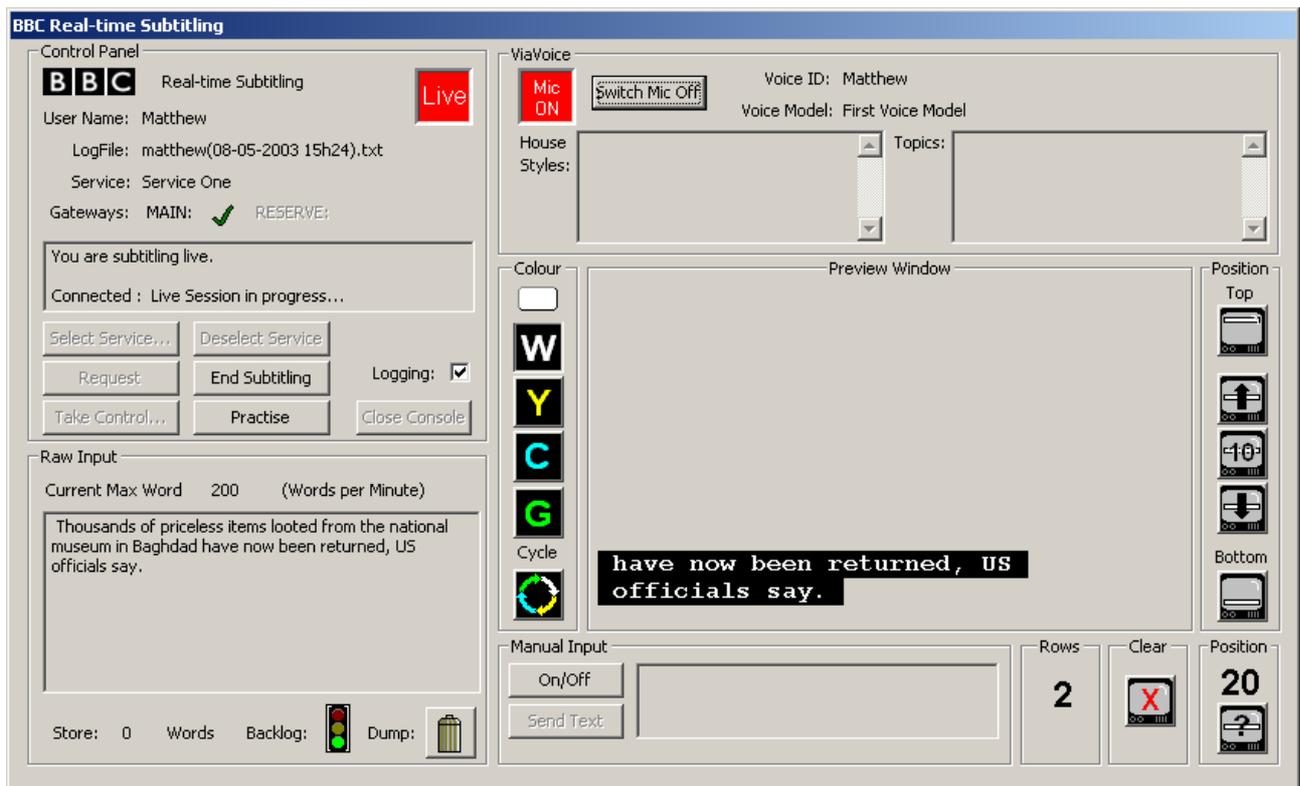


Figure 2 – Workstation in speech mode

Figure 2 shows the appearance of the Workstation when it is being used by a speech subtitler. Speech subtitlers will listen to the programme on headphones and will re-speak the words, précising if necessary. Because they have spent a considerable time training the speech recogniser, and background noise and varying acoustics are eliminated, about 98% recognition accuracy can be achieved, which is on a par with stenography. The PC keyboard or a separate keypad can be used to change colour, move the subtitles and insert punctuation, thus increasing the speed at which subtitles can be generated.

In speech input mode, the Workstation shows the current voice ID, model and topic list, as these are critical for accurate recognition. Recognised words appear in the Raw Input box, and text can also be inserted here a sentence at a time from the Manual Input box. The text is buffered, and released at a preset word rate to be formatted into “snake” subtitles which appear in the Preview Box and are sent to the Gateways. Before formatting, the text passes through filters which intercept commands, and apply any number of House Styles. House Styles do such things as hyphenate particular word pairs, or remove unacceptable words.

Stenography Mode

Figure 3 shows how the Workstation appears in stenography mode. It is designed to be small and simple, because most of the activity will be occurring in the stenography translator. There is a large status indicator, a reassurance in text, giving the names of the other subtitler on that service where applicable, and a graphical display showing where and in what colour the subtitles will appear on the television screen. The two control buttons can be operated from the steno keyboard, and change their function depending on the status of the Workstation. Commands to change colour and position are also issued from the steno keyboard.



Figure 3 – Workstation in stenography mode

ROLL-OUT

The system has to be installed into an existing subtitling infrastructure with minimum disruption.

Gateways and Nameservers

At the time of writing, 80 Gateways are installed on 10 rack-mount PCs located in the Digital Technical Area at BBC Television Centre, London (Figure 4). Their outputs pass through a serial (RS232) router, to enable the system to be activated in phases.

The router outputs are multiplexed into the Energis SDH network to be distributed to subtitle inserters around the country. As this facility was already available and so did not represent an additional cost, the Gateways have been co-sited to reduce the number of machines required and to facilitate maintenance.

There is installed capacity for 40 broadcast services (main and reserve), and 30 are occupied, including three temporary simulcast services. These are conventional Gateways whose outputs are fanned out by the serial router as required, because the Distribution Gateways are to be installed at a later date.

Figure 4 also shows the four nameserver/Namesync machines. Eventually, these would be spread out around the country to protect the system against network splits as well as hardware failure.



Figure 4 – Gateways (left) and Nameservers (right) in BBC Television Centre

Workstations and Direct Inputs



Figure 5 – a speech subtitler at work

The Workstation and ViaVoice have been integrated into the BBC Desktop IT infrastructure, and are automatically installed onto machines used by subtitlers. This provides security as well as versatility, because only authorised users can run the application. Initially there are about 40 users in BBC premises, with home workers and approximately 10 stenography-only Workstations to be added to this. Figure 5 shows a speech subtitler using a workstation in BBC Television Centre.

There are approximately 15 Direct Inputs, running on dedicated PCs where there is more than one co-sited; otherwise, on desktop PCs already used for subtitling purposes.

CONCLUSION

An overhaul of the BBC's live subtitling infrastructure has been achieved by designing a system of networked workstations and gateways which allow subtitlers anywhere to locate and subtitle any broadcast service, using speech recognition or stenography, at very little incremental cost. The system is resulting in more efficient production of live subtitles, as well as providing a valuable source of programme metadata. It will help the BBC to subtitle all of its television output, as well as being sufficiently versatile to accommodate changes to the services in the future.

REFERENCES

1. Assisted Subtitling. BBC R&D Annual Review 2000-2001, pp. 31 to 33, http://www.bbc.co.uk/rd/pubs/rev_01/bbcrdar01part2.pdf
2. <http://www.corba.org>
3. <http://www.cs.wustl.edu/~schmidt/TAO.html>

