



**FOR INFORMATION**  
**September 10,1999**

---

***WebMessenger: Developed***  
**Gives Animated Computer Graphic Characters Synthesized Speech with Emotions**

**\*\* Assists in Making Multimedia Content Familiar and Appealing \*\***

NTT has developed a new generation Internet interface, *WebMessenger*, which provides emotional expression in animated CG with synthesized speech that are capable of flexible prosodic adjustment. *WebMessenger* is designed to work with popular WWW browsers.

The spread of the Internet has introduced the age of information sharing. Individuals as well as organizations now can easily open various homepages, create content, and send information over the net.

Yet, looking at the ways information is sent, text is still the main medium. Though computer-generated animation is becoming popular, the animation of moving pictures is slow, and the synthesized voices have limited variation. The Internet has a long way to go before it can provide multimedia content that is familiar and attractive.

*WebMessenger* lets you edit the prosodic qualities of synthesized speech so that it can express emotions realistically. In addition, it allows you to link synthesized speech to animations of CG characters. Furthermore, *WebMessenger* recreates animated sequences without sending the raw images, i.e. far less data is actually transmitted. Thus, the user can enjoy animated CG images on an Internet terminal without the irritation of slow or jerky replay.

This new technology is an outgrowth of the wealth of software developed by NTT Cyber Space Laboratories to assist in the creation of multimedia content. Such software includes a system for synthesizing speech from text data, as well as Sesign98, a synthesized speech design tool. To those fundamental technologies, we have added a new mechanism for precisely but flexibly synchronizing speech to moving pictures.

The new technology is ideal for such applications as a friendly interface for tele-education systems, and a speaking agent that can read out the text of a homepage for easier understanding.

We will exhibit this new technology at ICCC'99 EXPO (\*) from September 14th through 16th.

<System Configuration> ([See the System Configuration Diagram.](#))

Seign98 offers two modes: automatic conversion and manual adjustment. The latter allows the creator to manually adjust the intonations and the speed of the speech according to his/her taste. Our newly developed content creation tool software, WebMessenger-Creator, creates the content, including the animated pictures and the picture-speech linkage. The speech data can be combined with any animated CG images and adjusted as desired. (The current system has ten CG characters, each with fifty to sixty expression patterns.) Each set of synthetic unit data for a synthesized speech is represented using a synthesis unit index, and each animated CG image has a moving picture index. Once multimedia content is created, just its index information is attached to an HTML document and sent over the net. This way, speech and moving pictures can be transmitted using very small amounts of information. The WebMessenger-Player uses this index information to obtain the speech and moving picture data. Obviously it is necessary for the user's PC to hold the data set of WebMessenger (including synthetic unit data and moving pictures) to allow playback. The speech and pictures are reproduced with good fidelity.

# The "synthetic unit data" is a set of phonemes that is used in synthesizing speech. As the speech synthesis engine grows more sophisticated, this data will be updated to yield more expressive synthesized speech.

### <Major Features>

- 1) Animated CG picture motions and synthesized speech are precisely synchronized.

Conventional similar software products synchronize speech to moving pictures roughly. Our new technology is both precise and flexible; it lets you set the timing exactly as you want. For instance, suppose you have a CG character that makes a greeting gesture. You can set precisely when this character should say hello during the gesture.

- 2) You can sharpen the emotion expressed by using the speech editing functions provided.

For example, you can set the tonal quality, speaking speed, intonation, etc. just as you can choose the color, font, size, etc. of text in word processing software. For instance, you can choose a whispered "Good morning" or an enthusiastic "Gooooo Morning!" at will.

- 3) You can combine speech and animated pictures as you like.

People occasionally show body language that runs contrary to what is being said. With WebMessenger, you can combine the verbal language and body language in any way desired. For instance, you can make a CG character say, "I'll do my best," while the character shows reluctance in his gestures. This lets you create more "human" expressions, and it thus enriches your level of communication.

- 4) *WebMessenger* transmits images using much less data.

What *WebMessenger* actually sends over the net is the index information of synthesized speech and moving pictures. Thanks to this system, we can send the same picture with the same quality using only 1/120 the amount of data of a conventional image transmission system.

## <Technological Key Points>

- 1) Finely adjustable synchronization between synthesized speech and animated CG pictures

Our new technology achieves this thanks to its description format, which can describe speech in great detail, as well as motion information of animated CG images. The speech information can contain such details as phonemes and intonation, and this enables you to create speech that can express various emotions.

- 2) Flexible editing of synthesized speech using *Sesign98*

The speech editing functions of *WebMessenger* are provided by *Sesign98*, a synthesized speech design system developed by NTT. *Sesign98* lets you adjust the loudness, pitch, quality, and other characteristics of synthesized speech easily. All you have to do is change the various parameters through the GUI tools. Also, the tools let you make a library of intonations. These features enable anyone to create speech easily. *Sesign98* has many other potential uses in addition to *WebMessenger*.

## <Major Uses>

Since *WebMessenger* is designed to provide a high degree of creative freedom, there are a great number of possible uses. In particular, it is ideally suited for a closed network of users such as a membership system, because the sender and the receiver share the same programs and data. Shown below are two examples of *WebMessenger* as used in education.

- 1) CALAT, a tele-educational system

CALAT (Computer Aided Learning and Authoring Environment for Tele-Education) is a tele-educational system developed by NTT. Its most prominent feature is the ability to provide the learning environment best suited to each student. For example, the system adjusts the progress and learning content to each student's learning level and understanding. Using *WebMessenger* as the interface of CALAT, allows encouragement and evaluation to be presented to the students using heart-warming animated CG pictures. This can stimulate their motivation to learn.

- 2) Cyber class diary system

*WebMessenger's* ability to synthesize speech from text can be used to create cyber class diaries. *WebMessenger* automatically converts any text into speech, which then can be infused with the desired emotions. In addition, it lets you integrate photos and pictures into the speech. This results in diaries that are much more expressive than the conventional, text-alone reports. Diaries made with *WebMessenger* can communicate the classroom experience of animal observations, study trips outside the school, and many more events, along with the emotions accompanying the experience.

\*ICCC'99 EXPO:Sponsored by the International Council for Computer Communication, the International Conference on Computer Communication has been held every other year since the first conference in 1972, and is primarily for communications operators. This year, Japan will host the conference for the first time since 1978. ICCC'99 EXPO is an exposition that accompanies the conference. Its theme is "Various Developments Based on Digital Integration of Computers, Communications, Broadcasting, and Consumer Electronics."

Date of the exposition: 10:00 am to 5:00 pm, Tuesday, September 14th through Thursday, September 16th, 1999

Site: Exhibition Hall, Tokyo International Forum (For further details, visit our Web site at [http://www.convention.co.jp/iccc\\_j/](http://www.convention.co.jp/iccc_j/) .)

**Attachment**

- [System Configuration](#)

For further information, please contact:

Kenya Nakatsuka  
Press Relation  
Nippon Telegraph and Telephone Corporation  
Telephone:(03)5205-5550  
E-mail: k.nakatsuka@hco.ntt.co.jp



[NTT NEWS RELEASE](#)