

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

FG AVA TR

Version 1.0
(10/2013)

Focus Group on Audiovisual
Media Accessibility
Technical Report

**Part 6: Final report of activities:
Working Group C "Visual signing and sign
language"**



FOREWORD

The procedures for establishment of focus groups are defined in Recommendation ITU-T A.7. The ITU-T Focus Group on Audiovisual Media Accessibility (FG AVA) was proposed by ITU-T Study Group 16 for creation in-between TSAG meetings and it was established on 22 May 2011. The Focus Group was successfully concluded in October 2013.

Even though focus groups have a parent organization, they are organized independently from the usual operating procedures of ITU, and are financially independent. Texts approved by focus groups (including Technical Reports) do not have the same status as ITU-T Recommendations.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Technical Report may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU-T Focus Group participants or others outside of the Technical Report development process.

© ITU 2014

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of contents

Introduction..... 4

Question 1:What is the state of the art of visual signing and sign language? 4

 1.1 Features of sign language and the key stakeholders..... 4

 1.2 Existing work in the area/domain..... 4

 1.2.1 Broadcasting services..... 4

 1.2.2 VOD service..... 5

 1.2.3 Supplementary sign language interpretation service..... 5

 1.2.4 Other services..... 5

 1.2.5 Technological developments of automatic translation..... 5

 1.3 Examples of good practice 5

Question 2:What is the vision for 2015 and 2020 for visual signing and sign language?..... 6

 2.1 Scenario explaining how the needs and interests of persons with disabilities (in its broadest sense) would be addressed in future (primarily for 2020)..... 6

 2.2 Design and production processes 6

 2.3 Key technologies 6

 2.3.1 Programme production technologies using broadband networks will be available 6

 2.3.2 Precision improvement of automatic translation from text to sign language and the increase in the topical domains that can be used 6

 2.3.3 Transmission and reception technology for closed sign language service 7

Question 3:What are the barriers in 2012 that currently prevent these visions from becoming a reality? 8

 3.1 Barriers related to current design and production processes 8

 3.1.1 Technology for closed sign language service has not yet been established 8

 3.1.2 There is not a sufficient number of proficient sign language interpreters..... 8

 3.2 Barriers related to current technologies..... 8

 3.2.1 Transmission and reception technologies for closed sign language service have not been established 8

 3.2.2 Research on the grammar of a sign language, automatic translation, and CG generation are definitely not mature 8

 3.3 Barriers related to current business models for accessible digital media..... 9

 3.3.1 Social awareness of the importance of sign language is not sufficient..... 9

Question 4:What actions are needed to break down current barriers to make the 2015 and 2020 visions a reality?..... 10

4.1	Actions related to current design and production processes	10
4.1.1	Guidelines on producing sign language interpretation programmes should be established.....	10
4.1.2	Education and training systems of proficient sign language interpreters should be established	10
4.2	Actions related to current technologies	10
4.2.1	A sign language program production system based on a broadband network should be developed.....	10
4.2.2	An IBB technology for closed sign language interpretation should be developed.....	10
4.2.3	Research on technologies for CG sign language generation should be accelerated	10
4.3	Actions related to current business models for accessible digital media	11
4.3.1	Social awareness of deaf support systems, such as those involving sign language interpreters or translation technology, should be promoted.....	11
4.4	Actions related to current legislative and regulatory frameworks needed to support these visions	11
References.....		12

Summary

This Technical Report of FG AVA was prepared by Working Group C "*Visual signing and sign language*" and outlines the conclusions of the work of this working group.

Introduction

FG AVA Working Group C discussed services based on sign language both by human sign language interpreters and automatically generated CG sign language images in order to enrich deaf supporting systems.

Question 1: What is the state of the art of visual signing and sign language?

1.1 Features of sign language and the key stakeholders

Sign language is the first language for deaf people who lost their hearing from birth or at an early age. Spoken language in their country or its caption text is usually their second language. For hearing impaired people who lost hearing in adulthood, however, their spoken language or its caption text is their first language.

There is often confusion with these two groups of hearing impaired people and there is a misunderstanding in that providing captions or sign language interpretation is enough for all hearing impaired people.

There is a variety of sign languages in the world, each of which has its own grammar and vocabulary, and they are usually independent from spoken languages. For example, American Sign Language and British Sign Language are quite different, although the spoken languages in both countries are categorized as English. This is because American Sign Language was invented based on French Sign Language. Even within a country, there are many dialects of a sign language. These differences sometimes make accessibility more complex and difficult.

Non-finger expressions, such as facial expressions and head motions, are as important as finger expressions. Non-finger expressions sometimes express grammatical information or modifying information, whereas finger expressions usually express a noun or a verb. This should be taken into account when the size of a signer's image on a display screen or an expression ability of computer graphics (CG) of a signer's avatar is determined.

1.2 Existing work in the area/domain

1.2.1 Broadcasting services

Sign language programmes for deaf people are provided by many broadcasters with various styles of services.

Picture-in-picture window superposed on an original programme is sometimes used to provide a video of a sign language interpreter. The size of the window, however, is crucial because too small a window makes the signer's image unreadable for deaf people while too large a window will be annoying to ordinary viewers.

Programmes specialized for deaf people usually use a full-size video of an interpreter in a studio with the original programme shrunk in the background. They include ordinary programmes superposed with a video of the sign language interpreter as well as those specially produced for deaf people.

They are usually provided as time-shared with other programmes in the same channel. Some advanced broadcasters provide such programmes via a separate channel in addition to the usual programme channel.

1.2.2 VOD service

Broadcasters or other organizations may provide sign language services through video on demand (VOD). In this case, video providers can set up a dedicated page for deaf viewers, where signed versions of TV programmes or videos are collected.

1.2.3 Supplementary sign language interpretation service

Sign language interpretation service for TV programmes can be supplementary provided through broadcasting channels or via the Internet by an organization independent from broadcasters. Deaf viewers utilize a special receiver which receives both broadcasting channels and supplementary sign language interpretation video, and synthesizes them in it.

1.2.4 Other services

Several types of sign language CG software are commercially available. The Japanese company Hitachi (JP) developed a sign language dictionary and sign language CG authoring tool based on a system called Mimehand. CG animation programmes made with this system are used for guidance at public facilities. One another company Vicom3D has also developed an authoring tool system. The ATLAS project¹ uses a CG engine provided by Vicom3D for sign language video image generation.

1.2.5 Technological developments of automatic translation

Research on technologies for automatic translation from text to sign language and CG generation are being conducted at several organizations such as NHK [1] and [2] in Japan, RAI and Politecnico di Torino [3] in Italy, and Chulalongkorn University [4] in Thailand.

1.3 Examples of good practice

Many broadcasters are producing sign language programmes and news reports, although the quantity may not be satisfactory.

The BBC (UK) offers more than 5% of their programmes with signing. It provides programmes with sign language interpreters via a separate channel, if available, or a different time slot such as late night hours on a mainstream channel.

The BBC also provides VOD service named iPlayer, in which broadcasted signed programmes can be accessed from a dedicated page².

Danish public broadcasters offer a visual signing channel sharing one virtual channel using multiplexing of the digital terrestrial TV service. The signing programme schedule is provided via the electronic program guide (EPG).

The Japanese broadcasting NHK provides specialized programmes for deaf people in its education channel in which specially produced news programmes make up the largest category, in which news scripts are modified so as to be easily translated into Japanese Sign Language.

A non-profit organization named Organization of Broadcasting for People with Disability (OBPD) in Japan provides supplementary sign language interpretation and caption service via satellite. It broadcasts sign language interpretation and captions for programmes which are being broadcasted in terrestrial TV channels. Viewers use a special set-top box which receives both one of the terrestrial TV channels and the satellite channel and synthesizes them in a display screen.

¹ Developed by RAI and “Politecnico di Torino”, Italy.

² See <http://www.bbc.co.uk/iplayer/tv/categories/signed>.

Question 2: What is the vision for 2015 and 2020 for visual signing and sign language?

2.1 Scenario explaining how the needs and interests of persons with disabilities (in its broadest sense) would be addressed in future (primarily for 2020)

In 2020, many broadcasting programmes will be accompanied by sign language interpretation by human sign language interpreters because the number of interpreters will have increased and production will have become much easier with a remote production system.

Automatic translation from text to sign language will be available, although the application domain may be limited, for example, in emergencies such as weather warnings and earthquakes. This technology will enable automatic translation of news text into visual signing even late into the night.

Sign language videos will be provided separately via a communication link while main programmes will be provided through radio waves, communication links, or cable. Viewers will be able to select ON/OFF, size, and location of the sign language video on the display screen using functions of a TV receiver. This sign language service, which is controllable at the receiver side, is called "closed sign language service".

This technology will greatly increase the number of sign language interpretation programmes because it will satisfy both deaf and able-hearing people.

2.2 Design and production processes

Due to the development of broadband networks, sign language interpreters at remote locations will be connected to broadcasters or production studios. This will make the production of sign language interpretation for ordinary programmes much easier and cost efficient.

Translation technology from text to sign language will be improved and personal computer (PC)-based systems of sign language CG generation from caption text will be possible in limited domains.

A speech-to-text conversion system based on speech recognition technology may be connected to the sign language CG translation system. The system can automatically interpret news reports into sign language, though the domain must be limited to urgent meteorological news, for example.

2.3 Key technologies

2.3.1 Programme production technologies using broadband networks will be available

The use a remote production system for sign language interpretation of broadcast programmes will be very effective. When producing a programme with sign language interpretation, an interpreter usually needs to go to a broadcasting station to interpret in front of a camera while watching the original programme. By using a broadband network, an interpreter, even at a remote site, can participate in the production of a sign language programme, which makes production much more effective and efficient.

A good example of remote production of sign language interpretation has been reported by Richard Gatarski in Sweden [5]. The system streams superposed video of a sign language interpreter and sends it back to the source of the stream. This kind of system is also applicable for remotely adding sign language interpretation to conferences and meetings. A higher quality video system may be developed for professional broadcasting use.

2.3.2 Precision improvement of automatic translation from text to sign language and the increase in the topical domains that can be used

Defining a notation system of a sign language is essential for using sign language with computers.

Sign languages usually do not have a written format except drawings or illustration, which makes it difficult to use them with computers. A notation with an alphabet enables this. Such a notation system has been developed by a research team of Japanese Sign Language. This notation system, called IGINDEX [6], can describe both manual description and non-manual signals of Japanese Sign Language.

Automatic translation between different languages requires a corpus, which stores many sentence pairs of the two languages in which correspondence relations among the elements of the sentence pairs are described word by word or phrase by phrase. To improve precision, a larger corpus is preferable.

2.3.3 Transmission and reception technology for closed sign language service

To create a closed sign language service, two key technologies have to be developed. The first one is for transmitting a sign language video signal through the broadband network within a certain amount of delay from the original programme broadcasted through radio waves. The second is for displaying two video streams, the original program and the sign language video, in at least a quasi-synchronous manner.

Question 3: What are the barriers in 2012 that currently prevent these visions from becoming a reality?

3.1 Barriers related to current design and production processes

3.1.1 Technology for closed sign language service has not yet been established

Sign language interpretation service is currently provided by superposing an interpreter video onto the original programme at the broadcasting side. This is called "open sign language service". This open service is very simple because it requires nothing from the viewers to use it. However, open sign language service is annoying for able-hearing viewers who want to watch the original programme. This may be one of the reasons sign language services have not increased.

Some broadcasters provide programmes with open sign language interpretation in another channel. However, the frequency spectrum has become stricter after the digitalization of TV broadcasting; therefore, it may be difficult to maintain additional channels for sign language services in many countries.

3.1.2 There is not a sufficient number of proficient sign language interpreters

Since professional announcers are required to have superb enunciation and live announcing skills, a sign language interpreter for TV programmes is also required to have superb signing skills such as fluency and vocabulary. To increase the number of programmes with sign language interpretation, it is necessary to increase the number of proficient sign language interpreters.

3.2 Barriers related to current technologies

3.2.1 Transmission and reception technologies for closed sign language service have not been established

To create closed sign language services, two pathways from a broadcasting station to the home are needed to separately transmit the original programme and sign language video. Although some digital television standards can send a sequence of images through a data broadcasting slot, the bandwidth is not wide enough to transmit sign language video with sufficient quality for understanding.

Integrated broadcast and broadband (IBB) services, which have recently been put on the market or are still under development, use both a broadcasting channel and a broadband network to separately transmit the main programme and supplementary information. They are hopeful candidates not only for creating closed sign language services but also for improving accessibility of broadcasting in general.

It should be clarified what kind of functions and specifications are necessary for accessibility, for example, synchronization between the main video and supplementary information and synchronisation accuracy.

3.2.2 Research on the grammar of a sign language, automatic translation, and CG generation are definitely not mature

Research on sign languages has been extensive. In some countries, there has been no research because its necessity is not fully understood.

Budget and the number of researchers are definitely deficient. Accumulation of research results, research resources, such as a dictionary and a corpus, and research know-how are not sufficient.

3.3 Barriers related to current business models for accessible digital media

3.3.1 Social awareness of the importance of sign language is not sufficient

Many people think subtitling is enough for deaf people. Because of the lack of understanding of viewers and sponsors, the incentive for deaf support activities based on sign language is not satisfactory. Very few TV programs, digital versatile disks (DVDs), and video streams are provided with sign language interpretation. This results in higher production costs for sign language services and fewer sign language interpreters.

Such costs have to be covered by broadcasters in addition to the burden of current digitalization of television broadcasting.

Question 4: What actions are needed to break down current barriers to make the 2015 and 2020 visions a reality?

4.1 Actions related to current design and production processes

4.1.1 Guidelines on producing sign language interpretation programmes should be established

An example of such a guideline is "ITC guidelines on standards for sign language on digital terrestrial television" [7]. It describes several requirements for adding sign language interpretation to broadcasting programme, such as interpreter qualification, synchronization with the original programme, and image quality of the interpreter on the display.

Another example is available from the web site of the Centre for Excellence in Universal Design in Ireland [8]. This rather extensive guideline describes not only sign language interpretation but also other access services such as subtitles and audio description.

These production guidelines should be compiled for less experienced broadcasters or producers to be able to produce higher quality sign language programmes.

4.1.2 Education and training systems of proficient sign language interpreters should be established

A system of training and qualification is required to increase the number of proficient sign language interpreters. It is important to maintain a certain number of sign language interpreters who have proficient interpretation ability to expand the service. Since professional announcers for broadcasting stations are specially trained in enunciation, proficient sign language interpreters should also be specially trained in signing vocabulary and fluency.

4.2 Actions related to current technologies

4.2.1 A sign language program production system based on a broadband network should be developed

A Swedish group led by Richard Gatarski has already developed a network-based sign language interpretation system. Basically, it can be used for broadcasting stations if stability and image quality of the system are verified.

4.2.2 An IBB technology for closed sign language interpretation should be developed

Technology for developing closed sign language interpretation services will be a key to expanding sign language service. IBB systems, some of which have already been put onto the market or are under development, are candidates for this purpose.

Important elemental technologies for implementing a closed sign language service should be clarified and standardized, for example, synthesizing two video streams, the original programme and that of sign language interpretation, and those of synchronizing these two videos. Important functions and parameters in a TV receiver required to implement closed signing are mentioned in the references.

4.2.3 Research on technologies for CG sign language generation should be accelerated

CG sign language generation needs to be accomplished by researches in various fields such as linguistics of sign language, machine translation technology, and high-quality CG generation technology. The research staff and budgets should be increased. Any kind of budget support, including governmental, for the research is preferable.

It should be noted that there is a variety of sign languages and dialects. Efficient research may be conducted by focusing on the commonality and differences among them.

A sign language is composed of manual motions as well as non-manual motions such as facial expressions and head motions. To make CG generation of a sign language possible, a CG generation engine with higher level of expression ability is required.

4.3 Actions related to current business models for accessible digital media

4.3.1 Social awareness of deaf support systems, such as those involving sign language interpreters or translation technology, should be promoted.

Enlightening able-hearing viewers to the importance of sign language services will lead them to properly evaluate broadcasting stations or sponsors who have already begun trying such services. This will in turn encourage more broadcasters or sponsors to support such initiatives.

Economic support for deaf support systems (sponsorship, donation, governmental support, etc.) will be effective in increasing the number of programmes with sign language interpretation. Incubation activities of production companies producing sign language content for broadcasting programmes may also be necessary.

Once the importance of deaf support systems is socially recognized, the number of TV programmes, DVDs, video streams with sign language interpretation, and conferences with sign language interpretation will increase. This is expected to encourage a better business environment for deaf support systems for sign language interpreters, production companies, and broadcasters.

4.4 Actions related to current legislative and regulatory frameworks needed to support these visions

Standardization of the mandatory functions of broadcasting receivers should be promoted. Specifications of IBB systems for creating a closed sign language service should also be clarified and standardized.

References

- [1] Report ITU-R BT.2207-1 (2011), "Accessibility to broadcasting services for persons with disabilities."
 - [2] Hiroyuki Kaneko, *et al.* "[Sign Language Animation Using TVML](#)", VRCAI '10, Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry, pp. 289-292, 12-13 Dec. 2010, COEX, Seoul, Rep. of Korea.
 - [3] Alberto Morello, *et al.* "ATLAS: Automatic Translation into Sign Languages".
 - [4] Teeranoot Chauksuvanit, *et al.* "Progress on an electronics Thai sign language dictionary development", 27th Annual International Technology and Persons with Disabilities Conference (CSUN 2012), DHH-018 (March 2012).
 - [5] Richard Gatarski, *et al.* "Remote sign language interpretation in live video productions", http://wiki.vidblaster.com/index.php/Remote_sign_language_interpretation_in_live_video_productions
 - [6] Kazuyuki Kanda, *et al.* "Notation System and statistical Analysis of NMS in JSL", Lecture Notes in Artificial Intelligence, Springer LNAI. 2298, pp.181-192, (March 2002)
 - [7] ITC (Independent Television Commission) guidelines on standards for sign language on digital terrestrial television (March 2002), http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/sign_language_dtt/index.asp.html
 - [8] Centre for Excellence in Universal Design, Section "Sign Language Interpreting" of "Guidelines for Digital TV equipment and services", <http://www.universaldesign.ie/useandapply/ict/irishnationalitaccessibilityguidelines/digitaltequipmentandservices/programmecontent/signlanguageinterpreting>.
 - [9] [Organization of Broadcasting for People with Disability, "Listen with your Eyes TV"](#).
-