REPORT ON THE SURVEY ABOUT
LIVE EVENT CAPTIONING/SPEECH-TO-TEXT INTERPRETING

Introduction

IFHOH and WFD have partnered on a subtitling project to develop baseline information about different types of meeting/live event subtitling around the world and to identify key issues in providing this service. A key goal of the project is to identify a list of acceptable speech-to-text parameters and providers for UN activities.

The provision of live-event speech-to-text/captioning involves the transmission of information on a visual display by a captionist/text provider. Live event refers to all sorts of live, in-person events such as meetings or conferences, the classroom, the workplace or ceremonies. The captions may be seen on a large screen, a separate large TV monitor, a smart tablet (such as iPad) or personal computer.

The project advisory committee consists of
Bruno Druchen (South Africa), World Federation of the Deaf

Filip Verstraete (Belgium) – World Federation of the Deaf

Marcel Bobeldijk (Netherlands), European Federation of Hard of Hearing People

Lidia Smolarek-Best (United Kingdom), European Federation of Hard of Hearing People

Louise Carroll (New Zealand), International Federation of Hard of Hearing People

Ruth Warick (Canada), International Federation of Hard of Hearing People

Lauren Storck (United States), CCAC (The Collaborative for Communication Access Via Captioning)

Mark WHEATLEY (Belgium) European Union of the Deaf

Survey

A global survey was prepared by a Subcommittee of Ruth Warick, Lauren Storck and Lauren MacDonald with the assistance of Colin Allen. Rosalind Ho prepared it in Survey Monkey format. All of the work was done on a volunteer basis.

The survey was launched on May 3 with an International Sign version, thanks to Colin Allen, for individual responses. The survey consisted of 29 questions in sections focusing on services, administrative aspects and demographics. As of end of May, 2016, 316 persons responded to the survey.

Respondent Profile

The respondent profile of disability was:
37% Hard of Hearing
34% Deaf
14% Other
10% Deafened
4% Hearing
1% Deaf/blind
Under other, several persons identified themselves as cochlear implant users; some individuals identified themselves as have dual disabilities or as orally deaf.

The user profile of the respondents was:
179   Hearing Aid Users
121   Assistive Device Users
116   Sign Language Users
94   Cochlear Implant Users
48   Oral Interpreting Users
Note: individuals were able to check off all of their applicable forms of access.

The country profile of respondents was:
37% Europe
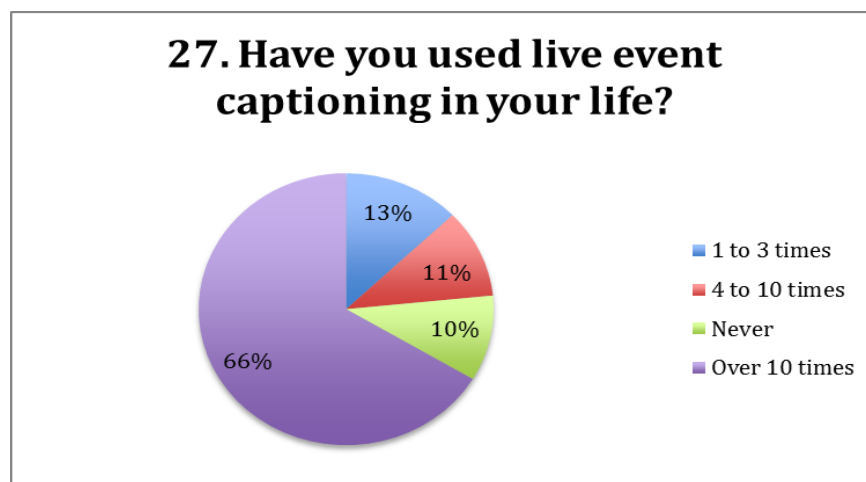30% United States
14% Australia and New Zealand
7% Canada
5% Asia
3% Africa

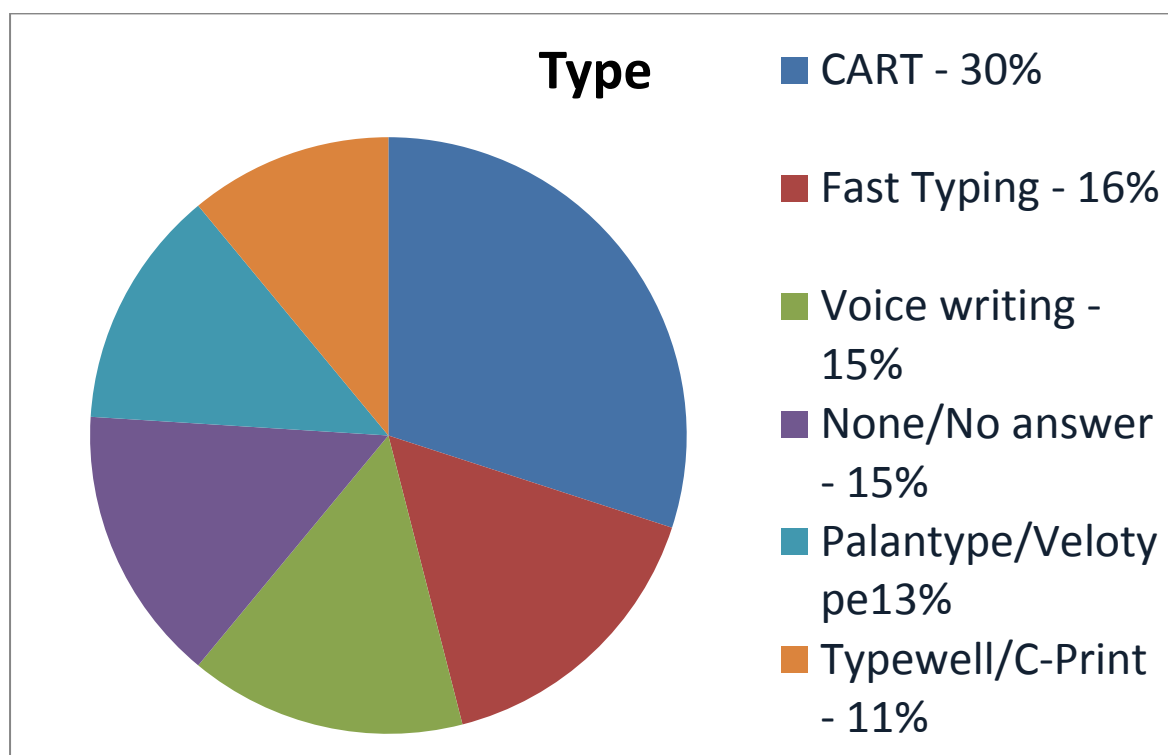Provision of Live-event Speech-to-Text/Captioning
Most of the survey respondents reported that they had used live event captioning/speech-to-text with 77% reporting use 4 times or more. Ten per cent reported never having used it.

Some persons reported use of remote event captioning, but not to the same extent as in-person service provision:

32%   Over 10 times

18%   1 to 3 times

12%   4 to 10 times

37%   Never used

There are different methods of speech-to-text/captioning and the survey endeavored to learn which method was most in use by respondents. Responses showed that there is a wide divergence in methods used. Respondents who were familiar with CART (see definitions below) tended to rate this method above the other methods but this was also the case for those familiar with Palantypist and Velotype. Sometimes it was dependent on the skills of the service provider or on the nature of the event and the rate of speech of the speakers. The answer to the question about the best method is complicated by the availability of individuals trained in specific methods since variations exist as a country level. Further, a number of individuals responded that they lack any form of speech-to-text/captioning in their country.


Type
- CART - 30%
- Fast Typing - 16%
- Voice writing - 15%
- None/No answer - 15%
- Palantype/Velotype 13%
- Typewell/C-Print - 11%

Note: CART refers to Communication Access Realtime Translation (uses a 'court' stenographic machine to produce text). Velotype and Palantype methods involve use of a special keyboard with shortcuts. Typewell and C-Print methods use a regular keyboard with hot keys to provide shortcuts.

The most common types of events for use of speech-to-text/captioning were conferences, meetings, on-line events such as webinars, theatre, post-secondary education and the workplace.

Respondents were asked how they view captions and responses were varied (more than one response was possible):

234	On a screen
165	On own laptop computer
164	On a monitor
126	On a laptop by the provider
78	Mobile phone
68	Tablet

One of the most significant questions of the survey was the question: Do you feel that there are enough captioning providers to meet demand? Almost 60 percent stated that there was a lack of service providers; adding in those who sometimes could not have service, the total number who felt there were insufficient numbers of service providers at times was close to 90 percent.

22. Do you feel that there are enough captioning providers to meet demand?

- No — 59%
- Varies – sometimes but not always — 28%
- Yes — 13%
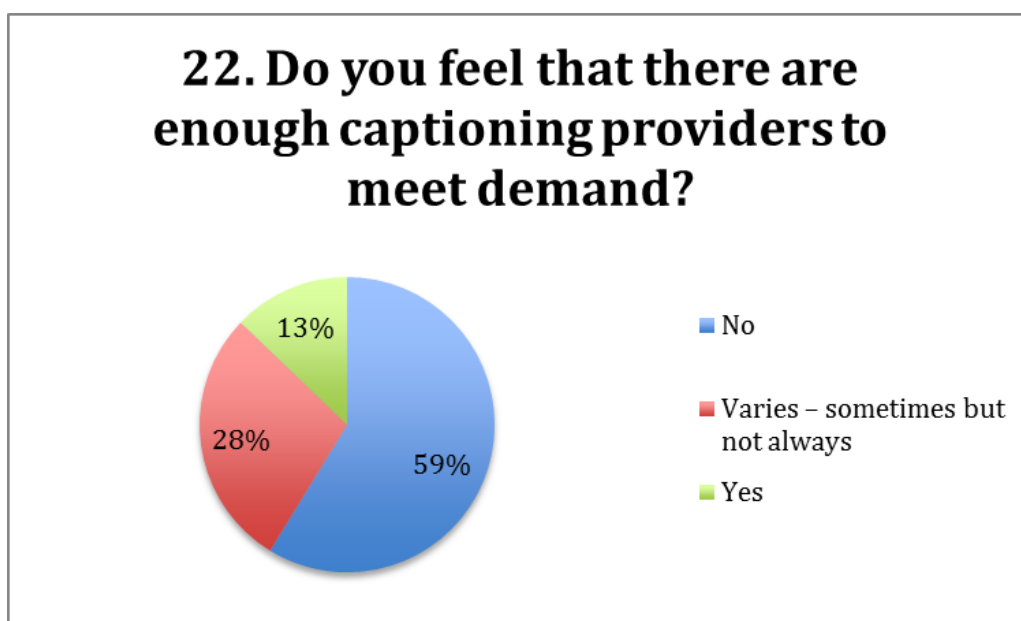
Two-third of respondents had not experienced voice recognition speech-to-text/captioning or did not know what it was. Most of the comments suggested that this form of captioning has quality problems and has not yet reached the potential where it can be a reliable form of access although some respondents were hopeful of future improvements.

Administration

Regarding administrative details, 61% stated that they were responsible for booking a captionist and the other 39% reported that someone else had this responsibility. That could particularly be the case for a conference or a webinar.

Respondents were asked who pays for captioning in their country. In the majority of cases respondents stated it was the Company or institution offering the event or the Government. However, it was also stated that the service agency, health service system, social system, education system or workplace also paid for it. 53 respondents reported that they paid for it themselves at

times. Some respondents reported that there is funding in their country or jurisdiction while other respondents reported no one paid for it in their country. There tended to be a difference on a regional basis with countries of the global south having less service than those of the global north.

Comments

The survey provided ample opportunity for respondents to write in comments. While it would be too lengthy to include them all, among some of them were:

> I am Deaf AND I ALSO have attention deficit disorder. One reason I like CART rather than sign interpreter in some contexts is because CART is more "ADHD friendly" for me, if I zone out for a few seconds due to my ADD then the words stay on the screen so I can catch up.

> Each method has advantages and disadvantages. Velotype is fast but the translators must be well trained and sometimes the German version of the software creates the wrong syllables. Corrections require extra time. Fast typing (using short cuts of office software) is easy to realize but not so fast. A good translator makes good summaries but sometimes I think this thinking somehow should be my work. Voice-writing either needs an extra room for the translators or special equipment like a mask in the face which is a bit irritating. And it too often needs time consuming corrections.

> Depending on the circumstances, the available sources and the audience, different methods may suit differently. In general, it is best to be able to select the particular method that suits best each time.

> We are woefully short of captions in our public and private venues. Many captions we do have are too slow and not accurate enough. It is very difficult to convince hearing people that timely, accur4ate captions are necessary for people with hearing loss. Please help.

> Technology is advancing so fast on many things. I hope that speech recognition can be achieved very soon to provide instant text anywhere.

> I suggest not to jump into conclusions or be convinced too soon that the voice recognition technology is superior or most effective. Stay with the stenography tech for now while studying the effectiveness of voice writing.

> In (country) there are various areas where it is defined clearly who has to pay. But for many types of live events there is nobody who must pay and then it is hard to persuade one. You have to collect donations, etc., tedious job.

> In (country) there was never any live event caption in any conference or meeting, etc. because no one such government, education system, was willing to pay for this service. Only the closed caption company sometimes was willing to provide the live closed caption event for free.

> I'm a kind of person that really concerns on detail. As a HoH person, I had many experiences where one or two words missed and it will change the meanings.

<u>Conclusion</u>

The following are among the issues raised about speech-to-text/captioning. Many of these issues are not new but are confirmed by the survey.

- No common terms and understanding of them – individuals use different terms and there is a lack of a common understanding of the terms.

- Not available in some countries – there is a lack of availability of any service in area of the global south, but even in developed countries almost 90% of respondents found insufficient providers.

- Lack of provision in the language of the country – software challenges of being available in different languages exist.

- Quality of service varies – qualitative differences arise regarding different methods.

- Funding challenges – funding for the service appeared good in some countries but totally lacking in other countries.

- Voice recognition potential unknown and not the answer at this time with some hoping that it will improve in the future.