

FFCDH: Solution to Enable Face-to-Face Conversation between Deaf and Hearing People

Othmar Othmar Mwambe¹, Shelena Soosay Nathan², Toan Nguyen-Duc³, Eiji Kamioka¹

¹Graduate School of Engineering and Science, Shibaura Institute of Technology, Tokyo, Japan

²Tun Hussein Onn University of Malaysia, Parit Raja, Malaysia

³Hanoi University of Science and Technology, Hanoi, Vietnam

Email: mg17007@shibaura-it.ac.jp, sn.shelena@gmail.com, toan.nguyenduc1@hust.edu.vn, kamioka@shibaura-it.ac.jp

How to cite this paper: Mwambe, O.O., Nathan, S.S., Nguyen-Duc, T. and Kamioka, E. (2018) FFCDH: Solution to Enable Face-to-Face Conversation between Deaf and Hearing People. *Journal of Computer and Communications*, 6, 1-14.
<https://doi.org/10.4236/jcc.2018.65001>

Received: April 10, 2018

Accepted: May 13, 2018

Published: May 16, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A real time communication between deaf and hearing people is still a barrier that isolates the deaf people from the hearing world. Over ninety percent of deaf children are born to hearing parents. However, most of them can only learn how to communicate using sign language at school. One of the reasons is that the hearing parents have neither enough time nor support to learn sign language to communicate and support their children. Not surprisingly, the deaf finds difficulties in the oral-only education. Since many other hearing pupils do not even know about the existence of sign language, they cannot communicate directly with the deaf without a sign language interpreter. Therefore, to enable a face-to-face conversation between deaf and hearing people, it is important not only to sustain real time conversation between the deaf and their hearing counterparts but also to equip the hearing with basics of sign language. However, speech to sign conversion remains a challenge due to dialectal and sign language variation, speech utterance and lack of sign language written form. In this paper, a solution named Face-to-Face Conversation Deaf and Hearing people—FFCDH is proposed to address above issues. FFCDH supports real time conversation and also allows the hearing to learn the signs with the same meaning as the deaf understand. Moreover, FFCDH records the speech of the hearing and converts it into signs for the deaf. It also provides deaf with an option to adjust volume of their speech by displaying volume of their voice. The performance of the system in supporting the deaf has been evaluated by using a real test-bed. The obtained results show that English and Japanese daily conversation phrases can be recognized with over 90 percent accuracy on average. The average coherent of simple content is over 94 percent. However, when the speech includes long and complex phrases, the average accuracy and the coherent are slightly lower because the system could not comprehend long and complex context at large scope.

Keywords

Face-to-Face Conversation, Deaf, Sign Language, Finger Spelling, Assistive Technology

1. Introduction

In deaf education, the most formidable challenge to deal with is to find the best means of communication. Since 1500s, deaf people are acknowledged that they are able to be taught. They can learn and understand written symbols by pairing them with the depicted objects. However, many deaf people have been isolated in the society and poorly being taught how to communicate, even though sign languages are available and bilingual-bicultural education is growing widely [1] [2] [3]. However, some deaf students have had a difficult experience to attend regular schools, where verbal or written language is dominantly used as the means of communication and instructional tool. When deaf people study or join a meeting with hearing ones, the communication is generally accomplished by one or more sign language interpreters [4] [5]. The presence of the sign language interpreter aims at bringing the equality to both deaf and hearing people. However, many hearing people know little about the sign language due to fact that verbal language dominates communication means in the hearing world. In the absence of sign language interpreter, the deaf and the hearing cannot communicate directly as a result deaf students are compelled to attend special schools where they can find a community with whom they can use sign language in their daily life conversation. Majority of the deaf individuals use sign language as their first language and learn verbal or written languages later in their lifetime. This makes sign language be an essential tool of instruction for deaf students as they can benefit more from sign language than text in learning environment [6] and sign language has proven better performance in supporting kids at kindergarten [7]. Due to the crucial role played by sign language in supporting deaf people, many studies have focused on utilization of sign language in communication systems [8] [9].

The number of deaf and hard of hearing people, which has increased to over five percent of the global population [10], and the prevalence of deaf children who were born to hearing parents, is higher than ninety percent [11]. Unfortunately, hearing parents of a deaf child commonly do not have either enough time or support to learn a full sign language. Consequently, their deaf children often begin to learn sign language at school. Thereby, the work [12] shows that the students who are being identified as of hearing loss score significantly lower than the hearing ones. In contrast, deaf children born to deaf parents have an early access to sign language. The deaf parents are also proficient at managing the visual gaze of their children, especially, when sharing books or their language and cognitive development duration [13] [14]. Therefore, it is necessary to equip

the hearing parents with a means of communication so that they can communicate and support their children in their critical periods. The communication means should be able to allow any hearing parents to get used to sign language. Additionally, the communication means should provide a unique meaning of the sign to both hearing parents and their deaf children. This is important since the sign language varies not only among separated deaf institutions but also among teachers in the same institution [15].

In this paper, a solution to support real time communication between deaf and hearing individuals is introduced. As an extension of our previous work [16], the proposed solution takes into account dialectal and sign language variation, speech utterance and coherent of the speech to sign conversion process. The current implementation targets English and Japanese. However, this generic solution can be extended for any other target sign languages. The obtained results prove that the system can support real-time communication among deaf individual and their hearing counterpart.

The rest of the paper is organized as follows. In Section 2, the theoretical background study, related work and state of art is present. In Section 3, the detail of the requirement and the design of the proposed speech to sign system architecture will be explained. Section 4 will describe the evaluation on the performance of the system. In Section 5, discussion and implications will be explained. Finally, the paper will be concluded in Section 6.

2. Background and Related Works

Attempt to bridge communication means for the deaf people to communicate with the hearing people started since 1964. Robert H. Weitbrecht, a born deaf, invented the first teletype writer (TTY). The TTY is to allow typed messages to be sent over a telephone line. In the telephone relay centers, messages between deaf and hearing people are relayed by specially trained operators. Also European eSIGN project paid huge contribution to the synthesis of signs in communication systems. With advancement in technology, the number of means of communication for the deaf people has increased and in turn several studies are now focusing on supporting speech to sign translation systems so as to sustain real time communication between deaf and hearing people.

Nguyen-Duc *et al.* [16] proposed a local smart network for speech visualization shown in **Figure 1**. The network supports speech to text conversion in wearable devices. The network uses an automatic speech recognition (ASR) engine for conversion of speech to text. The network was evaluated by testing accuracy of speech to text conversion, calculating the number of spoken words converted into text and the response time. The solution paves a way for real time speech to sign conversation. However, the solution does not implement sign language utilization and the evaluation does not take into account the comprehension of speech to sign conversion process so as to handle dialectal and sign language variation. This makes the solution unfeasible in real time communication [16].

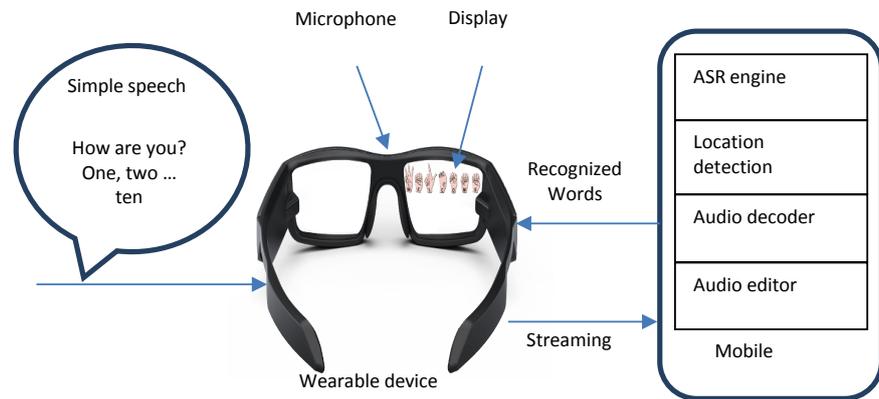


Figure 1. A local smart network for visualizing spoken language.

San-Segundo *et al.* [17] proposed system architecture for translating speech into Spanish sign language in real time domain. The proposed architecture consists of three modules: speech recognition module, natural language translation module and animation module. The natural language translation module operates based on statistical translation and rule-based approach. The translation works well in restricted domain. However, it induces time delay which makes it unfeasible in real time communication [17].

López-Ludeña *et al.* [18] proposed a user-centred methodology for developing a communication system for deaf that consists of four steps: requirements analysis, parallel corpus generation, technology adaptation and system evaluation. However, the methodology relies on parallel corpus generation which can cause delay. This makes it unfeasible in real time communication [18].

Zhao *et al.* [19] proposed a machine translation system from English to American Sign Language (ASL). The proposed system uses input data (text) to derive semantic and morphological information. However, the proposed solution lacks evaluation [19].

Unlike previous approaches, the system architecture proposed in this study is designed to handle dialectal and sign language variation and speech utterance in real time communication. The architecture adopts automatic speech recognition (ASR) engine and uses direct translation that enables it to display signs in real time using finger spelling approach.

3. The Proposed FCDH Solution

Taking into account dialectal and sign language variation, FCDH is proposed to overcome such shortcomings of our previous work [16]. FCDH is designed to enable face to face conversation between deaf and hearing people in real time without sign language interpreter.

In this work, the deaf and hearing people are assumed to understand at least the hand arrangements, or sign alphabet, corresponding to the alphabet letters of the hearing's native language. Additionally, they are assumed to use only one language, *i.e.*, English or Japanese, during their conversation. The communication can follow a turn-talking mechanism [20] [21], or talking in a group simul-

taneously. In the latter case, to record the speech of the hearing person individually, an array of microphones will be used [22]. When the deaf and hearing people do not want to use finger-spelling mechanism [23], they can switch the proposed system to the advance mode in which animated signs from dictionary [24] will be used. In this case, a sign language dictionary must be installed in the proposed system and the dictionary is assumed to have enough vocabulary to support the communication.

3.1. Design Requirements

The proposed system is to let deaf people live almost fairly within the world of others. The key aspects in the requirements aim at allowing the deaf to communicate like hearing people, to have a proper intellectual, and to find themselves happy and productive members of society. Therefore, the proposed system needs to meet the following requirements:

1) Enable the hearing and deaf people to use sign language

The deaf people with or without support of technology, still need the support from the hearing people. Especially when they are at their early ages, the support of their hearing parents plays an important role in their early development. The proposed system must enable hearing parents to support their deaf children. The system needs to have at least two modes, namely simple and advance ones. In the simple mode, finger-spelling will be used to represent words. While the advance mode uses sign language vocabulary to allow the deaf and the hearing people to communicate using natural sign language.

2) Support real time conversation

The proposed system must be able to support the deaf and hearing people to talk to each other directly in a real time manner. When the hearing person speaks, their speech will be recorded and converted to signs for the deaf in real time. When finger-spelling is used, the speed of the conversation can be low. However, when sign vocabulary is used, the speed of the conversion must be fast enough to support normal communication. Besides, the proposed system must support the deaf even when they do not understand the native language of the hearing people whom they are talking to. In these cases, the gestures generated by the deaf will be captured and converted into sound.

3) Eliminate the fear of being labeled “impaired” for the deaf

The deaf have been overlooked and labeled “impaired” or “handicapped”. Thus, almost all of them refuse to use hearing support devices in their daily life [25]. Therefore, the proposed system must take into account this matter not only in terms of technical solution but also its design. Technically speaking, the proposed system must inconspicuously support the deaf anywhere on their daily basis activities particularly when they are at school as well as in the community at large.

3.2. Proposed System

Figure 2 shows the diagram of the proposed system architecture. In order for

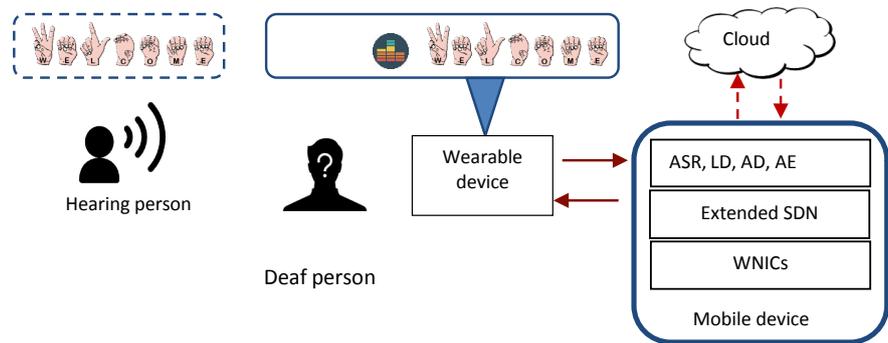


Figure 2. Proposed system architecture.

the deaf person to see the signs correspond to the speech of the hearing person, the speech of the hearing person is recorded by a microphone embedded in a wearable device. Then wearable device worn by the deaf then streams the recorded speech to a mobile device. Although the voice stream does not require high band-width, the audio stream is still encoded for faster transmission and energy saving purpose. An audio decoder (AD) engine on the mobile device will decode received audio and automatic speech recognition (ASR) will recognize spoken words from the received decoded audio stream. If the accuracy of the recognition process is low, an audio editor (AE) is used to reduce the tempo of the speech. When the tempo of the recorded speech is reduced, the audio will be played in the similar way as the speaker speaks with a lower speed. If the ASR still unable to recognize the received audio, it will ask a cloud-based service. In this paper, the proposed system operates only in one mode in which the recognized words are sent back to the wearable device in text format. When the text messages reach the wearable device, the device will display each character in sign alphabets format on a hand-free display using Gallaudet fonts developed by David Rakowski [23]. For example, Optical Head Mounted Display (OHMD), which is used in existing wearable glasses, *i.e.*, Epson, Vuzix and Google glasses, is considered. However; this work can be extended to an advance mode in which a location detector (LD) engine will detect the appropriate sign database to be used. The signs in the advance mode are not sign alphabets. They are sign vocabularies and in GIF format to show the animation of hands. Finally, the matched signs will be sent to the wearable device to display on the hand-free display OHMD.

To let the deaf respond to their hearing counterparts with a suitable volume, the volume of the deaf person's voice is recorded and then is turned into an animation and displayed on the OHMD. This is because the deaf can talk but they cannot hear their voice, and the volume of the voice is commonly very high. The deaf can adjust the volume of his or her voice by observing the animation, which shows how loud the voice is.

Motivated by the face to face real time speech to the sign conversion, the algorithm for the proposed system was designed as shown in Figure 3. The system operates based on the proposed algorithm. The conversion process goes through

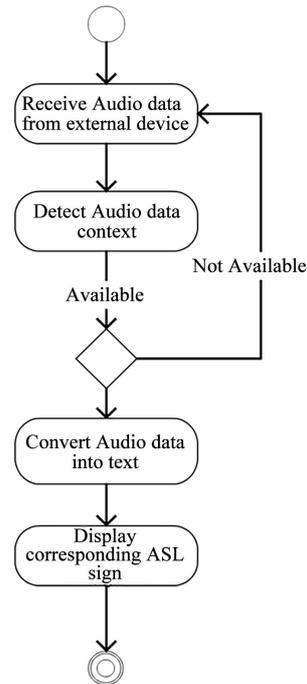


Figure 3. Proposed system algorithm.

two core stages: speech to text and text to sign. The audio data is received by the system, and it is converted into text and then into sign. In pursuance of the speech coherent and handling of dialectal variations, the context awareness is applicable whereby the received audio data is matched with existing context in the cloud server database. The perfect match is converted into text. In favor of common understanding and control over sign language variations, direct translation is applicable whereby sign corresponding to the converted text is displayed. The system enables the end user to see the sign corresponding to the converted text in real time. Thus improving the response time, web socket is applicable whereby quick interactive communications between the end user's mobile device and the cloud server can be sustained. The feasibility of the proposed system has been evaluated and demonstrated in the following section.

4. Evaluation

To evaluate the performance of the proposed system, speech-to-text accuracy and the coherent based on the recognized text will be observed. In addition, the proposed system will be evaluated using English and Japanese languages; evaluation also includes subjects of various nationalities to show system ability in handling dialectal differences. The performance of the proposed system is compared between when a limited database, *i.e.*, offline database, is used as shown in [16] and when a cloud-based database [26] [27], *i.e.*, Google cloud speech recognition, is used.

The speech-to-text accuracy here is defined as the number of correct recognized words over the total number of spoken words. To calculate the coherent of

the recognized content, we firstly split the original content into single sentences. A single sentence is the simplest form of a sentence in term of grammar. For example, a simple sentence must consist of a subject and a verb. The coherent is then defined as the percentage of the number of correct single sentences out of total single sentences.

4.1. Experiment Setup

The evaluation is performed on a test bed as illustrated in **Figure 4**. For the sake of simplicity, the function of the wearable device and the mobile device are implemented using two conventional computers (Core i5-3437U @2.4 GHz processor and 4GB RAM, Windows 7), called WD and MD, respectively. On the WD, an embedded microphone was used to record the speech of the hearing people as well as the deaf. To enable the signs to be able to be displayed in any hand-free display, the area on the screen of the WD used to display the sign has the same size as a hand-free display. The WD connects to the MD using Bluetooth 4.0. On the MD, a web based application has been developed to allow the ASR engine to recognize a speech using either the limited database, *i.e.*, an offline database [16] or cloud-based database (Google cloud speech recognition engine). This is also to make a fair comparison between the previous study [16] and the proposed system in this paper.

4.2. Evaluation Using English

The experiments were performed by twenty-two participants from fourteen different countries including America, England, Japan, China, Nigeria, Saudi Arabia, Malaysia, Cape Verde, Thailand, Mozambique, Sierra Leone, Barbados, Tanzania and Kenya. The participants were aged between 22 and 38, six of them were females. The participants were invited to join two tests. In the first test each participant read a simple content shown in **Table 1**. This content consists of simple daily life conversation words. In the second test, a complex content was used. The complex content consists of scientific English words taken from a computer science lecture class. The complex content has two hundred words, fourteen long and complex sentences.

The procedure of the experiment was as follows: each participant is asked to read the content with the normal speed. The speech is then recorded by the audio handle component in the wearable device and is streamed to the mobile device. The ASR engine on the mobile device recognizes the spoken words from the received audio. The recognized words are sent back to the wearable device. On the wearable device, each recognized word is shown by a sequence of sign characters. The participant also sees the English alphabet character to validate the results.

The accuracy of the recognition processes and the coherent of the recognized content when either limited database or the cloud-based database is used are given in **Figure 5** and **Figure 6**, respectively. The average accuracy when the

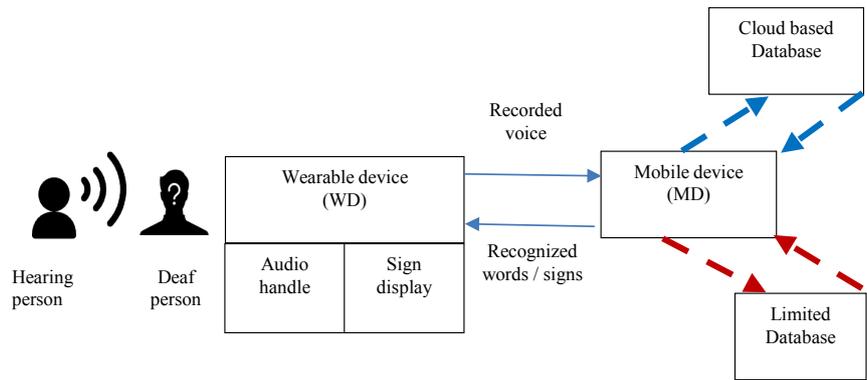


Figure 4. System evaluation environment.

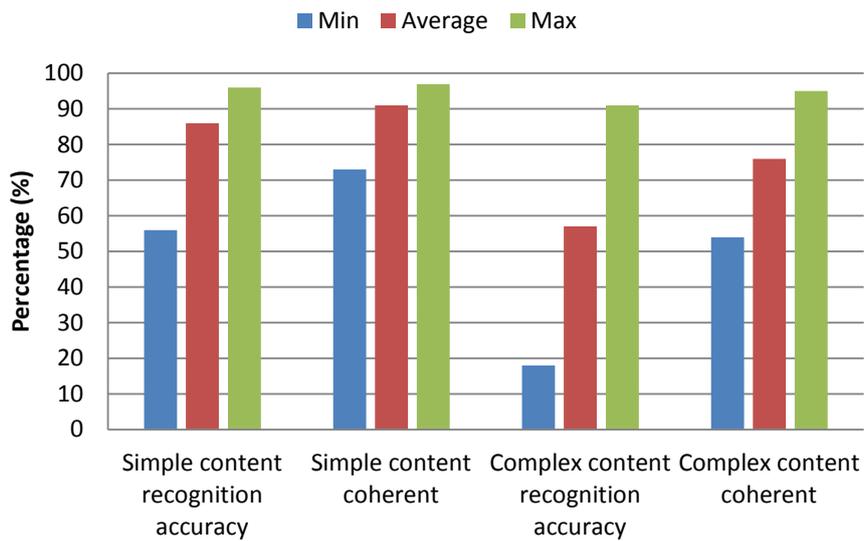


Figure 5. Performance of the proposed system when the limited database is used to recognize English content.

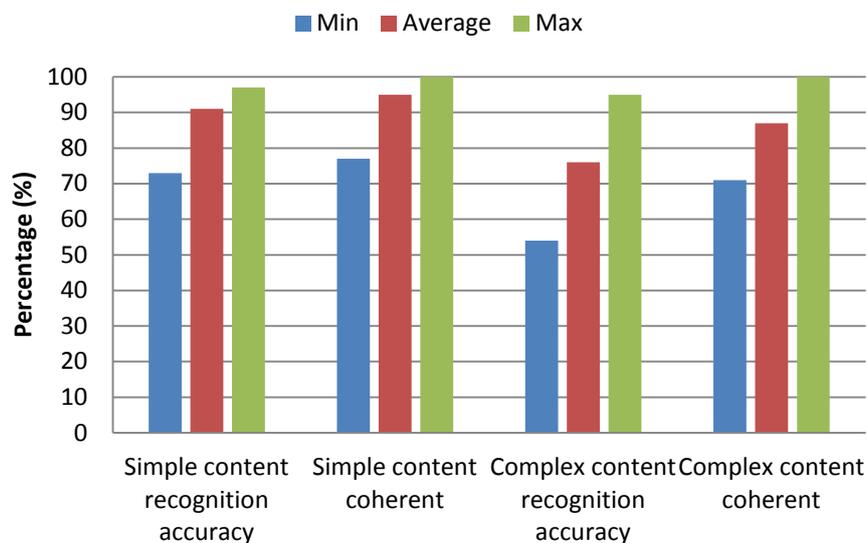


Figure 6. Performance of the proposed system when the cloud based database is used to recognize English content.

Table 1. English contents used in the evaluation.

Type	Content %
Simple	Hi Jimmy, this is Tom from SIT. You and I haven't spoken yet, but I've been doing some research on your company and I think you're a great fit for our networking solutions-sensors. We can provide you with smart sensors and I know you'll be happy if we spend just a couple of minutes discussing how this can help you. When you get this message, please call me back at 02102770033. I look forward to speaking with you, and I guarantee you'll be glad you returned this call
Complex	<p>Ubiquitous Computing is a model in which information processing has been thoroughly integrated into everyday objects and activities. It engages many computational devices and systems simultaneously, and may not necessarily even be aware that they are doing so.</p> <p>It integrates computation into the environment, rather than having computers which are distinct objects. Promoters of this idea hope that embedding computation into the environment and everyday objects would enable people to interact with information-processing devices more naturally and casually than they currently do, and in ways that suit whatever location or context they find themselves in. Ubiquitous computing encompasses wide range of research topics, including distributed computing, mobile computing, sensor networks, human-computer interaction, and artificial intelligence.</p> <p>A context adaptive system typically enables the user to maintain a certain application while roaming between different wireless access technologies, locations, devices and even simultaneously executing everyday tasks like meetings, driving a car etc. This involves making the navigation functionality available for different availability of output devices, input devices and location sensors as well as adapting the user interaction operability to the current speed, noise or operator handicaps while keeping in mind the overall applicability depending on the user preferences, his knowledge, current task etc.</p>

limited database was used is lower with 86 percent as shown in **Figure 5**. The **Figure 5** shows that the simple content can be recognized with 91 percent accuracy on average when the cloud based database was used. Similarly, the obtained results show that the average coherent of the recognized simple content in both cases is 91 percent and 95 percent, respectively. When the complex content was used, the unknown scientific terms were the main reason that reduced the accuracy of the recognition process. When the limited database was used, the average recognition accuracy is as low as 57 percent, thus, the coherent is 76 percent. The average recognition accuracy has been improved to 76 percent, raising the average coherent to 87 percent.

4.3. Evaluation Using Japanese

The experiments were performed by twenty-one Japanese natives. The participants were aged between 20 and 44, two of them were females. The same procedure used in section 3.2 was repeated for this evaluation. Also, the same contents used in the section 3.2 were re-used, but they have been translated into Japanese as shown in **Table 2**.

The accuracy of the recognition processes and the coherent of the recognized content when the limited database or the cloud-based database is used are given in **Figure 7** and **Figure 8**, respectively. **Figure 7** shows that the simple and com-

plex content can be recognized with a similar average accuracy when the limited database was used. The average accuracy was improved in any cases when the cloud based database was used as illustrated in **Figure 8**. Similarly, the obtained results show that the average coherent of the recognized simple content has also been improved by using the cloud based database. The average accuracy in both cases is similar and the average coherent is high because all participants are Japanese native speakers. The coherent is smaller than the accuracy of the recognition process because subjects and verbs in Japanese can be formed by using several words. Therefore, if the whole phrase that represents a subject or verb could not be recognized, the coherent of the recognized content is low.

5. Discussion and Implications

In the obtained results, coherent and accuracy were observed to be of high percentage, this has been interpreted as the ability of the proposed system (FFCDH) to be resilient to dialectal differences and sign variation. Thus, FFCDH can be

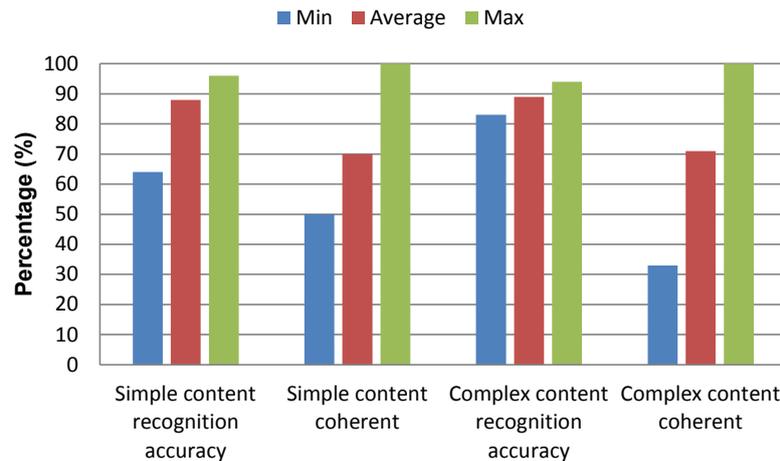


Figure 7. Performance of the proposed system when the limited database is used to recognize Japanese content.

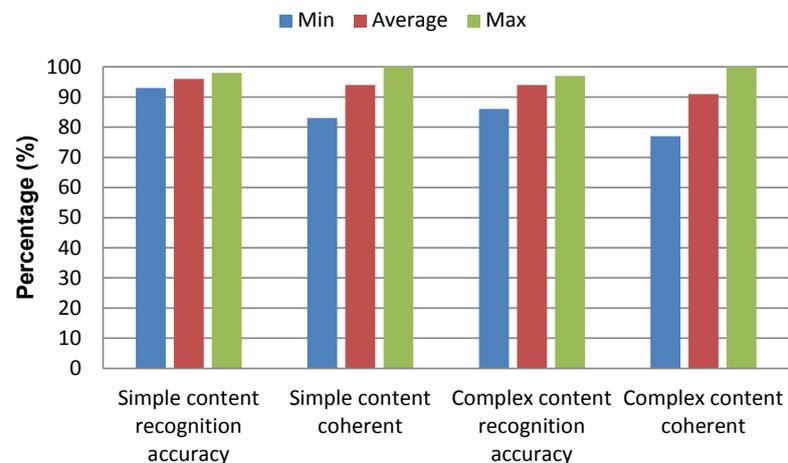


Figure 8. Performance of the proposed system when the cloud based database is used to recognize Japanese content.

Table 2. Japanese contents used in the evaluation.

Type	Content %
Simple	金川様、こんにちは。芝浦工業大学 (SIT) の北見です。お互いまだ話したことありませんが、あなたの会社について少し調べて調査させていただいたら、私たちのセンサーを利用した課題解決に適していると思いました。私たちはあなたにスマートセンサーを提供できますし、どのように手助けになるのか説明する時間をいただければ理解してもらえますと思います。この連絡を確認したら 02102770033 に連絡をしてください。あなたと話すことを楽しみにしているので、折り返し連絡していただくと幸いです。
Complex	現在インターネット上に存在する情報量は爆発的に増加しており、利用者はすべての情報を閲覧することが困難になりつつある。推薦システムはそのような状況のなかで、利用者にとって有用と思われる対象情報、または商品などを選び出し、それを利用者の目的に合わせた形で提示するシステムである。推薦システムの発達に伴い、ユーザは自身の好みに合う情報を高い精度で取得できるようになった一方で、自身の好みに合わない情報や興味のない情報に触れる機会が低下している。このような状況はフィルターバブルと呼ばれる。フィルターバブルはパリサーが2011年に提唱したもので、推薦システムの発達によりユーザが触れる情報がユーザの好みに合う情報ばかりになり、知らず知らずのうちに興味がない情報や新しい情報に触れる機会が失われているような状態を表す。その結果ユーザは推薦システムを利用することで閲覧する情報に偏りや範囲の縮小が生ずる。また、ユーザ側からでは推薦システムがどのようなアルゴリズムにおいて情報を提示するのかを把握することができず、自身に提示された情報がどの程度偏りをもったものか、どの範囲内において収集された情報なのかを理解することができない。ユーザは自分が見たいものだけを見ることができる情報空間を得た一方で自分とは異なる意見や興味外の情報などに触れる機会が失われるようになった。このような状況は社会的リアリティの共有を困難にしたりイデオロギーの極性化、想像性の低下などにつながる可能性がある」と指摘されている。

used to sustain short and long conversations without altering the integrity of the spoken speech. Coherent and accuracy were slightly lower in long and complex phrases than short phrases; this has been interpreted that the proposed system can perform better in daily short conversations than in long speech.

The study also implies that real time speech to sign conversion is feasible and can support daily life conversations between deaf and hearing individuals. Also the study implies that cloud-based databases are comprehensive enough to support speech to sign conversion.

6. Conclusion

In this paper, a solution to enable real time face-to-face communication between deaf and hearing people has been introduced. Experimental results have confirmed that when the database is large enough to support real time speech to sign conversion, English and Japanese languages speech can be recognized with more than 90 percent accuracy on average. The average coherent of the recognized content is also around 90 percent. Using our proposed system, the results have confirmed that the deaf can understand almost all the spoken languages. Also, hearing people can start studying sign language by using finger-spelling.

References

- [1] Mayer, C. and Akamatsu, C.T. (1999) Bilingual-Bicultural Models of Literacy Edu-

- cation for Deaf Students: Considering the Claims. *Journal of Deaf Studies and Deaf Education*, **4**, 1-8. <https://doi.org/10.1093/deafed/4.1.1>
- [2] Mayer, C. and Akamatsu, C.T. (2000) Deaf Children Creating Written Texts: Contributions of American Sign Language and Signed Forms of English. *American Annals of the Deaf*, **145**, 394-403. <https://doi.org/10.1353/aad.2012.0135>
- [3] Mayer, C. and Wells, G. (1996) Can the Linguistic Interdependence Theory Support a Bilingual-Bicultural Model of Literacy Education for Deaf Students? *Journal of Deaf Studies and Deaf Education*, **1**, 93-107. <https://doi.org/10.1093/oxfordjournals.deafed.a014290>
- [4] Witter-Merithew, A. and Dirst, R. (1982) Preparation and Use of Educational Interpreters. In: Sims, D.G., Walter, G.G. and Whitehead, R.L., Eds., *Deafness and Communication: Assessment and Training*, Williams & Wilkins, Baltimore, 395-406.
- [5] Stewart, D.A. (1988) Educational Interpreting for the Hearing Impaired. *British Columbia Journal of Special Education*, **12**, 273-279.
- [6] Marschark, M., Leigh, G., Sapere, P., Burnham, D., Convertino, C., Stinson, M. and Noble, W. (2006) Benefits of Sign Language Interpreting and Text Alternatives for Deaf Students' Classroom Learning. *Journal of Deaf Studies and Deaf Education*, **11**, 421-437. <https://doi.org/10.1093/deafed/enl013>
- [7] Hurdich, J. (2008) Utilizing Lifelike, 3D Animated Signing Avatar Characters for the Instruction of K-12 Deaf Learners. *Proceedings of the Exploring Instructional and Access Technologies Symposium National Technical Institute for the Deaf*, New York.
- [8] Huenerfauth, M., Zhao, L., Gu, E. and Allbeck, J. (2007) Evaluating American Sign Language Generation through the Participation of Native ASL Signers. *Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility*, Tempe, 15-17 October 2007, 211-218.
- [9] Smith, R., Morrissey, S. and Somers, H. (2010) HCI for the Deaf Community: Developing Human-Like Avatars for Sign Language Synthesis. *Proceedings of the Fourth Irish Human Computer Interaction Conference*, Dublin, 2-3 September 2010.
- [10] World Health Organization (2018) Deafness and Hearing Loss. <http://www.who.int/mediacentre/factsheets/fs300/en/>
- [11] Mitchell, R.E. and Karchmer, M.A. (2004) Chasing the Mythical Ten Percent: Parental Hearing Status of Deaf and Hard of Hearing Students in the United States. *Sign Language Studies*, **4**, 138-163. <https://doi.org/10.1353/sls.2004.0005>
- [12] Blackorby, J. and Knokey, A. (2006) A National Profile of Students with Hearing Impairments in Elementary and Middle School: A Special Topic Report of the Special Education Elementary Longitudinal Study. <http://www.seels.net/grindex.html>
- [13] Lieberman, A.M., Hatrak, M. and Mayberry, R.I. (2014) Learning to Look for Language: Development of Joint Attention in Young Deaf Children. *Language Learning and Development*, **10**, 19-35. <https://doi.org/10.1080/15475441.2012.760381>
- [14] Singleton, J.L. and Crume, P.K. (2010) Socializing Visual Engagement in Early Childhood Deaf Education. *21st International Congress on the Education of the Deaf*, Vancouver, July 2010.
- [15] Keep, J.R. (1857) The Mode of Learning the Sign Language. *American Instructors of the Deaf*, Bedford, 133-153.
- [16] Nguyen-Duc, T., Mwambe, O.O., Nathan, S.S., Taiye, M.A., Hussain, A., Hashim,

- N.L. and Kamioka, E. (2017) Visualization of Spoken Language for Deaf People. *6th International Conference on Computing and Informatics*, Kuala Lumpur, 25-27 April 2017, 166.
- [17] San-Segundo, R., Barra, R., Córdoba, R., D'Haro, L.F., Fernández, F., Ferreiros, J. and Pardo, J.M. (2008) Speech to Sign Language Translation System for Spanish. *Speech Communication*, **50**, 1009-1020.
<https://doi.org/10.1016/j.specom.2008.02.001>
- [18] López-Ludeña, V., González-Morcillo, C., López, J.C., Ferreiro, E., Ferreiros, J. and San-Segundo, R. (2014) Methodology for Developing an Advanced Communications System for the Deaf in a New Domain. *Knowledge-Based Systems*, **56**, 240-252. <https://doi.org/10.1016/j.knosys.2013.11.017>
- [19] Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N. and Palmer, M. (2000) A Machine Translation System from English to American Sign Language. In: White, J.S., Ed., *Envisioning Machine Translation in the Information Future, AMTA 2000, Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg, 54-67.
- [20] Sacks, H., Schegloff, E.A. and Jefferson, G. (1974) A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, **50**, 696-735.
<https://doi.org/10.1353/lan.1974.0010>
- [21] Van Herreweghe, M. (2002) Turn-Taking Mechanisms and Active Participation in Meetings with Deaf and Hearing Participants in Flanders. In: Lucas, C., Ed., *Turn-Taking, Finger Spelling and Contact in Signed Languages*, Gallaudet University Press, Washington DC, 73-103.
- [22] Hioka, Y., Kingan, M., Schmid, G. and Stol, K.A. (2016) Speech Enhancement Using a Microphone Array Mounted on an Unmanned Aerial Vehicle. 2016 *IEEE International Workshop on Acoustic Signal Enhancement (IWAENC)*, Xi'an, 13-16 September 2016, 1-5.
- [23] Keane, J., Brentari, D. and Riggle, J. (2012) Coarticulation in ASL Fingerspelling. In: Schafer, A.J., North Eastern Linguistic Society, et al., Eds., *Proceedings of the North East Linguistic Society*, UMass Amherst, Amherst, 261-272.
- [24] Tennant, R.A. and Brown, M.G. (1998) *The American Sign Language Handshape Dictionary*. Gallaudet University Press, Washington DC.
- [25] McCormack, A. and Fortnum, H. (2013) Why Do People Fitted with Hearing Aids Not Wear Them? *International Journal of Audiology*, **52**, 360-368.
<https://doi.org/10.3109/14992027.2013.769066>
- [26] Google (2017) Google Cloud Speech Recognition Engine.
<https://cloud.google.com/speech/>
- [27] Microsoft (2017) Microsoft Cognitive Services.
<https://www.microsoft.com/cognitive-services/en-us/speech-api>