

Converting South African Sign Language to Verbal

Shingirirai Chakoma and Philip Baron

Abstract—There is a significant population of hearing-impaired people who reside in South Africa; however, South African Sign Language (SASL) has not yet been recognized as South Africa’s 12th official language, resulting in slow uptake of this important language. Since most people do not know SASL, there is a need for gesture recognition systems that convert Sign Language (SL) to verbal and/or text to reduce the communication barriers between the hearing and the hearing-impaired. This study presents an application for gesture recognition in converting SASL to both a verbal format and a textual format. By using gesture recognition from a single wearable glove, hand gestures were quantified, categorized, and then converted into an auditory format and played on a speaker, as well as the equivalent textual information displayed on an LCD screen. The complete prototype consists of a wearable glove with a transmitter and an associated receiver box which were all designed to cost less than \$150. The glove consists of five flex sensors that measure the handshape and an inertial measurement unit which measures the hand motion. The handshape and motion data are processed and wirelessly transmitted to a receiver box. This then displays the associated English character on an LCD while also playing the audio on a speaker. The SL converter can convert the 26 letters of the SASL manual alphabet with an overall accuracy of 69%. It can also convert common words and phrases, as well as proper names when fingerspelled.

Index terms— flex sensors, gesture recognition, hearing impaired, South African Sign Language, wearable glove

Open License: CC BY-NC-ND

I. INTRODUCTION

South African Sign Language (SASL) is the language primarily used by the South African Deaf community. Just like spoken languages, Sign Language (SL) is also a natural language and signers use it to learn, communicate, and express their viewpoints. While SASL is still not understood by the majority of South Africans, in 2018, it was officially recognized as a home language in the South African education system [1]. While there are initiatives aimed at expanding the use of SL, there is a dire need for alternative modes of communication for

the Deaf community. Two modes include the conversion of SL to visual text and/or audible voice. Either of these two modes (text or voice) could reduce the communication barrier between hearing and hearing-impaired people. For a SL converter to offer a viable solution, some design goals need to be attained; namely, devices ought to fully represent all the linguistic properties of SL; such devices must be comfortable for the signer to wear and portable to use; and finally, these devices must be affordable.

The conversion process of SL to verbal starts with a hand movement and a measurement of this movement, which must be recognized by a computing device. There are two main types of gesture recognition techniques for hand motion, namely, image-based, and non-image-based. Image-based techniques make use of a camera to capture the hand configurations. Non-image-based techniques make use of sensors to capture hand configurations which are implemented as wearable gloves and bands [2].

Once the gesture data has been captured, it is then processed into meaningful information by microprocessors—acting as the interface between the signs made by the signer and the representation of those signs into a verbal format. An example of a microprocessor that is viable for this application is the Arduino open-source programmable microcontroller board [3]. Arduino-based devices can process the analog information from the wearable glove and convert it into text and speech. In this article, a SL to text and audio converter is presented, which is called the AudibleSigns Sign Language converter.

II. LITERATURE REVIEW

A. South African Sign Language

South African Sign Language, just like any other SL, is characterized by its phonology. The SL phonological parameters include handshape, palm orientation, location, movement, and non-manual features which are not executed using hands, such as facial expressions [4]. As shown in Fig. 1, the SASL fingerspelling alphabet consists of 26 manual letters of the English alphabet. These manual alphabet signs are divided into static and dynamic gestures. The dynamic gestures are for signing the letters “J” and “Z”, as shown by the different handshapes and hand movements. The remaining letters are static gestures shown by handshapes only. Since SL does not have all words represented as signs, fingerspelling is an important part of SL. The English language, for example, has

S. Chakoma is with the Department of Electrical Engineering Technology, University of Johannesburg, Gauteng, South Africa (e-mail: shingiemichael1352@gmail.com)

P. Baron is with the Department of Electrical Engineering Technology, University of Johannesburg, Gauteng, South Africa (e-mail: pbaron@uj.ac.za)

over 400 000 words, although only about 170 000 are in current use [5]. SASL does have over 1000 words, but fingerspelling is still very important for names, places, and objects. Fingerspelling often evolves into generally accepted signs as well. For example, the sign for the month of July, is the fingerspelling of J-U-L-Y. There are many of these cases.



Fig. 1. South African Sign Language manual alphabet [1].

B. Gesture recognition methods

The reason why hand gestures are important in SL is the critical role the hand plays in displaying a significant number of recognizable configurations. Various techniques have been implemented to capture and recognize these hand gestures. Smart gloves, non-wearables, and wearable bands are the types of non-image-based techniques used for hand gestures. Smart gloves usually require wire connections and sensing devices. Sensors in the form of accelerometers, gyroscopes, flex sensors, proximity sensors, and abduction sensors are common in glove applications [6]. These sensors are mounted on a wearable glove, and they aim to precisely calculate the hand and finger orientation and configuration.

Non-wearable sensors make use of radio frequency (RF) signals to detect gestures without physical contact with the human body [2]. An example is Google's *Project Soli*, which makes use of low-power radar sensors for gesture recognition. The radar emits a beam of electromagnetic waves that physical objects (the hand) reflect back to the radar. It can capture information about the object's shape, size, orientation, distance, and velocity [7].

Another method called *SignFi* uses Wi-Fi signals to recognize hand gestures. The signer makes signs in front of a computer at the Wi-Fi station, which in turn sends data packets over Wi-Fi to the computer via the access point. The packets

are then classified by a convolutional neural network to give an output word [8].

Band-based sensors, which is another type of wearables, rely on sensors placed on a wristband. An example of this technique is the *Myo Armband* that makes use of arm muscle patterns to recognize sign gestures. An electromyography sensor on the armband measures the arm muscle's electrical pulses which classify the hand gestures with the help of support vector machines [9].

Image-based techniques make use of a camera as the input device and includes markers, stereo cameras, and depth sensors. Markers are placed on the human hand to detect hand motion. These markers are either reflective—they are passive, and they shine when exposed to a strobe—or they may be active—light-emitting diodes that have a sequence of lighting [10]. A camera captures the hand positions and these positions are interpreted onto a 3D space. Detection and segmentation of the predefined colors from the captured image would show which fingers were active when the gesture was made.

The early researchers who used image-based methods used single cameras in gesture recognition applications. However, these single cameras were found to be inaccurate for vigorous hand gestures due to restricted viewing angles. Stereo cameras were then introduced to capture vigorous hand gestures. They were used to build a 3D environment that has the depth information of objects. A depth sensor is a non-stereo device with its output as 3D information, which is better suited to preventing problems of calibration and illumination—a known challenge with stereo camera rigs [2].

Time-of-Flight technology makes use of depth sensing by measuring the time of flight of short light pulses emitted by an infrared projector, then reflected and captured by a sensor. This technology has been implemented in Microsoft's *Kinect* camera system [11]. The *Kinect* relies on a depth sensor that detects motion and creates a 3D scan of captured objects. It has an infrared projector that illuminates the object and an infrared sensor that reads the infrared light and measures the depth of the pixels of the captured images.

C. Comparison between image and non-image-based gesture recognition techniques

The main advantage of image-based techniques is the use of a camera to detect hand motion, which is preferred over having a wearable glove [6]. The unconstrained nature, in terms of making handshapes, of the image-based technique allows the signer to naturally execute the sign without the hindrance of a glove with sensors worn on the hand. Important depth and color data of images can be obtained by using a camera [12]. If used with the Microsoft *Kinect* camera, there is an increased capability to detect handshapes [6]. Despite its successful application in 3D human action gaming, human skeleton estimation, healthcare [13], sports [14], and facial recognition [15], there is still a problem with its application in vigorous hand gesture recognition.

According to [16], they have found that the Microsoft *Kinect* camera has a low-resolution depth map (640×480) which works well for tracking large objects like the human body. However,

there are some difficulties in detecting and segmenting small objects from images using this low resolution. With the human hand/fingers occupying a small space of the image/frame, the process becomes complex. The signer's position might also change in front of the camera which results in variable hand size. Thus, for camera-based methods to be effective as hand gesture recognition, the signer must keep their hand in front of the camera in the correct predefined position for a considerable amount of time [17]. With SL relying on a fairly rapid and continuous hand motion, this limitation of momentarily paused hand motions is thus not ideal. This critical challenge is overcome by non-image-based techniques—wearing a glove, for example—which does not require the signer to be in the view of a camera, in turn improving the signer's mobility [6].

Non-image-based gesture recognition techniques require physical sensing of the movements using sensors on the person. Historically this was bulky and limited the person's mobility. However, modern-day technology has enabled the design of wearable devices that do not need to have a physical connection to a computer, thus allowing for portability and flexibility of movement. For instance, a wireless data glove offers reasonable flexibility [18]. With the miniaturization of electronic devices, wearable gloves have become lightweight, making them more convenient to wear [6]. Accelerometers and flex sensors, unlike optical sensors in the image-based techniques, are not affected by a cluttered background or illumination fluctuations [16], which are both challenges for the image-based gesture recognition systems.

Despite most sensor-based gloves being unaffected by ambient lighting, the sensors do sometimes require calibration to distinguish gestures. This calibration requirement does hinder the natural gesture articulation [18]. Most modern-day data gloves make use of flex sensors, which are low cost and quite hardy², but these sensors often have non-linearity and provide less sensitivity at minute bending angles. For example, finger pivot joints provide a wide angular movement ranging from small angles to large angles [19], with all these angles being important in the accurate recognition of the hand sign. With the complex phonology of SL and the added challenge of handshape variability, there is a need to improve the linearity and sensitivity of flex sensors if they are to be used with more accuracy.

From the above comparison, each mode of gesture recognition has advantages and disadvantages, but for this study, the non-image-based method of the wearable glove was found to be the most feasible option. The feasibility is defined in terms of overall development cost, affordability for the user, and increased mobility for the signer. In this study, a wearable glove and a converter device utilizing wireless communication were developed. The prototype presented in this article was designed to convert the 26 letters of the SASL manual alphabet and common words and phrases that have specific signs in the SASL dictionary, as well as words that do not have specific

signs in the SASL dictionary. The description of the prototype follows next.

III. CONVERTING SOUTH AFRICAN SIGN LANGUAGE TO VERBAL: AUDIBLESIGNS

AudibleSigns is a device that converts SASL to text and speech to improve communication between hearing and non-hearing people. The main parts of the prototype are the wireless glove and the standalone wireless converter device (Figs. 2 & 3). Wireless technology was used to improve the comfort of the signer by removing any wires from the glove to the converter. The 2.4 GHz nRF24L01 wireless transceiver pair enables wireless communication between the glove and the converter device; thus, this is a two-part system—a transmitter side and a receiver side. The type of microcontroller board used for the glove and converter device is the Arduino Nano, based on the ATMEGA328 microprocessor. The prototype was designed for converting the SASL manual alphabet as well as words that do not have specific signs in the SASL dictionary. The messages were first converted from text to speech using an online text to speech service, and then the audio files were downloaded and stored on a microSD card. Rechargeable lithium polymer batteries were used owing to their smaller footprint and longer lifespan.

A. Method

The scope of the research and the prototype design were centered on four SL phonological parameters of the hand: handshape, palm orientation, movement, and location. This excludes non-manual features or non-hand signs like facial expressions and head movements. The block diagram and the hardware design for the two-part prototype are shown in Figs. 2 and 3.

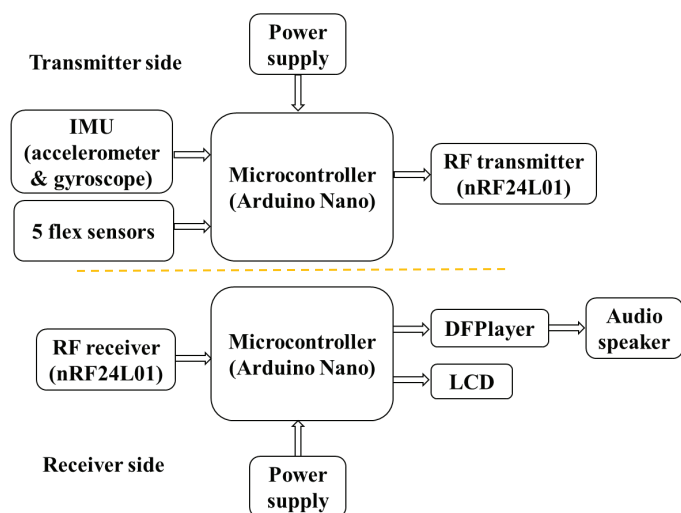


Fig. 2. Block diagram showing the hardware implementation of the AudibleSigns Sign Language converter.

² The resistive flex sensors can be bent, dropped, stood on, and even wetted. However, they do wear out from extensive use which is discussed as one of their limitations.

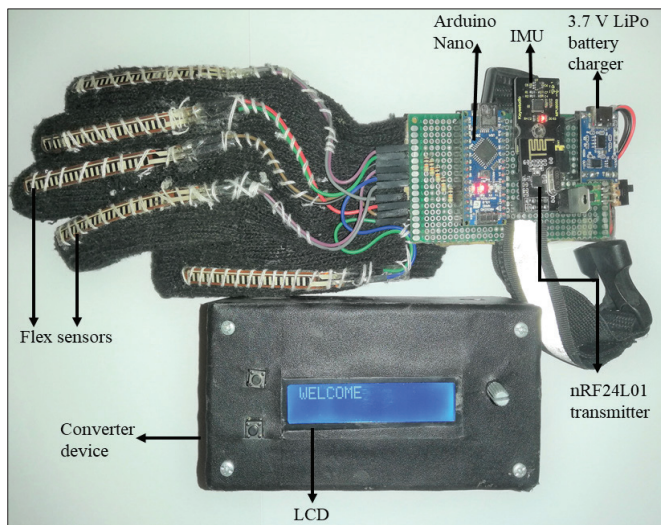


Fig. 3. Image of the glove and converter device of the AudibleSigns Sign Language converter prototype.

B. Transmitter Side

Five flex sensors were sewn on top of each glove finger compartment to allow the flex sensors to capture the handshape by measuring the finger flexion. Flex sensors behave like variable resistors whose resistance is directly proportional to the amount of bend. Each flex sensor is connected between a 5V supply and a 10k Ω fixed value resistor to form a potential divider circuit to create a variable voltage that is dependent on the change in resistance. This output voltage is inversely proportional to the flex sensor resistance. The output from each potential divider circuit, a varying voltage, is then fed into the Arduino Nano, which reads this as an analog signal.

The MPU6050 was used as the inertial measurement unit (IMU) to measure the hand orientation and movement. It consists of a three-axis accelerometer and a three-axis gyroscope that measures the hand's acceleration and rotation, respectively. These were measured with respect to the pitch, roll, and yaw hand orientations as shown in Fig. 4. The acceleration and rotation are both converted to pitch, roll, and yaw angles measured in degrees. The values from the flex sensors and IMU are read as analog signals by the Arduino Nano. The Arduino Nano converts these analog signals to digital data using the onboard analog-to-digital converters. The nRF24L01 transmitter sends the data from the sensors as RF signals to the nRF24L01 receiver on the converter side.

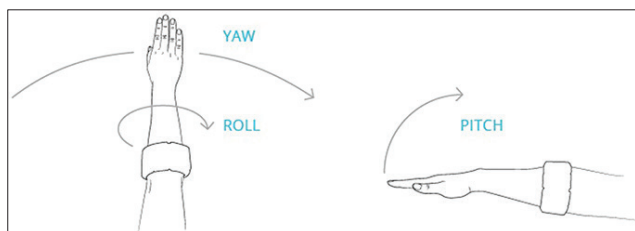


Fig. 4. Pitch, roll, and yaw hand orientations [20].

C. Receiver side

The receiving and converting device uses the nRF24L01 receiver, which picks up the signals transmitted from the glove. A local SASL database was created using the data from the sensors and is stored inside the memory of the Arduino Nano on the receiver side. When a sign is made by the signer, the Arduino Nano compares the received data with the one in the SASL database and looks for the closest match. When there is a match, the microcontroller sends serial commands to the LCD to print the letter or message that corresponds to the sign which is then played audibly on the speaker. The audio messages played by the mini MP3 player are stored on a 2GB microSD card formatted as FAT32. The audio files are played via a serial interface from the Arduino.

D. Power source

The power supply design is the same for the transmitter side (glove) and the receiver side. Both the glove and the receiver device are powered by 3.7V rechargeable lithium polymer batteries. Each battery terminal voltage is boosted by a 5V DC booster to power each Arduino microcontroller on the glove and receiver. The 5V from the DC booster is regulated down by a 3.3V voltage regulator to power the 3.3V-tolerant components like the IMU and nRF24L01 transceivers. Shunt capacitors on the input and output pins of the 3.3V regulator smooth any AC ripple that might be present in the DC supply. The dual power supplies of 3.3V and 5V are manually controlled by slide switches to turn the devices ON or OFF.

E. Functionality

Examples of the performance of the AudibleSigns SASL converter are shown in Figs. 5 to 7. Figs. 6 and 7 show the conversion for words that do not have specific signs in the SASL dictionary. For such words to be converted to verbal, the letters of the words are individually finger-spelled. The LCD displays each letter of the word, one after the other, while the audio speaker mounted on the receiver plays the corresponding sound.

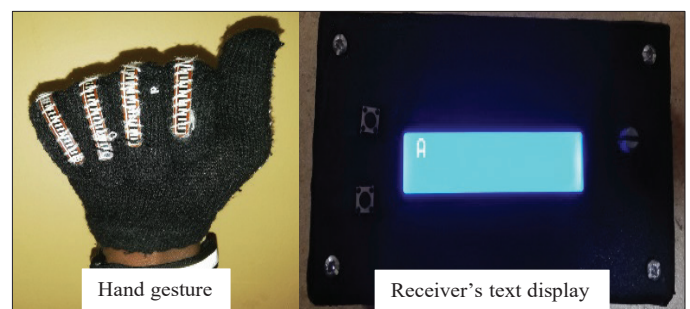


Fig. 5. Signing letter "A" and its associated text on the display at the receiver.



Fig. 6. Signing the word "read" by fingerspelling



Fig. 7. Corresponding LCD text display for fingerspelling the word “read”

F. Sensor calibration and test results

The flex sensors and IMU were calibrated using the Arduino serial monitor. Some of the calibration results for letters “A” and “B” are shown in Figs. 8 and 9. The labels FLX1-FLX5 are values for flex sensor readings measured in degrees, while ax, ay, and az are accelerometer readings. The readings shown as gx, gy, and gz are gyroscope values measured in degrees per

second. The x-sets represent the pitch of the hand, the y-sets represent the roll, and the z-sets represent the yaw. Comparing Figs. 8 and 9, the gyroscope readings are almost the same, and owing to “A” and “B” being static gestures, there is no rotation of the hand. The accelerometer readings were mostly utilized for distinguishing static gestures.

Owing to the high sensitivity of the sensors to slight finger flexion and hand movements, the sensor values for each gesture are not constant despite maintaining the same handshape for each case. To improve the accuracy of distinguishing the signs, certain ranges were allowed for each set of sensor values while making sure that the handshape still represents the correct sign. Giving a range of acceptable values for each sensor reading helps to minimize fluctuation of readings that occurs with inconsistent bending of fingers and hand orientation for the same sign. This small deviation from the ideal handshape for each sign is such that the deviating handshapes still resemble the intended sign. Table 1 shows how the ranges were set for letters “A” and “B” in the SASL database.

COM3										
A										
FLX1	FLX2	FLX3	FLX4	FLX5	ax	ay	az	gx	gy	gz
36	42	54	31	115	6	265	324	90	79	99
36	40	55	33	114	5	268	318	95	80	100
36	43	54	30	113	3	265	325	93	80	99
35	42	56	35	113	2	267	324	94	80	99
37	42	56	43	113	5	270	328	100	79	99
37	44	56	41	115	4	269	329	96	80	98
36	43	56	38	116	5	269	320	99	80	99

Fig. 8. Arduino serial monitor display for letter “A” for the AudibleSigns Sign Language converter.

COM3										
B										
FLX1	FLX2	FLX3	FLX4	FLX5	ax	ay	az	gx	gy	gz
112	98	126	100	93	1	252	321	99	79	100
112	98	126	99	94	0	255	318	96	79	100
111	98	126	100	94	0	254	324	97	80	99
113	97	126	101	93	1	254	323	98	79	99
113	98	126	101	92	2	253	323	99	80	100
112	97	128	99	94	3	256	324	98	80	99

Fig. 9. Arduino serial monitor display for letter “B” for the AudibleSigns Sign Language converter.

TABLE I
DATABASE FOR LETTERS “A” AND “B” WITH RANGES SHOWN IN DEGREES

Letter	FLX1(°) Little finger	FLX2(°) Ring finger	FLX3(°) Middle finger	FLX4(°) Index finger	FLX5(°) Thumb	ax(°/s) Pitch	ay(°/s) roll
A	30-50	35-55	45-65	29-50	110-125	ax≤10	ay≥245
B	100-120	85-105	120-135	95-105	85-100	ax≤10	ay≥245

At least 5 trials for each letter were undertaken for the slight variations in handshape from the ideal handshape. When performing the trials, the position of the hand in front of the signer (determined by the elbow flexion and extension), as well as the extension and flexion of the wrist and fingers were considered for each letter. The corresponding sensor values were captured, as shown in Fig. 9. All the values for each sensor in all the trials were compared and their minimum and maximum values were then determined. This method was repeated for all the letters, and this was how the database for each letter was created (Table 1 shows the ranges for letters A and B). The ranges shown in Table 1 were chosen after trying out possible and acceptable deviations of each handshape from the ideal handshape for each sign and capturing the values for all the trials. These ranges of values were also chosen in such a way that they would improve the accuracy of distinguishing one sign from the others. For some letters a trade-off was made; if the range was set too wide, the categorization of the correct letter may overlap into another letter.

G. Signing accuracy

Fig. 10 shows the prototype's accuracy when signing all of the 26 letters of the manual alphabet. To determine the accuracy, 10 repetitions were completed for each letter, and a percentage accuracy was ascertained. Of the 26 letters, 18 had an accuracy of at least 70%; eight had an accuracy of 40% or lower. Of the 18 letters with 70% accuracy or better, the average accuracy was 86.8%. The average accuracy of all 26 letters was 69.6%. It was easier to sign words that had letters with an accuracy higher than 40%. These words are those with signs that are not found in the SASL dictionary and require fingerspelling of each constituent letter. The word "read" was tested as an example. Furthermore, the phrase "thank you" with specific dynamic signs in the SASL dictionary was tested and had 100% accuracy.

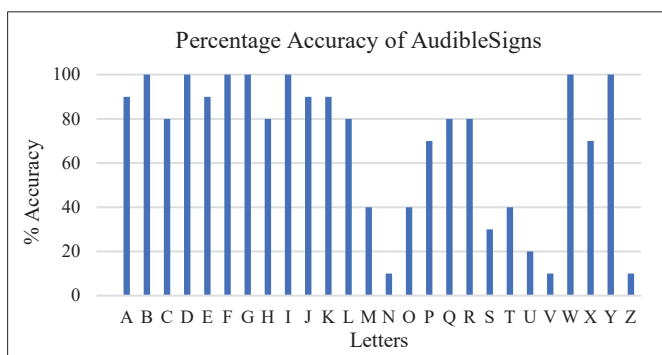


Fig. 10. Accuracy of AudibleSigns for each letter.

H. Costing

The breakdown of costs for the prototype is shown in Table II. The total cost is under R2500 (<\$150).

Description	Cost (ZAR)
Flex sensor × 5	900
Electronic components and enclosure	521
Arduino Nano × 2	278
DfPlayer	118
Liquid Crystal Display	115
LiPo charger + boost converter × 2	110
LiPo battery × 2	90
MPU6050 (IMU)	89
Protoboards × 4	65
Speaker	65
Glove	30
Total	R 2381
	(\$144)

IV. LIMITATIONS

The prototype was designed for the following configurations:

- Converting 26 letters of the SASL manual alphabet.
- Converting common words and phrases that have specific signs in the SASL dictionary. For this prototype, "thank you" was used for demonstration purposes.
- Converting some words that do not have specific signs in the SASL dictionary, such as people's names.

The prototype was not designed to convert the entire SASL dictionary, which has over 1000 signs. Due to only one glove being used in this prototype, the prototype cannot capture all aspects of the SASL phonology. The prototype was specifically designed to convert the manual alphabet as used in fingerspelling as fingerspelling is an important part of SL. Thus, the focus for this prototype was on words and phrases that non-signers would not readily understand.

Due to the constraints caused by the flex sensors, there are some accuracy limitations in distinguishing similar signs. These are for signs of A, E, M, N, O, S, T, U, and V. While the use of flex sensors did manage to solve the problem, flex sensors alone are not the best method for capturing handshape. Flex sensors measure finger flexion, not abduction and adduction. This means that they are not accurate at signing letters like "U" and "V" which are distinguished by finger abduction and adduction.

The flex sensors are susceptible to wear and tear owing to repeated extensions and flexions, resulting in reduced sensitivity of the sensors. Improvements in the flex sensor materials would need to be investigated to increase the lifespan.

V. DISCUSSION

Methods of reducing communication barriers between hearing and non-hearing people is an important field of research. Deaf people face additional daily challenges owing to communication barriers between hearing-impaired and non-hearing-impaired people, resulting in reduced employment opportunities and difficulties in accessing healthcare [21]. Deaf communities may be assisted with a SL to verbal device to reduce possible marginalization of Deaf communities owing to

a lack of interactions within hearing communities. The AudibleSigns prototype presented was designed to convert the manual signs of the 26 letters of the SASL manual alphabet. Thus, words that do not have specific signs in the SASL dictionary can also be converted to audible language. This prototype can also convert common phrases in the SASL dictionary. This prototype achieves an accuracy in distinguishing the manual alphabet signs using the available sensors despite their limitations, achieving an average success rate of 69.6%. Most letters achieved high conversion accuracy; however, there were a few letters that did not. The main reason behind the lower success rate is that the flex sensors used were 55.8 mm in length and were cheaper than the longer ones, measuring 115 mm. The shorter flex sensors used were not long enough to entirely cover the length of fingers; thus, there was a problem with the flex sensors not being bent enough at the distal interphalangeal joints. This made it difficult to distinguish between signs with similar handshapes. Longer flex sensors would achieve better accuracy than shorter ones.

There was a trade-off between accuracy and the deviation of handshapes during signing. Narrowing down the range of values for each flex sensor improves the signing accuracy but limits the deviations of handshapes. These deviations are important considerations because they result from the signer being inconsistent with their handshape every time they make the same sign. In practice, it is very difficult to maintain the exact handshape repeatedly. Therefore, an optimum trade-off was chosen during design to make sure that all the letters could at least be converted correctly despite the prototype achieving low accuracy for some letters. The results show that to a greater extent, the designed prototype of a single glove system is an affordable pathway for bridging the communication barrier between hearing and non-hearing people which can be used as a basis for future design improvements. The AudibleSigns converter offers portability owing to it being battery operated but there is still a need for improvements in gesture recognition of certain manual alphabet letters such as the letters N, S, U, V, Z. Improvements in the type of flex sensors used, as well as the addition of a second glove are possible methods for increasing the range of usability and accuracy for a SL to verbal converter. Flex sensors can be used together with contact sensors placed on fingertips and between fingers to measure the finger abduction and adduction which will improve the accuracy of the conversions. Thus, improvements to the range of the signs to verbal are achievable.

The focus for this prototype was on words and phrases that non-signers would not understand. This is particularly useful in healthcare settings whereby specific medical terms or descriptions need to be communicated to the healthcare practitioner. In such cases, the hearing-impaired individual could use such a device to fingerspell her symptoms to her healthcare worker.

VI. RELATED WORK

The work on AudibleSigns builds on ideas from previous prototypes that were made for other countries' SL. As SL is not

a universal language but rather country specific, AudibleSigns was developed in the South African context. Within the SASL context, work presented in [22] from the University of Cape Town focused only on developing a SASL dataset from a Bluetooth glove using machine learning algorithms. The work was then improved in [23] where a smart phone was used to convert and display only 24 static signs and 7 numerical digits from the Bluetooth glove. In comparison, AudibleSigns captures both static and dynamic gestures and goes beyond single letter conversion by demonstrating phrase or word construction through fingerspelling. In addition, with a smartphone, the user must hold the phone with the non-signing hand in a position that would show the phone's display to the person that the user is communicating with. This can be strenuous and uncomfortable for the user. The converter device that AudibleSigns uses can be hung from one's neck using a string/strap and can rest on the chest area making it more comfortable to use. This is also safer than holding a smartphone, which in South Africa could be a risky activity.

A single glove prototype was developed for Arabic SL called the *Sign-to-Speech/Text* [24]. This prototype uses flex sensors only, whereas the AudibleSigns' glove has flex sensors and an inertial measurement unit (IMU) for better sign capturing of the SASL. It also uses an Arduino Nano and an Arduino Mega2560 compared to the two Arduino Nano microcontrollers used for AudibleSigns. For the audio processor, the Emic 2 text to speech module was used in this Arabic to SL converter prototype, while a mini MP3 player was used for the AudibleSigns.

A group of researchers at Notre Dame University developed their *S2L System*, which is a Sign to Letter translator system [25]. This affordable system uses a five-flex sensor-based glove and has an LCD to display the corresponding signed letters. The system has an overall signing accuracy of 94%; however, the crucial "J" and "Z" signs have a lower accuracy of 85%. This lower accuracy is due to the dynamic nature of these signs, which flex sensors cannot yet accurately detect. Owing to the S2L system lacking an audio output, the researchers recommended a future improvement that replaces the LCD with a mobile phone to incorporate verbal speech into their system. The AudibleSigns prototype has an audible output.

Scholars from the American International University in Bangladesh developed the *Speak Up* prototype for speech-impaired people [3]. The system translates signs made by a flex sensor-based glove into both text and speech using an LCD and a speaker. The system does not make any reference to any SL; thus, the user makes four random gestures (defined by the designer) that translate into audible and visual phrases.

Scholars from the University of Washington in the United States won the Lemelson-MIT Student Prize for their work on the *SignAloud* project [26]. This is a system of a pair of gloves that recognizes American SL signs into words and phrases. It uses Bluetooth communication between the gloves and a computer that will produce the corresponding phrases via the speaker.

The success of having devices that converts SL to verbal being marketable relies on how much they replicate the art of SL. For instance, some devices have been developed where any random signs (with no reference to SL) were made to give the verbal output of words and phrases that already have specific signs in the SL dictionary. Such devices were criticized for not representing SL as a Natural Language, implying that they could not be marketed.

VII. FUTURE STUDY

The overall accuracy of the single glove prototype can be improved by adopting more sensitive approaches to sensor movement tracking. An example includes metallic nanomeshes that make soft conformal electrical interfaces with the skin. This approach has been used in recent wearable electronics as highly precise sensing platforms [27]; however, their fabrication costs are high. New technologies such as microfibre sensors have emerged in gesture recognition for data glove applications. They are a potential replacement for flex sensors as they have higher sensitivity, a faster response, and a smaller footprint than flex sensors. This technique was developed for the *InfinityGlove* for gaming applications [28]. Future research could investigate this technology for SL translation.

To improve the portability of a SL converting system, improvements can focus on adding the functionalities of these systems to our smart devices. Some of the discrete components of the converting device can be miniaturized and fabricated on nanochips to reduce the number of standalone components, which in turn reduces the overall footprint. These nanochips can then be integrated into smart watches. In addition, Bluetooth connectivity can be used to link the glove to a mobile phone. The mobile phone would then provide both the audio and visual outputs.

If these SL devices were to be marketed as a complete solution, they should address most aspects of SL, including sentence construction and non-manual features like facial expressions. The devices should be accurate when signing both static and dynamic gestures. Since not all SLs across the world are the same, there is a requirement for the SL converter to be specifically designed for the country of use. Overcoming these challenges underscores the need for more research in trying to expand the reliability and accuracy of these devices. The application of artificial neural networks in gesture recognition also promises to be a viable solution [29].

Lastly, it would be best to design systems that not only convert SL to verbal but also convert speech to gestures at the same time to have two-way communication between the Deaf and non-Deaf people.

VIII. REFERENCES

- [1] Umalusi. *Sign of the Times The Quality Assurance of the Teaching and Assessment of South African Sign Language*. Pretoria, SA. Umalusi, Council for Quality Assurance in General and Further Education and training. 2018
- [2] H. Liu and L. Wang, "Gesture recognition for human-robot collaboration: A review," *International Journal of Industrial Ergonomics*, vol. 68, pp. 355–367, Nov. 2018, doi: 10.1016/j.ergon.2017.02.004.
- [3] S. Ahmed, R. Islam, M. S. R. Zishan, M. R. Hasan and M. N. Islam, "Electronic speaking system for speech impaired people: Speak up," *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, 2015, pp. 1-4, doi: 10.1109/ICEEICT.2015.7307401.
- [4] E. Holmer, M. Heimann, and M. Rudner, "Evidence of an association between sign language phonological awareness and word reading in deaf and hard-of-hearing children," *Research in Developmental Disabilities*, vol. 48, pp. 145–159, Jan. 2016, doi: 10.1016/j.ridd.2015.10.008.
- [5] A. Dexter. "How Many Words are in the English Language?" *wordcounter.io*. <https://wordcounter.io/blog/how-many-words-are-in-the-english-language> (accessed Jan. 02, 2023).
- [6] M. A. Ahmed, B. B. Zaidan, A. A. Zaidan, M. M. Salih, and M. M. bin Lakulu, "A Review on Systems-Based Sensory Gloves for Sign Language Recognition State of the Art between 2007 and 2017," *Sensors*, vol. 18, no. 7, p. 2208, Jul. 2018, doi: 10.3390/s18072208.
- [7] Google. "Soli" *atap.google.com*. <https://atap.google.com/soli> (accessed Jan. 02, 2023).
- [8] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "SignFi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, pp. 1–21, Mar. 2018, doi: 10.1145/3191755.
- [9] J. G. Abreu, J. M. Teixeira, L. S. Figueiredo and V. Teichrieb, "Evaluating Sign Language Recognition Using the Myo Armband," *2016 XVIII Symposium on Virtual and Augmented Reality (SVR)*, 2016, pp. 64-70, doi: 10.1109/SVR.2016.21.
- [10] S. S. Rautaray and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," *Artificial Intelligence Review*, vol. 43, no. 1, pp. 1–54, Nov. 2012, doi: 10.1007/s10462-012-9356-9.
- [11] H. Sarbolandi, M. Plack, and A. Kolb, "Pulse Based Time-of-Flight Range Sensing," *Sensors*, vol. 18, no. 6, p. 1679, May 2018, doi: 10.3390/s18061679.
- [12] N. B. Ibrahim, M. M. Selim, and H. H. Zayed, "An Automatic Arabic Sign Language Recognition System (ArSLRS)," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 470–477, Oct. 2018, doi: 10.1016/j.jksuci.2017.09.007.
- [13] R. Lun and W. Zhao, "A Survey of Applications and Human Motion Recognition with Microsoft Kinect," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 05, p. 1555008, Jul. 2015, doi: 10.1142/s0218001415550083.
- [14] M. Eltoukhy, A. Kelly, C.-Y. Kim, H.-P. Jun, R. Campbell, and C. Kuenze, "Validation of the Microsoft Kinect® camera system for measurement of lower extremity jump landing and squatting kinematics," *Sports Biomechanics*, vol. 15, no. 1, pp. 89–102, Jan. 2016, doi: 10.1080/14763141.2015.1123766.
- [15] L. B. Neto *et al.*, "A Kinect-Based Wearable Face Recognition System to Aid Visually Impaired Users," *IEEE Transactions on Human-Machine Systems*, pp. 1–13, 2016, doi: 10.1109/thms.2016.2604367.
- [16] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor," *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1110–1120, Aug. 2013, doi: 10.1109/tmm.2013.2246148.
- [17] D. Patil, S. Sahane, V. Soma and A. Francis. Interactive Glove. *International Journal of Industrial Electronics and Electrical Engineering (IJIEEE)*, 3(11), pp.68-70. 2015
- [18] P. Havalagi and S. Nivedita. The amazing digital gloves that give voice to the voiceless. *International Journal of Advances in Engineering & Technology*, 6(1), p.471. 2013
- [19] G. Saggio and G. Orengo, "Flex sensor characterization against shape and curvature changes," *Sensors and Actuators A: Physical*, vol. 273, pp. 221–231, Apr. 2018, doi: 10.1016/j.sna.2018.02.035.
- [20] S. Manton, G. Magerowski, L. Patriarca, and M. Alonso-Alonso, "The 'Smart Dining Table': Automatic Behavioral Tracking of a Meal with a Multi-Touch-Computer," *Frontiers in Psychology*, vol. 7, Feb. 2016, doi: 10.3389/fpsyg.2016.00142.
- [21] A. Kuenburg, P. Fellingner, and J. Fellingner, "Health Care Access Among Deaf People," *Journal of Deaf Studies and Deaf Education*, vol. 21, no. 1, pp. 1–10, Sep. 2015, doi: 10.1093/deafed/env042.
- [22] B. McInnes, "South African sign language dataset development and translation : a glove-based approach," *open.uct.ac.za*, 2014, Accessed:

- Jan. 09, 2023. [Online]. Available: <https://open.uct.ac.za/handle/11427/13310>
- [23] M. Seymour and M. Tšoeu, "A mobile application for South African Sign Language (SASL) recognition," *AFRICON* 2015, 2015, pp. 1-5, doi: 10.1109/AFRCON.2015.7331951.
- [24] D. Abdulla, S. Abdulla, R. Manaf and A. H. Jarndal, "Design and implementation of a sign-to-speech/text system for deaf and dumb people," 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA), 2016, pp. 1-4, doi: 10.1109/ICEDSA.2016.7818467.
- [25] H. El Hayek, J. Nacouzi, A. Kassem, M. Hamad and S. El-Murr, "Sign to letter translator system using a hand glove," *The Third International Conference on e-Technologies and Networks for Development (ICeND2014)*, 2014, pp. 146-150, doi: 10.1109/ICeND.2014.6991369.
- [26] anon. "UW undergraduate team wins \$10,000 Lemelson-MIT Student Prize for gloves that translate sign language," *UW News*. <https://www.washington.edu/news/2016/04/12/uw-undergraduate-team-wins-10000-lemelson-mit-student-prize-for-gloves-that-translate-sign-language> (accessed Jan. 02, 2023).
- [27] J. A. Rogers, "Nanomesh on-skin electronics," *Nature Nanotechnology*, vol. 12, no. 9, pp. 839–840, Sep. 2017, doi: 10.1038/nnano.2017.150.
- [28] N. U. of Singapore, "Researchers develop smart gaming glove that puts control in your hands," *techxplore.com*. <https://techxplore.com/news/2020-08-smart-gaming-glove.html> (accessed Jan. 02, 2023)
- [29] F. Solís, D. Martínez, and O. Espinoza, "Automatic Mexican Sign Language Recognition Using Normalized Moments and Artificial Neural Networks," *Engineering*, vol. 08, no. 10, pp. 733–740, 2016, doi: 10.4236/eng.2016.810066.



Shingirirai Chakoma (M' 22), received his Bachelor of Engineering Technology and Bachelor of Engineering Technology Honors degrees in electrical engineering from the University of Johannesburg, South Africa in 2020 and 2021, respectively. He is a Ph.D student in the electrical engineering and computer science department at the University of California, Irvine, Unites States, where he is currently working in the Integrated Nano Bioelectronics Innovation Lab. His research interests include wearable devices for health monitoring, microfabrication, biosensors, radio frequency and electronic circuit design. In 2019, he became a member of the Golden Key International Society. He is also a recipient of the 2021 Henry Samueli Endowed Fellowship awarded by the Henry Samueli School of Engineering at UC, Irvine.



Philip Baron has been published across several academic disciplines and has an active social-media presence with over 85,000 subscribers and over 30 million views on his educational YouTube channel. Philip has postgraduate degrees in electrical engineering, psychology, and religious studies.