

Representing American Sign Language Letters and Numbers with Humanoid Robot Arm

Mehmet GÜL^{1,*} 

¹ Department of Computer Engineering, Şırnak University, Şırnak, 73000, Turkey, **ORCID:** 0000-0002-4819-4743

Article Info

Research paper

Received : February 17, 2021

Accepted : July 07, 2021

Keywords

Humanoid Robotic Arm
Sign Language
Deep Learning
Human-Robot Interaction
Speech-Recognition

Abstract

The widespread use of robotic systems is an example of increasing technological developments in daily life. The robotic systems that are accustomed to widespread use in the industry show huge impact in education, health, and many other areas nowadays. Robotic systems are no longer expensive and are not difficult to access systems that can be drawn and designed with design software like CAD programs. Within the scope of the study, major changes are made in the reference design, and Python software is used to send data to the Arduino control card that is used for control of the robotic system and to provide movement. The humanoid robotic arm developed in the study is printed with the 3D printer, sign languages that can be realized by one-handed letters among sign languages were investigated. All 26 alphabetical letters of English in American Sign Language can be demonstrated with the right hand alone. It has nearly the same principle that represents the demonstration of sign language expressions with a robotic hand, and the American Sign Language is the most widely known sign language. If experience is gained with this study, it can be used for the other sign language. In this study, audio and text messages sent over Python were performed on the humanoid robotic arm.

1. Introduction

Sign language is considered the language that helps deaf people communicate with the outside world by converting the letters, numbers, words, and sentences of a natural language into hand and head gestures. Hearing-impaired individuals use hand gestures and facial expressions instead of sound to express themselves. Sign language varies from country to country and from region to region. The Ministry of Family Labor and Social Services of the Republic of Turkey published a statistics bulletin in 2020. There is a statistical table in that bulletin and according to the disabled people table, there were about 1.5 million hearing and speech impaired people in 2011 (Table 1) [1].

Hearing and speech loss bring about significant communication problems for individuals in daily life. Most people with this condition use sign language to

communicate with normal individuals. On the other hand, this situation is not understood by many individuals and is limited to individuals with hearing or speech impairment. Hearing-impaired individuals deserve all the help they need to continue their daily lives like normal people. One way to tackle this situation is to use advanced technology to overcome some of the challenges it faces. For example, high-tech machines that perform sign language, which is a combination of hand and head movements and facial gestures, and act as mediators between them and normal people, such machines are an example of human-machine interaction.

Today, humanoid robots are used in education [2], therapy [3], and entertainment. These developments are provided by encouraging experiments on humanoid robots. Studies on the use of robots as friends in the education and treatment processes of children with disabilities and special needs can be given as an example [4]. Teaching sign language to hearing-impaired children with the help of a humanoid robot can be given as an example.

Learning sign language is the most basic communication tool for people with hearing and speech

* Corresponding Author: mgul@simak.edu.tr



impairments to communicate with each other or other normal individuals. It has also been observed that sign language contributes significantly to the cognitive and intellectual development of individuals with hearing and speech impairments [5]. The special bond that a child will create with the humanoid robot can be an ideal tool for the use of the developed robot for both educational and therapeutic purposes for the child with whom it is in communication. Accordingly, it has been proposed to use humanoid robots as a game-based learning method in

education. It has been found that the quality of education and the willingness to play and interact with a child's autonomy or disability can be increased [6]. In some studies, robotic systems have been used in education, and it has been noted that they can be more effective than computer-based teaching systems [7, 8]. Kinematic limitations of the NAO and Robovie 3 robotic systems used in these studies, which greatly limit the number of signals, have been determined.

Table 1 Population of disabled individuals across the country according to TURKSTAT 2011 [1].

Disability group	Total Population Rate (%)	Male (%)	Female (%)	Total (Number of People)	Male (Number of People)	Female (Number of People)
Those who have difficulty seeing	1,4	1,3	1,5	1.039.000	478.000	561.000
Those who have difficulty in hearing	1,1	1,1	1,2	836.000	406.000	429.000
Those who have difficulty speaking	0,7	0,8	0,6	507.000	278.000	229.000
Those who have difficulty walking, climbing / descending stairs	3,3	2,4	4,1	2.313.000	861.000	1.452.000
Those who have difficulty carrying / holding things	4,1	3,2	5,1	2.923.000	1.136.000	1.787.000
Those who have difficulty in learning / doing four simple operations / remembering attention according to their peers	2,0	1,6	2,4	1.412.000	565.000	847.000

Today, only a few ingenious anthropomorphic hands have been developed for humanoid robots. Examples of these robots are 13 degrees of freedom (DoF) fifty and Honda's ASIMO Humanoid1. On the other hand, the open-source iCup humanoid platform with 11 DoF [9], the ARMAR III upper body mobile with 8 DoF [10], and NASA's 11 DoF Robo-naut-2 humanoid robots can be cited as examples [11].

The mechanical and software stages of the developed humanoid robot arm in this study are introduced. The letters used in American Sign Language were realized with the humanoid robot arm developed in the study. Compared to the letters made in Turkish Sign Language (TSL) and also British Sign Language, the letters used in American Sign Language (ASL) can only be performed with one hand such as French Sign Language or Irish Sign Language.

2. Materials and Methods

The use of humanoid robotic technology in education has been successful, and the results of given examples are extracted from the studies [12, 13]. Letters of some sign languages can be performed with just one hand such as ASL, in the study, the developed robot arm has performed the letters and 0–9 numbers of ASL. There are seven servomotors used to control the robotic arm, two servos for the thumb, and the other four servos for other fingers, and the last one for the wrist.

Sign language, the realization of signs, requires the analysis of structures, especially phonological ones. Phonology in sign language expresses how a sign is created and arranged. Examining the ASL points out that each sign in the ASL can be divided into five components: hand shape, hand shape movement, the position of the mark, palm direction, head direction, and facial expressions, etc. [14]. Research has shown that these components, which can be applied to TSL, together form the meaning of a sign.

2.1. Electronic Architecture

The humanoid robotic arm is designed taking into account other design guidelines, such as low development cost, as well as its ability to mimic ASL markings. Palms and fingers are used in the study referenced on the web page of Thingiverse open-source sharing [15]. Other elements of the referenced design are not used, but redesigned due to insufficient actuators. In the reference design, a total of five actuator slots are opened and used for the movement of the fingers. On the other hand, for the movement of the fingers on the palm, a total of six actuators must be used. The reason for this is the unique movement of the thumb. The metacarpal bone in the thumb gives the thumb a different movement feature than other fingers. For this reason, two actuators need to be used to control the thumb (Figure 1).

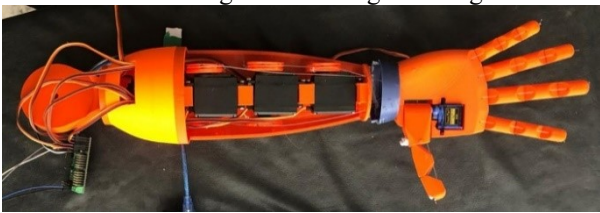
Table 2 The used electronic components

Servo Motor	Mini Servo Motor	Servo Control Card
MG996R	SG90	Mini Maestro Pololu 24 Channel Servo Control Board

The design referenced is shown in Figure 1.a. and Figure 1.b is the developed design. The components of the humanoid robot arm are developed using SolidWorks design software and printed on the 3D printer. PLA filament is used for printing the parts. 25 parts of a humanoid robot arm are printed by a 3D printer. In the study, a central computer and a control card providing movement are used to realize the letters and 0–9 numbers in ASL. An Arduino Uno is used to provide the movement. A MG996R servomotor is used for the movement of the fingers and an SG90 mini servomotor for the movement of the thumb from the palm. A Mini Maestro 24 channel servo motor control card is used to ensure the movement of servomotors. The movement is achieved by transferring the messages sent to the Arduino board via Python software to the servo motor control card. The software developed for providing movement is described in the next section.



a) Arm design consisting of 4 servo motors used for the movement of the fingers in the original design



b) redesigned robot arm shape

Figure 1. Robot arm printed with the 3D printer.

2.2. The Used Actuators and Other Mechanical Parts

It is necessary to use a servo horn to provide finger movements in the developed humanoid robot arm. The servo horn designs used in Thingiverse and other similar open-source web pages are examined, and it is decided to redesign the servo horn based on the experience gained from the applications. The servo horns used in many designs are shown in Figure 2. The humanoid robot developed by InMoov Company uses two different servo horns. In the first design show in Figure 2.a. the classical servo horns are used, and two fishing lines are used for the movement of the finger attached to the servo horn. The first design using servos used for finger movement can stretch the maximum amount of "r" with the servo horn. Figure 2.b shows the next design printed on a 3D printer.

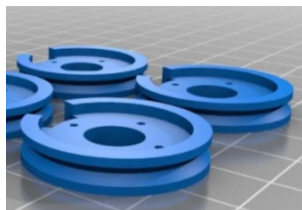
The biggest disadvantage of this design is that the line threads used for stretching and unloading are positioned and connected to the same horn. It has been observed that the line thread is wasted during the movement of the servo motor. It is observed that the InMoov design is used in the reference design. Due to the mentioned disadvantages, the servo horn has been redesigned and shown in Figure 3.

With the developed servo horn, the line thread is wound at least two cycles. In this case, it is noticed that the line thread that is stretched sufficiently fell into nothing. And on the contrary, the line thread is wound more effectively when winding.

Seven actuators are used in the robot arm. Considering easy maintenance and installation design factors, ready-made modular servomotors are judged to be the ideal solution for actuator selection. The robot arm has 16 degrees of freedom. Designs are prepared with the SolidWorks design software used for design and converted to STL format for printing on a 3D printer. In the next step, it is converted into Gcode with Simplfy3D slicing software and printed by a 3D printer. Finger movements are provided with a 90° movement.



a) Servo horn used for finger movement of the InMoov humanoid robot arm [15]



b) Another servo horn developed for InMoov humanoid robot arm [16]



c) Servo horn used in the referenced humanoid robot arm [17]

Figure 2. Servo horns used in the humanoid robot developed by InMoov company and referenced robot arm

2.3. Finger Design, Flexion/Extension Mechanism

For the palm and finger mechanism, the open-source design on the Thingiverse web page has been used [15]. In particular, the wired finger mechanism used in the iCub robot developed by S. Hirose [18] is taken as an example in the study. Rubber is used to help the finger actuator and replace it when it is idle. The tensile force created by the tire meets the force required by the servomotors used in the movement of the fingers to restore the finger. The finger structure is shown in Figure 4.

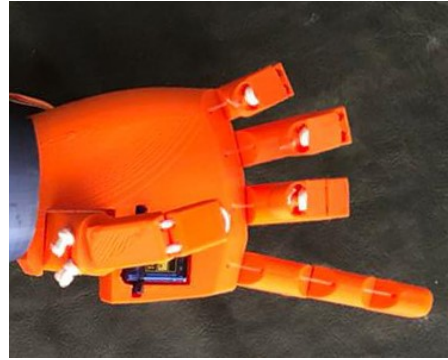
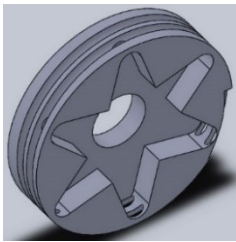


Figure 4. Fishing line and rubber used for the movement of the fingers



a. SolidWorks design



b. servo horn printed with the 3D printer

Figure 3. Developed servo horn

There is a DoF feature that allows the human wrist to move up and down. In the kinematic structure of the robot, the lateral movement of the wrist is not taken into account, and the tendon mechanism of the wrist movement and the cable routes are shown in Figure 4.

For finger movements, the angular positions of the servomotors are determined and positioned separately. One end of the line rope is knotted at the fingertip and the other end is knotted to Servo Horns attached to the servomotor. A rubber rope is used to restore the unilateral pulling procedure to its original position (see Figure 4). The tensile force accumulated in the rubber rope contributes to the transfer of the fingers to their initial positions.

3. Developed Software

Two software packages are used for the integrated software developed for the humanoid robot arm, Python and Arduino. Figure 5 shows the flow chart of the integrated software.

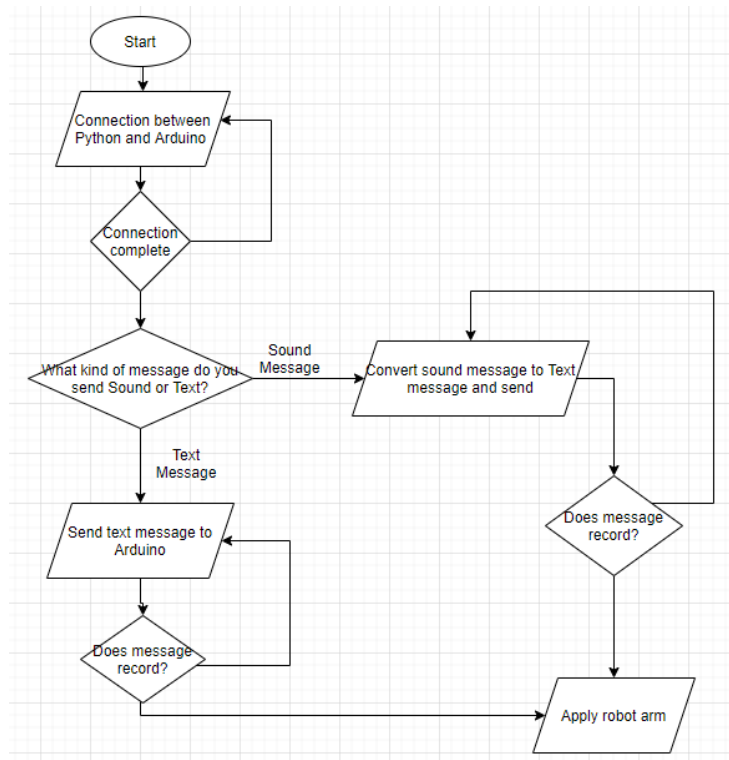


Figure 5. Flowchart of integrated software

The speech recognition (SR) system can be speaker-dependent or independent. The speaker-connected system is designed to be usable by a single speaker. In a speaker-independent system, it is designed to be used by any speaker and is naturally more difficult to obtain. These systems have a 3 to 5 higher error rate compared to speaker-connected systems [19]. The most common method used in SR systems is the probabilistic approach. A score is calculated for matching the expressed words to a speech signal. Any word in the vocabulary that is mentioned during the speech is compared to its probability value [20]. From the phonemes in the comparative acoustic model, which words can follow other words through grammar, are calculated. The word string with the highest score is selected as the recognition result. The SR process consists of four stages: preprocessing, feature extraction, decoding, and post-processing. The steps specified in different SR systems are implemented in different ways. In Figure 6, the four main stages are represented in detail.

3.1. Pre-Processing

The speech signal recorded through the microphone in the SR process is then extracted with a sampling frequency. According to the Shannon sampling theorem, if the sampling frequency is 16 kHz, the bandwidth limit can be formed to reconstruct a signal, meaning that frequencies up to 8 kHz can be correctly generated [21]. Since the data transmitted over the telephone network is in the range of 5–3.7 kHz, 8 kHz frequency generation is more than enough for speech recognition. In the speech recognition process, frequency values below 100 Hz are considered as noise and are eliminated by using a high pass filter. A significant part of preprocessing takes into account the sections between the start of the recording and the start of the user speech.

3.2. Feature Extraction

Acoustic observations made during the feature extraction process are extracted over uniform length time frames. The frame length is 25 ms for the signal that is considered constant between these frames. For acoustic observations in this window, the multidimensional feature vector is calculated and a fast Fourier transform is applied to the spectral field of the vector. The microphones on which speech will be recorded may differ. Therefore, linear filter effects are discarded with Cepstral Mean Subtraction (CMS) to compensate for the difference [22]. The information that can be found on the time change can be obtained by combining sequential feature vectors and creating a very high dimensional controller.

3.3. Decoding

The computational process in which word strings are most likely matched to feature vectors representing the acoustic signal is called decoding. In this process, there is a need for three important blocks consisting of an acoustic model containing a Hidden Markov Model (HMM) for each word or phoneme, a dictionary consisting of a word list and their phoneme strings, and a language model consisting of word probability or word strings [23]. The language model, which is not necessarily used in the decoding process, increases the success rate if it is used [24]. Usually fixed grammar or n-gram model consisting of words and probabilities; unigram, bigram, trigram model are used.

3.4. Post-Processing

In the last stage, the processed speech expression Viterbi is applied to the search and the best list of n is created. This list is re-scored using additional sources of information that improve the recognition accuracy of the highest-scoring result. There are more resources needed for computation and memory than unigram or bigram models. Often, the bigram model is combined with the trigram model in the recalculation.

Because of the four steps mentioned above, the processed audio is now ready to be converted into text. There are many different audio processing software packages available today that use similar processing processes. The success rate of sound processing software is inversely proportional to the minimum error rate achieved. The accuracy of human voices defined by sound libraries is measured by the Word Error Rate (WER) obtained because of sound processing [19].

$$WER = \frac{\text{Number of Substitutions+Insertions+Deletions}}{\text{Total number of words}} \quad (1)$$

In a study conducted in 2019, Microsoft Research stated that 5.1% WER was obtained in the NIST 2000 Switchboard task, 6.6% WER was obtained in the lower Switchboard cluster in the IBM Research Hub5 2000 evaluation test set, and finally, Google Research stated that 4.1% WER was obtained [26]. In another study conducted in 2017, Google Speech Recognition (GSR) and Siri, developed by Apple, were compared. In that study, 8-word samples taken from 41 people were processed through both voice recognition software, and it was determined that GSR gave much more successful results than Siri [27].

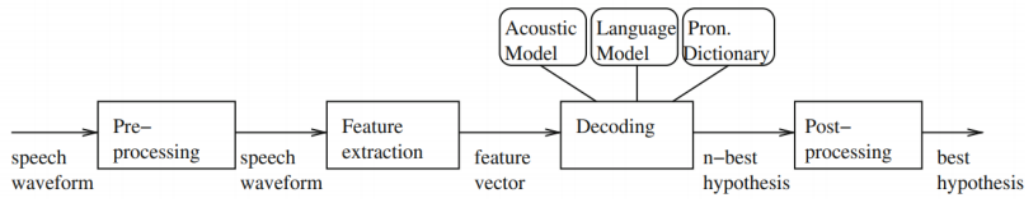


Figure 6. Speech Recognition process [25]

In this study, the speech recognition processes are examined, and found that GSR voice processing is successful compared to other similar softwares. In the Python software used in the prepared integrated software, GSR is called in the speech recognition library used for speech recognition, and speech expressions are processed over the GSR.

The Arduino Uno card and Mini Maestro 24 channel servomotor control card are controlled by the integrated Python and Arduino software. First of all, the user is asked to send a text message or a voice message with the right of choice. The GSR sound library is used in voice messages, and spoken messages are converted into text. In both cases, the messages obtained are sent with PySerial to the motherboard, where the Arduino software is processed via Python. Figure 6 shows the letters ASL and the movement of the humanoid robot arm. In the study, 26 letters and numbers (0–9) in English are carried out with a humanoid robot arm (table 3).

4. Results and Discussion

This work focused on mechatronics and electronic design guidelines and integrated software of the humanoid robot arm. First, sign language letters that can be executed with one hand are investigated and ASL is chosen because of its applicability in this respect. The robot arm with integrated software realized all letters and numbers in ASL in Table 2.

The sound data obtained with Python or the entered

written letters and numbers were sent to the Arduino control card with PySerial, and the comparison is made against the movement recorded on the Arduino and performed on the robot arm. There are eight bones in the human wrist to provide rotational movement. The existing eight bones are located between the forearm and comb bones and are in a structure that is connected by tight ligaments. Rotational motion of the human wrist around its axis does not seem possible in many humanoid robots today. For this reason, the representation of some letters in ASL is created in the closest form. Although there are some limitations in the robotic arm, both the printing on 3D printers and the use of open-source software in the development of the software is shown to be the prominent features of this study compared to the existing studies. The letters and numbers mentioned below with a humanoid robot arm have shown that robotic systems can be used successfully in imitating sign language.

The main features of the developed humanoid robot arm are the ability to use sign language, low development cost, and easy maintenance. The next stage of the study aims to develop a humanoid robot to teach sign language to hearing and speech-impaired children. Other units of the humanoid robot that are hoped to be developed, such as the upper arm, body, and head, are still being researched. The humanoid mobile robot to be developed will have features that can imitate the TSL. It is assumed that this work will contribute to robotic systems that can be printed with 3D printers and shared as open source, and that the work done in this field will continue.

Table 3. Letters and numbers are used in American Sign Language [28]

a	A	b	B	c	C

Table 3. (Cont.) Letters and numbers are used in American Sign Language [28]




































































 d	 D	 e	 E	 f	 F
 g	 G	 h	 H	 i	 I
 j	 J*	 J*		 k	 K
 l	 L	 m	 M	 n	 N
 o	 O	 p	 P	 q	 Q
 r	 R	 s	 S	 t	 T

Table 3. (Cont.) Letters and numbers are used in American Sign Language [28]

 U	 U	 V	 V	 W	 W
 X	 X	 Y	 Y	 Z	 Z
 0	 0	 1	 1	 2	 2
 3	 3	 4	 4	 5	 5
 6	 6	 7	 7	 8	 8
 9	 9				

* To symbolize the letter J in ASL, the little finger is first pointed upwards like the letter I, and after a short while the little finger closes towards the palm. This movement is raised as indicated, the little finger is lifted upwards and after 500 ms it is broken towards the palm.

Declaration of Ethical Standards

The authors of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] Republic of Turkey Ministry of Family Labor and Social Services (Web Page: <https://www.ailevecalisma.gov.tr/media/31492/istatistik-bulteni-kasim2019.pdf>) (Date Accessed: 01.05.2020)
- [2] Alemi M., Meghdari A., Ghazisaedy M., 2014. Employing humanoid robots for teaching English language in Iranian junior high-schools. *Int. J. Humanoid Rob.* **11**(3), pp.1–25. doi: 10.1142/S0219843614500224
- [3] Alemi M., Ghanbarzadeh A., Meghdari A., Moghadam L.J., 2015. Clinical application of a humanoid robot in pediatric cancer interventions. *Int. J. Soc. Rob.*, pp. 1–17. doi:10.1007/ s12369-015-0294-y
- [4] Meghdari A., Alemi M., Taheri A., 2013. The effects of using humanoid robots for treatment of individuals with autism in Iran. In: 6th Neuropsychology Symposium, Tehran, Iran.
- [5] Mayberry R.I., 2002. Cognitive development in deaf children: the interface of language and perception in neuropsychology. In: *Handbook of Neuropsychology*, **8**(Part II), pp. 71–107
- [6] Besio S., et al., 2007. Critical factors involved in using interactive robots for play activities of children with disabilities. In: *Proceedings of AAATE 2007 on Challenges for Assistive Technology*, pp. 505–509
- [7] Janssen J.B., van der Wal C.C., Neerinx M.A., Looije R., 2011. Motivating children to learn arithmetic with an adaptive robot game. In: Mutlu, B., Bartneck, C., Ham, J., Evers, V., Kanda, T. (eds.) *ICSR 2011. LNCS*, vol. 7072, pp. 153–162. Springer, Heidelberg
- [8] Nalin M., et al., 2012. Children's adaptation in multi-session interaction with a humanoid robot. In: 2012 IEEE on RO-MAN. IEEE
- [9] Schmitz A., et al., 2010. Design, realization and sensorization of the dexterous icub hand. In: 10th IEEE-RAS International Conference on Humanoid Robots (Humanoids). IEEE
- [10] Albers A., et al., 2006, Upper body of a new humanoid robot-the design of ARMAR III. In: 2006 6th IEEE-RAS International Conference on Humanoid Robots. IEEE
- [11] Diftler M.A., et al., 2011. Robonaut 2-the first humanoid robot in space. In: 2011 IEEE International Conference on Robotics and Automation (ICRA). IEEE
- [12] Kose H., Akalin N., Uluer P., 2014. Socially interactive robotic platforms as sign language tutors. *International Journal of Humanoid Robotics*, **11**(01), 1450003.
- [13] Köse H., Uluer P., Akalin N., Yorgancı R., Özkul A., Ince G., 2015. The effect of embodiment in sign language tutoring with assistive humanoid robots. *International Journal of Social Robotics*, **7**(4), pp.537-548.
- [14] Valli C., Lucas C., 2000. *Linguistics of American Sign Language: An Introduction*. Gallaudet University Press, Washington DC
- [15] Open-Source Thingiverse Design. <https://www.thingiverse.com/thing:2269115> (Access Time: 01/05/2020)
- [16] Design of Humanoid Robot prepared by InMoov company. <https://www.thingiverse.com/thing:17773> (Access Time: 01/05/2020)
- [17] Design of Humanoid Robot prepared by InMoov company. <https://www.thingiverse.com/thing:65274> (Access Time: 01/05/2020)
- [18] Hirose S., Umetani Y., 1978. The development of soft gripper for the versatile robot hand. *Mech. Mach. Theory*, **13**(3), pp. 351–359
- [19] Stenman M., 2015. *Automatic speech recognition An evaluation of Google Speech*, Corpus ID: 13963607, Published
- [20] Gruhn R.E, Minker W, Nakamura S, 2011. *Statistical pronunciation modeling for non-native speech processing*. Springer Science & Business Media
- [21] Jerri A.J., 1977. *The shannon sampling theoremits various extensions and applications: A tutorial*

review. Proceedings of the IEEE, **65**(11), pp. 1565–1596

- [22] Westphal M., 1997. The use of cepstral means in conversational speech recognition. In: EUROSPEECH
- [23] Rainer E. Gruhn W.M., Nakamura S., 2011. Statistical pronunciation modeling for non-native speech processing. In: Springer Science and Business Media
- [24] Tanigaki K., Yamamoto, H., Sagisaka, Y., 2000. A hierarchical language model incorporating class-dependent word models for oov words recognition. In: INTERSPEECH
- [25] Soudamalla S.K., “Implications of Conversational AI on Humanoid Robots”, site: <https://nbn-resolving.org/urn:nbn:de:bsz:ch1-qucosa2-724268>
- [26] Kim J.Y., Liu C., Calvo R.A., McCabe K., Taylor S.C.R., Schuller B.W., Wu K., 2019. A Comparison of Online Automatic Speech Recognition Systems and the Nonverbal Responses to Unintelligible Speech. arXiv.org > cs > arXiv:1904.12403
- [27] Daniels P., Iwago K., 2017. The suitability of cloudbased speech recognition engines for language learning. JALT CALL Journal, **13**(3), pp.229-239
- [28] American Sign Language. https://en.wikipedia.org/wiki/American_manual_alphabet (Access time: 01/12/2020)