Development of a Support System for Voice Outputs and Character Recordings in Communication via the "Mouth-shape Character" Method

Tomoya Miyasaka, Norio Kato, Mamiko Kamoshida, Keiko Kawashima, Toshiaki Tanaka

Abstract— **The "mouth-shape character" is one of the communication means of ALS (Amyotrophic lateral sclerosis) patients and caregivers. Although this means allows communication between the patient and the caregiver without using any tools, it is a burden on caregivers to utter speech sounds and record characters. The purpose of this study is to develop a support system for speech utterances and character recordings in communication via the mouth-shape character method. The prototype support system (proto-system) consisted mainly of an electric character board and a speech voice recognition module. Using the proto-system, the communication experiments were conducted by healthy subjects role-playing both the patient and the caregiver, and the system operability was evaluated. A control experiment was performed using the conventional mouth-shape character method and the results using these two methods were compared. Although the communication speed in the proto-system was slightly inferior to that of the mouth-shape character method, the accuracy of the confirmed character recordings by the former method was high, indicating the possibility of reducing the burden on caregivers. In addition, since speech recognition adapted to the caregiver's voice can be used as an input trigger, both hands of the caregiver remain free, and the potential for a reduction in the caregiving workload was obtained.**

Index Terms— **Amyotrophic lateral sclerosis, communication support, Mouth-shape character, speech recognition**

I. INTRODUCTION

Neuromuscular disorders, including ALS, decrease the motor function of the whole body as the disorder progresses [1]. Onset of motor function decline is correlated with impaired verbal communication, indicating the need for support to enable patients to express their intentions [2]. There have been many distinct ways to support their communication, from advanced to low technologies [3]. Examples of advanced technology use include eye-tracking [4], cerebral blood flow [5], and brain waves [6]. Meanwhile, examples of low technology communication support [7]

Tomoya Miyasaka, Department of Physical Therapy, Faculty of Health Sciences, Hokkaido University of Science, Sapporo, Hokkaido, Japan.

Norio Kato, Department of Physical Therapy, Faculty of Health Sciences, Hokkaido University of Science, Sapporo, Hokkaido, Japan.

Mamiko Kamoshida, Department of Physical Therapy, Faculty of Health Sciences, Hokkaido University of Science, Sapporo, Hokkaido, Japan.

Keiko Kawashima, Department of Physical Therapy, Faculty of Health Sciences, Hokkaido University of Science, Sapporo, Hokkaido, Japan.

Toshiaki Tanaka, Department of Physical Therapy, Faculty of Health Sciences, Hokkaido University of Science, Sapporo, Hokkaido, Japan.

include the transparent character board [8] and the mouth-shape character [9]. We compared the characteristics of communication via the transparent character board method and the mouth-shape character method [10]. While the mouth-shape character method requires no tools and is capable of communicating in a simple manner, the burden on the caregiver is large, and it is the caregiver's task to utter the speech sound and record the "confirmed" characters. The purpose of this study is to develop a support system for speech utterances and character recordings in communication via the mouth-shape character method.

II. METHOD

A. Communication via the mouth-shape character method

The mouth-shape character method is used by caregiver to communicate with patients who have difficulty speaking. We showed two types of mouth-shape character method. Method (a) was initiated by the expression of patient's intention, and Method (b) was initiated by the utterance of caregiver's Hiragana characters. Assuming that Fig. 1 is a typical chart of Hiragana characters, the procedure for expressing "tsu" is illustrated in Table 1.

	(row) →									
	a	k	s	t	n	h	m	y	r	w
a	あ	か	さ	た	な	は	ま	や	ら	わ
	a	ka	sa	ta	na	ha	ma	ya	ra	wa
i	い	き	し	ち	に	ひ	み		り	
	i	ki	shi	chi	ni	hi	mi		ri	
u	う	く	す	つ	ぬ	ふ	む	ゆ	る	
	u	ku	su	tsu	nu	fu	mu	yu	ru	
e	え	け	せ	て	ね	へ	め		れ	を
	e	ke	se	te	ne	he	me		re	wo
o	お	こ	そ	と	の	ほ	も	よ	ろ	ん
	o	ko	so	to	no	ho	mo	yo	ro	n

Fig. 1 Chart of Hiragana characters

Method (a) is characterized by the ability to initiate communication from the patient. Since the row of vowels containing the patient's intended character can be determined first, the caregiver reads out the character column one step less than Method (b). The lesser the steps, the faster it takes to confirm a single character, and the caregiver can reduce the burden of reading out characters. Meanwhile, when the voluntary muscle force around the patient's mouth and face

decreases due to the progression of the disease, the patient will have difficulty in shaping the mouth indicating a vowel (Fig. 2), and at the same time, it will be difficult for the caregiver to confirm it. The feature of Method (b) is that even if the voluntary muscle force of the patient decreases, communication is possible if there is a muscle contraction which can be visually confirmed by the caregiver. However, the caregiver has one more step to read out in the character column than in Method (a), and it takes a long time to determine a single character. It also increases the burden on the caregiver to read out characters. For both Methods (a) and (b), it is necessary for the caregiver to record confirmed characters one by one.

Table 1. How to express "tsu" using the mouth-shape character method

---

Method (a); Initiating from the expression of the patient's intention
1. When the patient needs to express "tsu", he/she makes the mouth shape of the vowel "u" at the beginning of a row where the character "tsu" is present.
2. The caregiver confirms the patient's intention, and utters all characters in the row beginning with the vowel "u", expressed by the patient, as "u, ku, su, tsu..." from left to right.
3. When the patient hears the "tsu" uttered by the caregiver, he/she signals the caregiver by moving his/her eyelids and facial muscles.
4. The caregiver confirms the patient's intention, identifies the "tsu" and records it.

Method (b); Initiating from the caregiver's utterance of Hiragana characters
1. The caregiver utters the top character of each column, such as "a, ka, sa, ta..." from left to right in order.
2. When the patient hears the utterance of "ta" which is the first character in the row containing "tsu" intended by the patient, he/she signals the caregiver by moving his/her eyelids and facial muscles.
3. The caregiver confirms the patient's intention and utters all the characters in the " ta" column from top to bottom as "ta, chi, tsu...".
4. When the patient hears the "tsu" uttered by the caregiver, he/she signals the caregiver by moving his/her eyelids and facial muscles.
5. The caregiver confirms the patient's intention, identifies the "tsu," and records it.

Fig. 2 Mouth shapes for five Japanese vowels

B. Support methods for communication via the mouth-shape character method

A number of attempts have been made to use advanced technology to compensate for the shortcomings of the mouth-shape character method. These include the detection of images in the form of vowel mouth shapes expressed by the patient (using Method (a) in Table 1) [11], and the recording of characters expressed by the patient on a smartphone via a switch input circuit and wireless communication [12].

This study focused on Method (b) in Table 1 and examined the support for reading out and recording characters by caregivers, using voice output and speech recognition technologies [13] which are also applied in welfare engineering.

C. Support systems for the mouth-shape character method using speech recognition

The proto-system consisted of an electric character board (Let's Chat, Panasonic Co., Japan) (Fig. 3) [14], a speech recognition module (Voice Magic Standard, RayTron, Inc., Japan) (Fig. 4) [15], and Photo relay (TLP 596 A, Toshiba Co., Japan).

Fig. 3 Electric character board

Fig. 4 Speech recognition module

In the operation of the proto-system, when the caregiver uttered "confirm", the speech recognition module recognized his/her voice and generated a contact output signal. The electronic character board input the signal via Photo relay, and performed the definite operation (Fig. 5). The response time of speech recognition was 0.2 seconds, and the scanning speed of the electronic character board dial was 1 second/step. The operation process consisted of the following 8 steps, and as an example, the method of expressing Hiragana character

https://doi.org/10.31871/IJNTR.5.11.17

International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-5, Issue-11, November 2019 Pages 26-31
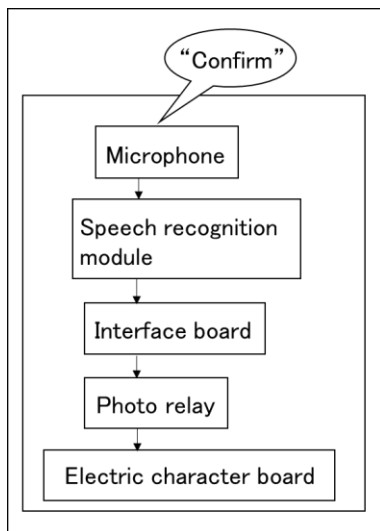
"tsu" is shown (Table 2).

Fig. 5 Operation process of prototype support system

Table 2. Procedure of expressing patient's intention via the prototype support system

1. The electronic character board is placed in a position not visible to the patient.
2. The caregiver utters "confirm". Then, the electronic character board scans one column at a time while uttering all five characters in all the columns from top down, such as "a, i, u, e, o", "ka, ki, ku, ke, ko", "sa, shi, su, se, so"...etc.
3. When the electronic character board utters the character column "ta, chi, tsu, te, to" including "tsu", the patient signals the caregiver by moving his/her facial muscles.
4. The caregiver confirms the patient's intention, and identifies the column by uttering "confirm".
5. The electronic character board scans the confirmed character column "ta, chi, tsu, te, to" while uttering one character at a time.
6. When the electronic character board utters the character "tsu", the patient signals the caregiver by moving his/her facial muscles.
7. The caregiver confirms the patient's intention, and identifies the character by uttering "confirm".
8. The electronic character board utters the confirmed character and records it on the device.

D. Evaluation of the prototype support system

The performance of this system was evaluated using communication experiments in which healthy subjects played the roles of the patient and the caregiver. A control experiment was performed using the conventional mouth-shape character method and these results were compared.

The experimental method was set based on a previous study [10]. Ten combinations of Hiragana characters, consisting of 3 characters that are not found in usual Japanese words, were created, and the order of appearance of the combinations was randomized. Subjects were 2 healthy persons over 20 years old. They were given the roles of the patient and the caregiver, facing each other in pairs and sitting in chairs at a distance of 1.2 m. One subject performed a trial of communication via the proto-system and the mouth-shape character method was used as a control. In this trail, a total of 30 Hiragana

characters consisting of 10 combinations times 3 characters were used (Fig.6). In the mouth-shape character method, the caregiver memorized the three characters that had been confirmed, and uttered the three characters that had been memorized at the end of a pair of tests. Furthermore, the patient and the caregiver were replaced, and one subject was given a 90 character trial per one communication method and two subjects were given a total of 180 characters. None of the subjects had any previous experience with these trials, and this study was conducted with prior explanation from the researchers. When the communication trial failed, the number of failures was counted and the trials were carried out again using the same Hiragana characters.

Fig. 6 Experimental scene

The communication process was recorded by a video camera placed on the lateral side of the subjects facing each other and an eye-movement measuring device (T.K.K. 2954, Takei Scientific Instruments Co., Ltd., Japan) (Fig. 6) was attached to the caregiver. The process number (N) was set for the intended characters on the character board because the steps to reach the column and the row are different for each character. For example, "ki" was positioned in a "ka" column (the 2nd column from left) and an "i" row (2nd row from top), and the process number was set to 4 as the sum of the number of columns and rows. A video file recorded from the video camera and the eye-movement measuring device was reproduced, and the process number of the character and the process time were totaled. In addition, the number of successes and failures in communication was recorded, and the subjects were asked for their impressions after the trials.

E. Research Ethics

This study was conducted based on the Research Ethics Application No.407 of Hokkaido University of Science after subjects had been informed in advance of the research and had given their consent. The experiment was carried out ensuring the safety of the subjects.

III. RESULTS

The process number was the same for the 2 types of trials, and was 8 ± 2.78 (N) with a minimum value of 2 (N) and a maximum value of 12 (N). The time of the proto-system was 7.07 ± 2.88 (s) with a minimum value of 1.50 (s) and a maximum value of 19.10 (s) (Fig. 7). With regard to the 2 points in Fig. 7 where the communication time was 15 seconds or more at the process number of 10 and 11,

respectively, it was shown that the communication was successful but the confirmation of the caregiver's expression was delayed, and the scanning was repeated. The time of the mouth-shape character method was 6.41 ± 2.40 (s) with a minimum value of 1.40 (s) and a maximum value of 12.3 (s) (Fig. 8). The number of communication failures via the proto-system was six, including four cases in which the patient failed to express his/her intention and two cases in which the caregiver failed to identify words. Also, The number of the failures via the mouth-shape character method was seven, including three cases of the patient's delay in expressing his/her intention and four cases of the caregiver's forgetting the confirmed characters.
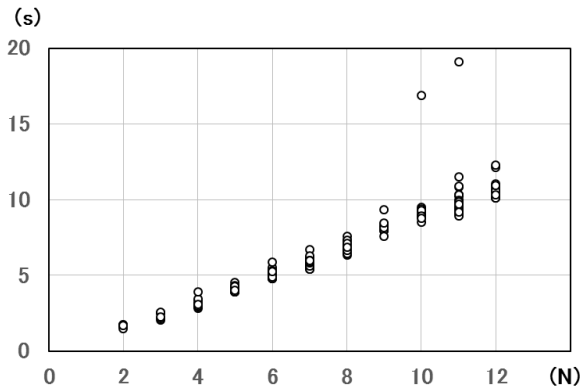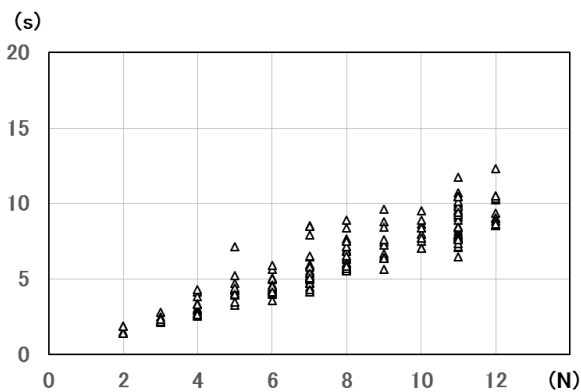

Fig. 7 Process number (N) and communication time (s) using the proto support system

Fig. 8 Process number (N) and communication time (s) using the mouth-shape character method
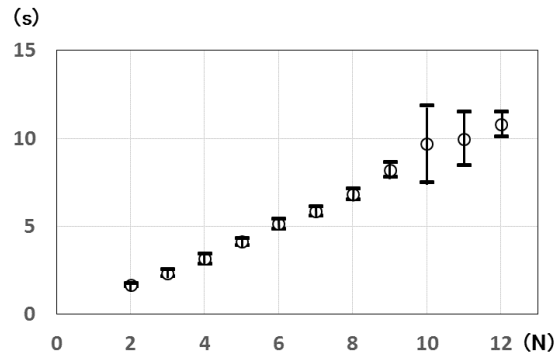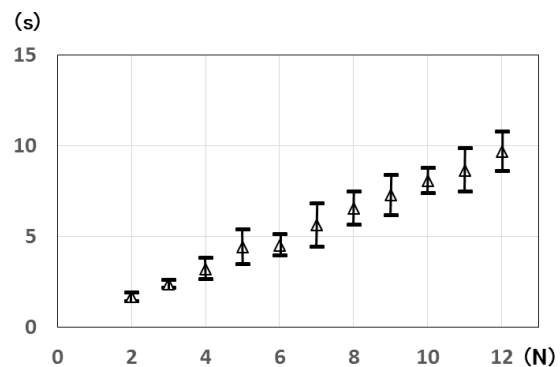
IV. DISCUSSIONS

A. Operation time

The inferential statistics showed no significant difference in communication time between the two methods. This suggests that the increase in communication time due to the increase in the process number was not large enough to make a significant difference. Therefore, regarding the 2 communication methods, the average values and standard deviations of communication time in each process number are shown in Table 3 and Fig. 9 and 10, the average time per process is shown in Fig. 11, and they were compared.

In the proto-system, the average communication time increased as the process number increased (Fig. 9).

Furthermore, the average time per process was about 0.9 seconds, and as the process number increased, the average time per process also tended to increase (Fig. 11). This can be explained by the fact that the scan time per process of the electric character board used in this study was 1 second and recognition of the caregiver's utterance of "confirm" took 0.2

Table 3 Total communication time for each process number

Process (N)	Time (s)	Prototype support system	Mouth-shape character method
2	mean	1.67	1.66
	SD	0.09	0.22
3	mean	2.35	2.37
	SD	0.19	0.23
4	mean	3.15	3.22
	SD	0.28	0.57
5	mean	4.13	4.43
	SD	0.20	0.95
6	mean	5.13	4.53
	SD	0.27	0.58
7	mean	5.86	5.62
	SD	0.28	1.19
8	mean	6.83	6.56
	SD	0.29	0.91
9	mean	8.21	7.27
	SD	0.40	1.12
10	mean	9.68	8.07
	SD	2.19	0.71
11	mean	9.98	8.65
	SD	1.51	1.20
12	mean	10.81	9.68
	SD	0.70	1.09

Fig. 9 Process number (N) and communication time (s) using the prototype support system

https://doi.org/10.31871/IJNTR.5.11.17

International Journal of New Technology and Research (IJNTR)
ISSN: 2454-4116, Volume-5, Issue-11, November 2019 Pages 26-31

Fig. 10 Process number (N) and communication time (s) using the mouth-shape character method

seconds, but the patient expressed his/her intention in the first stage of the vocalization of the board, and the average time per process was less than 1 second. In particular, when the process number was 10, the average communication time was 0.97 seconds, which could be attributed to the delay in confirming the caregiver's expression of intention.
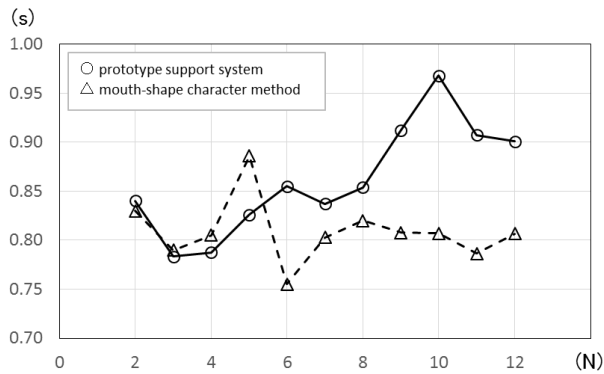

Fig. 11 Average time per process

The communication time using the mouse-shape character method (Fig. 10) also increased linearly as the process number increased. The average communication time per process (Fig. 11) was about 0.8 seconds. As the process number increased, the communication time became faster than the proto-system. In the mouth-shape character method, the patient expressed his/her intention at the early stage of the caregiver's utterance after the patient became accustomed to the operation through repeated trials, and the caregiver read the text faster, which may have resulted in faster communication than the proto-system.

B. Operation of prototype support system

The operation of the proto-system was examined from the failure contents, frequency of the communication and subjective view of the subjects. No error occurred for the speech recognition of the caregiver's "confirm" voice, because the registered voice was only one type and was not disturbed by the electronic sounds emitted from the electric character board. All six failures in the proto-system were due to delays in the expression of the patient's intention or in the action of the caregiver's confirmation of characters. Although the speed set in the proto-system (1 second/scan) was the standard setting for an electronic character board, both the patient and the caregiver caused a delay in confirmation, indicating that this setting was slightly too fast to communicate with voice alone without seeing any characters on the electronic character board. Optimization to set the fastest scan speed with minimal failure may require the caregiver to set the speed to match the patient's and caregiver's proficiency in use. Meanwhile, since the confirmed characters were recorded in the system, it was indicated that the system could prevent the failure caused by caregivers forgetting the confirmed characters.

Among the subjects, the subjective opinions as the caregiver roles were that they did not have to utter by themselves or memorize the confirmed characters. Therefore, they were less tired than when they were using the mouth-shape character method and could focus on communication with the patient. Furthermore, the caregivers were able to use both hands because they did not need tools for communication such as character boards, switches and memo pads.

Thus, although the speed of communication in the proto-system was slightly inferior to that in the conventional mouth-shape character method, the prototype system had the ability equivalent to the communication via the mouth-shape character method. These results suggest that speech utterances and character recordings using the proto-system have the capacity to reduce the burden of the communication support for the caregiver. In addition, the caregiving workload may be lowered by allowing him/her to use both hands freely.

C. Limitations of the study

There are limitations in this study. The subjects were healthy individuals. Hiragana characters employed in the trials did not cover the whole 46 characters, but random combinations were chosen. In this speech recognition environment, there were no voice outputs except for the speech utterances from the caregiver and the electronic character board.

V. SUMMARY

We developed the prototype support system for speech utterances and character recordings of caregivers, which is useful for communication between ALS patients and caregivers through the mouth-shape character method. In this system, an electronic character board was used for speech utterances and character recordings, and a speech recognition module was used for conforming characters. Although the speed of communication in this system was slightly inferior to that of the conventional mouth-shape character method, the accuracy of recording of confirmed characters was high, and the possibility of reducing the burden on caregivers was obtained. By using speech recognition as an input trigger, both hands of the caregiver were free, which also showed the possible reduction of the caregiving workload.

In the future, we will optimize the scan speed of communication, increase the number of characters, and conduct an evaluation assuming practical use. The development of smart speakers using speech recognition technology has made it easier to operate hands-free devices [16], and we will consider using them. Because communication consists of both patients and caregivers, support systems should be developed for both communication of patient's intention and reduction of caregiver's burden [17]. We expect that the enhanced support of communication between patients and caregivers will also provide an enhanced quality of life for both parties.

REFERENCES

[1] Salameh JS, Brown RH Jr, Berry JD, "Amyotrophic Lateral Sclerosis: Review," *Semin Neurol,* vol.35(4), 2015, pp.469–76.

[2] Neto LL, Constantini AC, Chun RYS, "Communication vulnerable in patients with Amyotrophic Lateral Sclerosis: A systematic review," *NeuroRehabilitation*, vol.40 (4), 2017, pp.561–568.

[3] Linse K, Aust E, Joos M, Hermann A, "Communication Matters-Pitfalls and Promise of Hightech Communication Devices in Palliative Care of Severely Physically Disabled Patients With Amyotrophic Lateral Sclerosis," *Front Neurol,* vol.9, 2018, pp.603.

[4] Spataro R, Ciriacono M, Manno C, La Bella V, "The eye-tracking computer device for communication in amyotrophic lateral sclerosis,"*Acta Neurol Scand,* vol.130(1), 2014, pp.40–45.

[5] Sitaram R, Zhang H, Guan C, Thulasidas M, Hoshi Y, Ishikawa A, Shimizu K, Birbaumer N, "Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface," *Neuroimage*, vol.34(4), 2007, pp.1416–27.

[6] Wolpaw JR, Bedlack RS, Reda DJ, Ringer RJ, Banks PG, Vaughan TM, Heckman SM, McCane LM, Carmack CS, Winden S, McFarland DJ, Sellers EW, Shi H, Paine T, Higgins DS, Lo AC, Patwa HS, Hill KJ, Huang GD, Ruff RL, "Independent home use of a brain-computer interface by people with amyotrophic lateral sclerosis," *Neurology*, vol.91(3), 2018, pp.e258–e267.

[7] Maresca G, Pranio F, Naro A, De Luca R, Maggio MG, Scarcella I, De Domenico C, Bramanti P, Conti Nibali V, Portaro S, Calabrò RS, "Augmentative and alternative communication improves quality of life in the early stages of amyotrophic lateral sclerosis," *Funct Neurol*, vol.34(1), 2019, pp.35–43.

[8] Toshiaki Tamura, Mitsuhira Iiguchi, Shinichi Noto, ""Hitorigoto" notes and a special kana characters board improved quality of life in a patient with Shy-Drager syndrome at the end of life", *JJSLandHR*, vol.12 (2), 2015, pp.87–93.

[9] Takashi Nakajima, "Assistive technology for supporting communication for patients with incurable and progressive neuromuscular diseases, including transparent character boards, a mouth-shape character method, and an advanced Cybernic Interface device," *J. Natl. Inst. Public Health*, vol.66 (5), 2017, pp.491–496.

[10] Tomoya Miyasaka, Mamiko Kamoshida, Norio Kato, Keiko Kawashima, "Comparison of Simple Communication Aids between Amyotrophic Lateral Sclerosis Patients and Caregivers," *IJNTR*, vol.5(9), 2019, pp.37–43.

[11] Maki Iinuma, Takeshi Saitoh, "Text input system using mouth shape pattern and vowel-based word," *IEICE Technical Report*, vol.115(100), 2015, pp.143–148.

[12] Kazuyuki Itoh, "A Study of Text Preservation Tool in the Vowel Mouse Shape and Scanning Communication Method for Persons with Severe Physical Disabilities," *IEICE Technical Report*, vol.118(270), 2018, pp.61–66.

[13] Anuar bin Mohamed Kassim, Takashi Yasuno, Hazriq Izzuan Jaafar, Mohd Shahrieel Mohd Aras, "Development and Evaluation of Voice Recognition Input Technology in Navigation System for Blind Person," *Journal of Signal Processing,* vol.19(4), 2015, pp.135–138.

[14] Hirotaka Kobayashi, "Communication Aids for Amyotrophic Lateral Sclerosis", *Jpn J Rehabili Med*, Vol.55(7), 2018, pp.564–572.

[15] RayTron voice recognition module http://www.raytron.co.jp/products/voice_module

[16] Ikkaku Kawaguchi, Hideaki Kuzuoka, Donald McMillan, "The Effect of Interaction Using Gaze Input / Output on Smart Speaker," *The Transactions of Human Interface Society,* vol.21(3), 2019, pp.269–278.

[17] Galvin M, Corr B, Madden C, Mays I, McQuillan R, Timonen V, Staines A, Hardiman O, "Caregiving in ALS - a mixed methods approach to the study of Burden," *BMC Palliat Care*, vol.15(1):81, 2016, pp.1–12.