Flexible bioelectronics

Silent speech interface

Oriane Peter, Philipp Spiess, Jan Frogg, Mathieu Vautey

Professor: S. Lacour
Supervisor: Florian Falleger

École Polytechnique Fédérale de Lausanne

December 23, 2018

Contents

1 Summary

Loss of speaking ability is a side effect to some radical surgery such as total laryngectomy. As the muscles involved remain intact, speech recognition by surface electromyography (EMG) could help compensate for this loss. Several systems using such technology have already been implemented. The biggest challenge is the creation of a model that yields reasonable recognition rates and which is based on a vocabulary large enough for basic everyday use. While advances in software has been steadily occuring, the hardware in all the proposed device remains very noticable and rigid. Such characteristics lead to uncomfort and potential stigmatisation of the patient.

Our device proposes a discreet and flexible solution to speech detection facial masks. It consists of both an earpiece and a disposable patch which are to be worn on one side of the patient's face. Both flexibility and discretion characteristics can be achieved through the use of Silicone rubber Polydimethylsiloxane (PDMS) for the patch, as it is a material with high critical strain and with a possibility for transparency. Critical strain of the patch is of 0,5 and minimum radius of curvature of 333 μm. The patch is 200 μm thick and about 15 cm long. Electrodes are made of a mixture of PDMS and silver particles to obtain a conductivity of 600 S/cm. Skin compliance is attained by pillar-shaped microstructures with flat tip, which ensures a bond of 0,1 Ncm^2 while not irritating the skin, even after a 10 hours use. This shape is given to conductive and non-conductive PDMS parts. On the electrodes it also gives rise to an impedance of 125 $k\Omega$, allowing for a clear recording of the muscle activity. As the lifetime expectancy of the electrode is limited the patch can be changed in-between uses and is connected to the earpiece using a magnet.

In term of feasibility, material and techniques used have proven their efficiency. The microstructured electrodes have been shown to yield acceptable measurements and hold comfortably on the user's skin. Manufacturing techniques used are also quite mainstream and should produce the expected results in a reasonable amount of time. Limitations lie mainly on the software aspect of the system which has not been developed here. Lastly, it would be relevant to note that this being a device which aims for medical patients, it must meet a lot of requirements in terms of safety for the patient of course, but also in terms of financial sustainability and regulation. It is subject to a large amount of examinations and tests, especially if we choose to export it to other countries, which all have their own laws and regulatory bodies.

2 State-of-the-art Review

The latest American Cancer Society estimates for laryngeal cancer in 2018 in the United States reports approximately 13'150 new cases and over 150'000 cases of laryngeal cancer are diagnosed each year in the world. To treat advanced laryngeal cancer radical surgery such as total laryngectomy is required. This involves removing the larynx, thereby eliminating the natural sound source for speech production. Patients undergoing such surgery therefore often lose their speaking ability. A study over 10 years in the United State showed that each year from 3000 to 5000 patients underwent such a procedure [1].

To compensate for this loss systems have been developed which infer and eventually retransmit intended speech. One of the relevant approaches is surface electromyography (sEMG) targeting specific facial muscles involved in words production. Several such research have already been performed with varying degrees of success.

In 2001, [2] obtained a 93% recognition rate on a vocabulary of digits from zero through nine. The vocabulary set was increased by [3], which achieved a mean 87% recognition rate on 60 vocalized words for 8 participants.

A noticeable advancement was attained by [4], working towards continuous speech recognition and ultimately achieving a Word Error Rate (WER) of 15.7% on a vocabulary of 108 words.They showed that a system trained on multiple recording sessions of one and the same speaker yields a reasonable performance, which is crucial to bring such device to a usable state.

[5] produced a portable silent speech recognition system called MUTE (Mouthed-speech Understanding and Transcription Engine). They focused on speaker dependent, unlimited vocabulary and continuous recording for military communication applications. Their results showed 87% accuracy for a vocabulary of 65 words. They targeted 4 muscles on the face and neck with a set of 2 electrodes for each. The electrodes used were custom made and it is not stated weather they required a gel application. They were connected to a wireless transmitter device using wires (Fig. 1 and 2). The raw data was transmitted to a mobile phone via a wireless transmitter module. The mobile phone carried the processing of the data (Fig. 2).

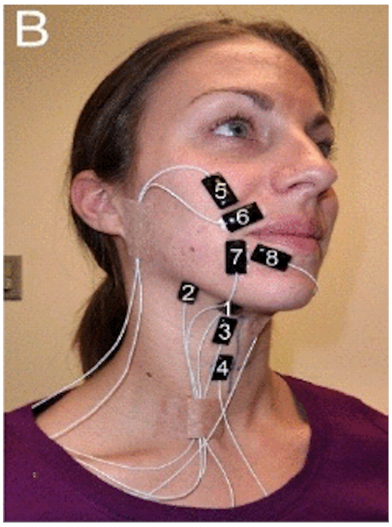

Figure 1: sEMG sensor locations. Sensors are on the submental neck (1-2), ventromedial neck (3-4), supralabial face (5-6) and infralabial face (7-8).

Figure 2: (1) four pairs of sEMG sensors custom designed; (2) a wireless sEMG transmitter module that transmits the raw sEMG data; (3) a wireless receiver module that receives the raw sEMG signals, and (4) an Android smartphone or tablet that acquires the raw data from the wireless receiver via USB, processes the data and ultimately conducts the recognition.

More recently a wearable interface called AlterEgo has been developed by [6]. They reported a 92% median word accuracy on a ~ 60 word vocabulary. This system exhibited robust accuracy even without deliberate muscle articulation. The signal was captured using passive dry Ag/AgCl electrodes. A reference electrode was placed on the wrist or on the earlobe. Sampling and amplification were performed on the device before the signal was wirelessly retransmitted to a external computing device for further processing.

The device was designed as a wearable worn around the back of the head which carry rigid arms landing on the targeted muscle on the face (Fig. 3). The arms rigidly support the electrodes but can be deliberately adjusted to fit different users.

It is to our knowledge the only one having been designed for comfort and practicality. However, it is quite bulky and noticeable. Rigid arms are likely to entrave the movement of the skin to some degree and to cause inconveniences. It is also important to consider that visible medical products can elicit a potential stigmatisation and emotional impact on the user and bystander. Therefore, a discreet device would be preferable.

Figure 3: Rendering of the AlterEgo wearable (Top). Front view of the user wearing the device (Bottom).

3 Device Description

3.1 Design

3.1.1 Layout

The device consists of two elements: an earpiece and an adhesive patch. The patch is about 200 μm thick and contains 13 electrodes (Fig. 4 and 6). To obtain a comfortable and discreet system we aimed for a transparent, thin and self-adhesive device in PDMS. The patch is composed of two main branches going respectively to mouth and the neck area targeting specific muscles.

The speech production mechanism involves facial muscle activity that leads to buccal and laryngeal movements. Studies of speech mechanism indicates that pertinent target muscles for the recording of the EMG be : the zygomaticus, depressor anguli oris, levator anguli oris, platysma, anterior belly of the digastric and orbicularis oris [7]. Lastly, the device comes with a smartphone app, to which it outsources the data processing upon signal reception via Bluetooth. The reason for doing this is that the device can take full advantage of the smartphone computational power, as speech recognition is quite demanding. Recording of the EMG signals starts upon powering on the device and stops when turned off.

Figure 4: 3D Modelisation of the mask

Figure 5: Earpiece-to-Patch connection

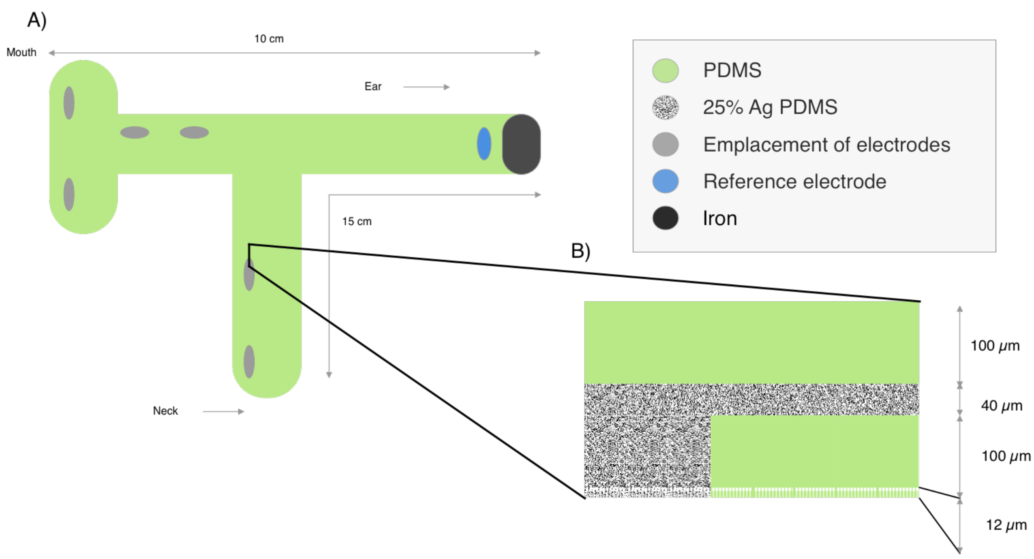

Figure 6: A) General layout of the patch with approximate emplacement of the electrodes (each grey oval is 2 electrodes) B) Cross section of the patch at intersection between electrode and non-conductive PDMS

3.1.2 Microstructures

Adhesion to the skin is achieved using pillar shaped microstructures with flat tip on the whole of the patch (Fig. 7). This adhesion technique is inspired from the Gecko and works by increasing the contact surface and therefore Van der Waals forces. This increase in surface is also beneficial for the electrode because it allows an increase in the capacitance of the electrodes.

[8] have shown that such an adhesive structure can stay for 10 hours without irritating the skin and can be attached and detached more than 20 times without causing any damage. The same structure is used on the whole of the patch, the electrodes are differentiated from the rest by the addition of silver particles in the PDMS.

The adhesion force will be of 0,1 Ncm^2 in accordance with the result of [8]. To reach this adhesion force, the microstructures need to have the following dimensions: base pillar : 10 μm height and 5 μm diameter. tip: of 2 μm height and 7 μm diameter. This allows for a very dense surface coverage.

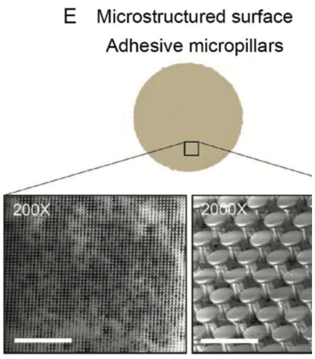

Figure 7: Microstructured surface - Adhesive micropillars for improved dry adhesion to the skin (thickness = 100 μm). The scale bars are 200 μm (200X) and 20 μm (2000X) [8]

3.2 Component

3.2.1 Electrodes

For each targeted muscle two electrodes are placed on its side to record electrical activity. As 6 muscles are targeted a total of 12 + 1 references electrodes will be used. They must at least have a 5mm distance between one another in order to avoid interferences [9]. The reference electrode is placed on the mastoid bone behind the ear where the earpieces is located. An electrode consists of a mix of approximately 25% of silver particles and PDMS, allowing an electrical conductivity of more than 100 $S.cm^{-1}$. The diameter of each electrode is of 0,5 cm. The microstructured shape of the electrode ensures a low impedance: at 100 Hz which is in the range of frequency of the measured signal, normalized impedance is of approximately 25 $k\Omega.cm^2$ [8]. For an electrode of 0,2 cm^2, the impedance is therefore of 125 $k\Omega$ which is less than a classical gel electrode as shown in [8].

3.2.2 Tracks

Electrodes are connected to the earpiece by Ag-PDMS composite tracks. All values for the tracks' design are taken from [10]. The filler volume fraction is of 25% Ag and allows for an elastic modulus of 1 MPa, which is close to the skin's elastic modulus (0.85 MPa in torsion [11]) and thus allows a good compliance. It also allows for an electrical conductivity of 600 S/cm which is enough for the intended purpose. The tracks are 1mm wide with a 40 μm thickness imposed by the metal mesh used in manufacturing. The resistance of the tracks was estimated as follows : The largest distance between an electrode and the ear is approximately 15 cm. At vol 25% of Ag the mean sheet resistance if 0.05 $[\Omega/sq]$. The resistance is therefore at most : R = Rs * L / w = 0.05 * 15*10^{-2} / 10^{-3} = 7,5 with w the width, L the maximum distance, Rs mean sheet resistance. This resistance is smaller than the impedance of the electrode. This goes to show that the noise of the system will not be dominated by the impedance of the tracks.

3.2.3 Earpiece

The earpiece houses all of the active electronics of the device. It has a size of 5x2x1 cm^2. The first stages of processing consist of operational amplifiers for signal acquisition and processing purposes. Once the signal is 'clean' and ready to be used, it is sent to the smartphone app via Bluetooth.

Outsourcing the data processing allows us to use fewer components inside the earpiece and especially it allows for low power consumption. There is only a need for signal analysis components, an antenna and a battery. Using existing Ultra-low-power operational amplifiers to perform operations on the signal, we would need a power supply ranging from 1.7 V - 3.6 V [12]. Batteries used in hearing aid products or wireless headphones can provide this in a minimalistic manner. To that effect, the device is fitted with an 2 x 2 array of such batteries using a 2s2p configuration -2 series, 2 parallel- in order to ensure that the voltage is high enough, as well as to improve battery life. Typical voltage values for said batteries is around 1.4 V and typical capacity is around 120 mWh ([13],[14],[15]). 2s2p configuration would result in a 21.4 V = 2.8 V power supply and a 2120 mWh = 240 mWh capacity, this second value gives us an approximate on the battery life. Lastly, dimensions for one battery cell of this type are: 7.9 x 3.6 mm, which means our battery will be twice this size: Battery Size = 7.93.6 mm * 2 = 15.87.2 mm.

3.2.4 Earpiece-to-Patch connection

The lifetime of the electrodes being limited, the device requires frequent replacement of the patch. Therefore it is necessary to implement a system allowing for an easy connection and disconnection of the patch and the earpiece. To this effect, the connection is done with a magnet and an iron thin film [2] (Fig. 8). The iron film is deposited around the track on the patch to ensure contact with the earpiece's tracks. A strip ending with a magnet will come out of the earpiece and make the connection to the end of the patch. As the tracks and the iron film are 40 μm and 160 μm thick respectively, the tracks in the earpiece will need to be at least 120 μm thick for the connection to be possible. The end of the copper track of the earpiece will have a pyramidal shape to ensure that even with relatively low pressure contact is correctly made. The magnet on the earpiece is just a regular ferrite magnet with a magnetic field of 0.2 T.

For a surface of 1,6 cm^2, this means a normal attraction force of F = B^2 * s / (2 * μ_0)= 2,04 N between the magnet and the iron thin film. This allows for a good adhesion without requiring too much strength to separate the patch from the earpiece as shear force is lower.

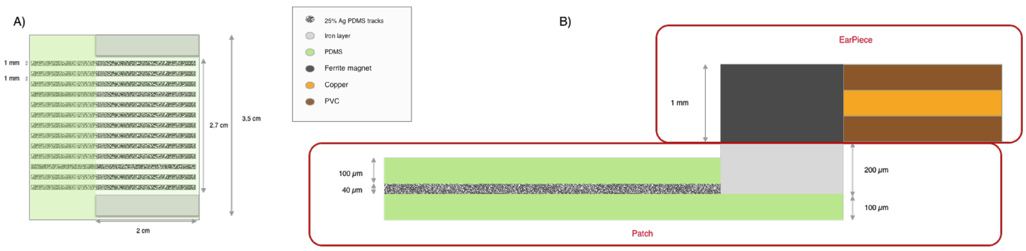

Figure 8: A) Top view and B) Cross-section of the interface between the patch and the earpiece (earpiece and patch are not to scale)

3.3 Stretchability

Strain at rupture is of 50% for the 25% filler fraction tracks and 50% for classical PDMS ([10], [16]). Stretchability is therefore well above what should be reached in a normal use of the patch. It should also be taken into account that when passing over the jaw, the patch will undergo a relatively large bending. By computing the minimum radius the patch can withhold using the smallest critical strain, we see that they can go over the jaw without any issue:

r = h / (2 * ϵ) = 200 * 10^{-6} / (2 * 0.3) = 333 μm with ϵ the critical strain of the PDMS, here about 50%, and h the thickness of the patch.

3.4 Energy Consumption

Seeing as the system uses active electronics only for signal recording, we may assume a maximum current of 1mA (typical drain current value for a hearing aid [15]).

Battery Life = CapacityDrain current = 240 mWh 1 mA 240 hours = 20 days.

Considering a 12 hours per day usage and 240 mWh Battery capacity.

3.5 Signal processing

Fully differential EMG signal is obtained in the earpiece by comparing first the two signals from one muscle and then the resulting signal to the reference electrode. This signal is then sampled at 600 Hz and a high-pass filter is applied with cutoff at 60 Hz. This allows removal of the ambient noise (60 Hz) and the motion artefact (0-20 Hz) [5]. Thermal noise would be of about 4.5 V for a bandwidth of 100 Hz to 10 kHz and absolute temperature of 20°C and can easily be ignored. Feature extraction is then performed based on time domain feature and recognized using fully continuous Hidden Markov Chain Model (Fig. 9).

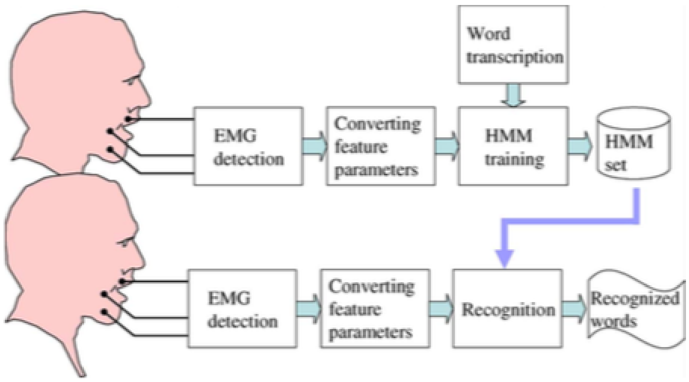

Figure 9: Possible process flow for Hidden Markov Chain Model [3]

3.6 Long term stability

The patch has been shown to rest comfortably on the skin for up to 10 hours without causing irritation. All electronics in the earpiece are stable and the battery can last for 20 days and can easily be changed after that. As the patch is constantly changed, long term stability of the device is guaranteed.

3.7 Usage

At first use the device requires training, to create the most reliable model possible. Training consist of the lecture of several sentences and can take up to several hours [5]. Following uses require the application of the patch using a gypsum mask, which is tailored to the user's face and records the patch positioning during the training process. A hole is cut in a rigid structure to keep track of the location of the patch during the first session and future patches can simply be applied by filling this hole. Electrodes positioning can vary the predicting ability of the model and the use of the gypsum mask therefore clearly increases both accuracy and precision of the results [17]. The user must turn on the earpiece before speaking and open the corresponding smartphone application and turn it off when done. This allows for retranscription of speech only when speech is intended and not during auxiliary face movement. The smartphone microphone will be used for speech retransmission. After each use the patch is removed and a new one is used.

4 Feasibility Assessment

4.1 Manufacturing

Step n°	Operation	Deposited Thickness	Equipment
1	Spin coating of a lift-off resist LOR 10B on a Silicon Wafer (Si), previously coated with Polyvinyl Alcohol (PVA) 15 μm	5 μm	Z13/SSE
2	Spin coating of another resist AZ4562 on top of the previous one	2 μm	Z13/SSE
3	Structuration of another resists with Photolithography	7 μm	Suss MjB4
4	Screen printing of the conductive Ag-PDMS composite silicone layer in the resist mold using a Polymethylmethacrylate (PMMA) screen to determine frontiers. PMMA stencil mask is a laser cut to have the correct shape [10]	100 μm	THIEME1010
5	PMMA screen removal		Manually
6	Printing of the non-conductive PDMS layer in the mold around the electrode and everywhere else, using direct light-based 3D printing method which allow to a 100 nm lateral resolution [18]	100 μm	Aerotech
7	Dissolution of LOR 10B and AZ4562 photoresists in Acetone		Acetone
8	Iron (Fe) film deposition on a 1 cm^2 bit of surface on the very edge of the patch. Deposition is done by thermal evaporation at a 15 nm/s approximate rate	160 μm	Moorfield M307
9	Ag-PDMS tracks deposition done by screen printing using a metal mesh	40 μm	THIEME 1010
10	Tracks-to-electrode connection following plasma oxygen treatment		PE-25
11	PDMS layer deposition done by screen printing to encapsulate and protect the whole device against oxydation	100 μm	THIEME 1010
12	Patch release from wafer using water bath to remove PVA		H2O

Table 1: Fabrication process flow

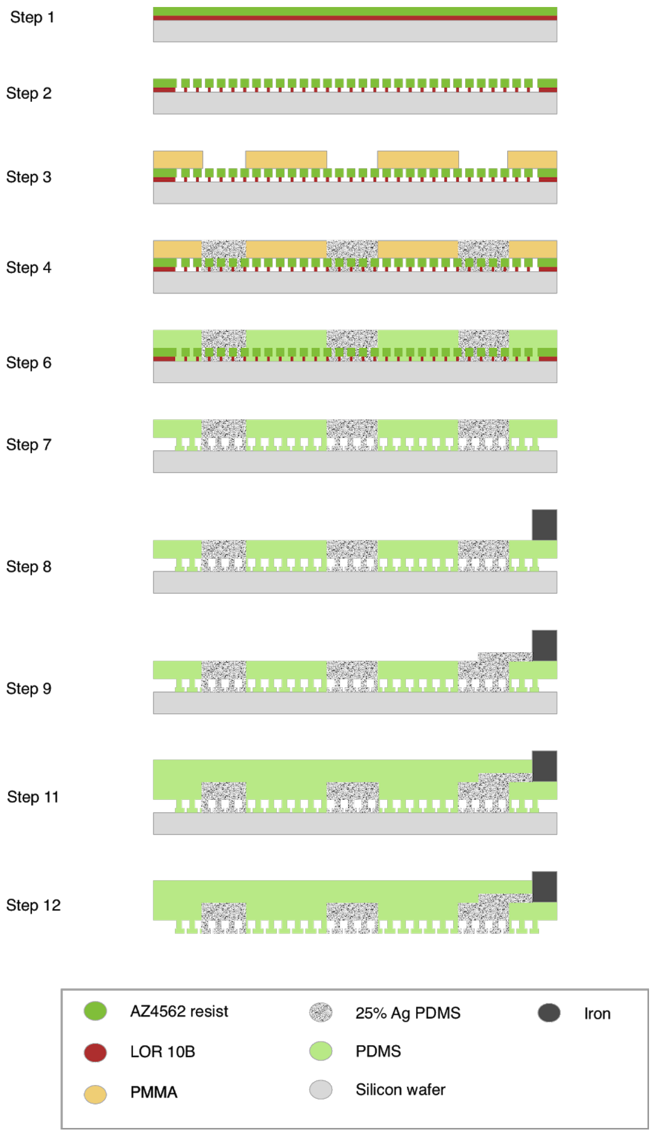

Figure 10: Possible process flow for Hidden Markov Chain Model [7]

4.2 Regulation

Serving a medical purpose, our device will need approval from several regulatory bodies, such as the European Commission Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs for Europe and Switzerland or the U.S. Food & Drug Administration (FDA) in the USA, depending on its worldwide application. As it has naturally been imposed by the medical sector, our device must face several challenges. It must be competitive and innovative, find a certain balance between patient needs and financial sustainability and lastly it must fit the public health system.

4.3 Future improvement

Overall, and especially from a fabrication point of view, the device along with its electronics is realistic. The biggest setback of the system lies in the limited amount of words recognized by the device and the high WER of about 20%. Those are mainly software dependent and might see an improvement with the creation of larger database and the perfectioning on data analysis system. Another enhancement could be in the integrability of the patch. By using antenna the size could be reduced and the system could be more discreet. Here we decided not to use this system as integrating a battery to the patch would increase the weight and size of the device reducing the overall comfort. We also explored batteryless systems, such as passive RFID, but the power supply at hand was around 1000 fold smaller than needed. Lastly, it is important to note that our device implicates an Ag-skin contact, which can cause allergies.

5 Conclusion

The device as proposed is feasible. PDMS and Ag-PDMS composites are widely used and outcomes from different studies have shown that this composite can be used as sEMG electrode. Compared to existing technologie, the biggest improvement of our device lies in its discretion and flexibility. Due to the complex shape of the patch, good adhesion and flexibility are essential to follow the movements of the face without entraving normal activity. The use of Ag-PDMS and near-transparent PDMS brings discretion and comfort as the elastic modulus is close to that of the skin's. Adhesion is managed by a micro-pillar structure mimicking the Gecko's setae which allows for an increase in the Van der Waals forces due to a larger contact surface. Most limitation of the device come from signal processing. Results in modelisation of sub-vocalised speech based on EMG signal are improving and may, in the future, reach the optimal goal of a fluent and real-time retranscription of the speech. Lastly, we believe that our device not only could impact the medical sector but could also benefit healthy people who could use the device for two other contexts: the need to talk in a loud environment or in an environment where one can not afford to make vocal noise: for instance, a person who would like to make a phone call while driving his motorcycle on the highway or a person in a conference who would like to take notes of what the speaker in saying. This device could in fact revolutionize the way people take notes in schools, universities or conferences. One could just tell the device what to note in a voiceless way.

References

[1] Patrick Tate Maddox and Louise Davies. Trends in Total Laryngectomy in the Era of Organ Preservation: A Population-Based Study. *OtolaryngologyâĂŞHead and Neck Surgery*, 147(1):85–90, July 2012.

[2] A. D. C. Chan, K. Englehart, B. Hudgins, and D. F. Lovely. Myo-electric signals to augment speech recognition. *Medical & Biological Engineering & Computing*, 39(4):500–504, July 2001.

[3] K. Lee. EMG-Based Speech Recognition Using Hidden Markov Models With Global Control Variables. *IEEE Transactions on Biomedical Engineering*, 55(3):930–940, March 2008.

[4] Michael Wand and Tanja Schultz. SESSION-INDEPENDENT EMG-BASED SPEECH RECOGNITION. page 7.

[5] Matthias Janke and Lorenz Diener. EMG-to-Speech: Direct Generation of Speech From Facial Electromyographic Signals. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 25(12):2375–2385, December 2017.

[6] Arnav Kapur, Shreyas Kapur, and Pattie Maes. AlterEgo: A Personalized Wearable Silent Speech Interface. In *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval - IUI '18*, pages 43–53, Tokyo, Japan, 2018. ACM Press.

[7] A surface EMG electrode for the simultaneous observation of multiple facial muscles - ScienceDirect.

[8] Flurin Stauffer, Moritz Thielen, Christina Sauter, Severine Chardonnens, Simon Bachmann, Klas Tybrandt, Christian Peters, Christofer Hierold, and Janos Voros. Skin Conformal Polymer Electrodes for Clinical ECG and EEG Recordings. *Advanced Healthcare Materials*, 7(7):1700994, 2018.

[9] Roberto Merletti, Matteo Aventaggiato, Alberto Botter, Ales Holobar, Hamid Marateb, and Taian M. M. Vieira. Advances in Surface EMG: Recent Progress in Detection and Processing Techniques. *Critical Reviews™ in Biomedical Engineering*, 38(4), 2010.

[10] Alexandre Larmagnac, Samuel Eggenberger, Hanna Janossy, and Janos Voros. Stretchable electronics based on Ag-PDMS composites. *Scientific Reports*, 4:7254, December 2014.

[11] Mariola Pawlaczyk, Monika Lelonkiewicz, and Michal Wieczorowski. Age-dependent biomechanical properties of the skin. *Advances in Dermatology and Allergology Postepy Dermatologi i Alergologi*, 30(5):302–306, October 2013.

[12] TLV8542 datasheet | TI.com.

[13] Danielle Glista Vijay Parsa Richard Marlene Bagatto Seewald, Susan Scollie. Case Study Outcomes of Hearing Impaired Listeners Using Nonlinear Frequency Compression Technology Marlene Bagatto Susan Scollie Danielle Glista Vijay Parsa Richard Seewald.

[14] Basic to Advanced Battery Information from Battery University.

[15] Hearing Tracker: Consumer Reviews for Hearing Aids.

[16] Dow Corning. Sylgard 184 Silicone Elastomer. (11):4, 2014.

[17] L. Maier-Hein, F. Metze, T. Schultz, and A. Waibel. Session independent non-audible speech recognition using surface electromyography. In *IEEE Workshop on Automatic Speech Recognition and Understanding, 2005.*, pages 331–336, November 2005.

[18] Ryan L. Truby and Jennifer A. Lewis. Printing soft matter in three dimensions. *Nature*, 540(7633):371–378, December 2016.